LINEAR COMPLEMENTARITY, LINEAR AND NONLINEAR PROGRAMMING

Internet Edition

Katta G. Murty
Department of Indutrial and Operations Engineering
University of Michigan, Ann Arbor

The Internet edition of this book has been prepared by

Feng-Tien Yu
Department of Indutrial and Operations Engineering
University of Michigan, Ann Arbor

Partial funding for presenting this book on the web has been provided by the Spring/Summer Research Grants Program Office of the Vice President for Research Horace H. Rackham School of Graduate Studies The University of Michigan, Ann Arbor

Katta G. Murty
Dept. of Industrial and Operations Engineering
The University of Michigan
1205 Beal Avenue
Ann Arbor, MI 48109-2117, U. S. A.
TEL: +1 734 763-3513
E-mail: murty@umich.edu
WWW: http://www-personal.engin.umich.edu/~ murty/

Feng-Tien Yu
Dept. of Industrial and Operations Engineering
The University of Michigan
1205 Beal Avenue
Ann Arbor, MI 48109-2117, U. S. A.
TEL: +1 734 764-8583
E-mail: fengtien@umich.edu
WWW: http://www-personal.umich.edu/~ fengtien/

Other books by Katta G. Murty :

—– **Linear programming, John Wiley & Sons, 1983**

—– **Linear and combinatorial programming, R. E. Krieger, 1985**

—– **Network programming, Prentice Hall, 1992**

—– **Operation research: deterministic optimization models, Prentice Hall, 1995**

PREFACE

INTRODUCTION

I am grateful for the enthusiastic reception given to my book *Linear and Combinatorial Programming* published in 1976. Many readers from all over the world commented that they liked Chapter 16 on the Linear Complementarity Problem (LCP) in this book, but found it too brief, and suggested that a new up-to-date book devoted exclusively to this topic, covering all aspects of linear complementarity would be worthwhile. This book is the result of the encouragement I have received from all these suggestions.

An important class of applications for the LCP stems from the fact that the necessary optimality conditions for a Quadratic Programming Problem (QP) lead to an LCP. Until recently, a practitioner of mathematical programming could have brushed off QP as an academically interesting generalization of linear programming which is not very useful. But the recent development of recursive quadratic programming methods for solving Nonlinear Programming Problems (NLP) has changed all that. These methods solve an NLP through a sequence of quadratic approximations, and have become extremely popular. They have suddenly made QP and thereby LCP an important topic in mathematical programming with a large number of practical applications. Because of this, the study of LCP is attracting a great deal of attention both in academic curricula and in the training of practitioners.

THE OBJECTIVES

1. To provide an in-depth and clear treatment of all the important practical, technical, computational, geometric, and mathematical aspects of the LCP, QP, and their various applications.
2. To discuss clearly the various algorithms for solving the LCP, to present their efficient implementation for the Computer, and to discuss their computational complexity.
3. To present the practical applications of these algorithms and extensions of these algorithms to solve general nonlinear programming problems.
4. To survey new methods for solving linear programs, proposed subsequently to the publication of [2.26].

BACKGROUND NEEDED

The background required to study this book is some familiarity with matrix algebra and linear programming (LP). The basics of LP are reviewed in Chapters 1 and 2.

SUMMARY OF CHAPTER CONTENTS

The book begins with a section titled 'notation' in which all the symbols and several terms are defined. It is strongly recommended that the reader peruse this section first at initial reading, and refer to it whenever there is a question about the meaning of some symbol or term.

Chapter 1 presents a clear geometric interpretation of the LCP through the definition of the system of complementary cones as a generalization of the set of orthants in $\mathbf{R}^n$. Applications to LP, QP, and nonzero sum game problems are discussed. There is a complete discussion of positive definiteness and positive semidefiniteness of square matrices, their relationship to convexity, together with efficient pivotal methods for checking whether these properties hold for a given matrix. Various applications of QP are discussed, as well as the recursive quadratic programming method for solving NLP models.

Chapter 2 presents a complete discussion of the many variants of the complementary pivot method and proofs of its convergence on different classes of LCPs. Section 2.7 contains a very complete, lucid, but elementary treatment of the extensions of the complementary pivot method to simplicial methods for computing fixed points using triangulations of $\mathbf{R}^n$, and various applications of these methods to solve a variety of general NLP models and nonlinear complementarity problems.

Chapter 3 covers most of the theoretical properties of the LCP. There is extensive treatment of the various separation properties in the class of complementary cones, and a complete discussion of principal pivot transforms of matrices. In this chapter we also discuss the various classes of matrices that arise in the study of the LCP. Chapter 4 provides a survey of various principal pivoting methods for solving the LCP. Algorithms for parametric LCP are presented in Chapter 5.

Chapter 6 contains results on the worst case computational complexity of the complementary and the principal pivoting methods for the LCP. Chapter 7 presents a special algorithm for the LCP associated with positive definite symmetric matrices, based on orthogonal projections, which turned out to be very efficient in computational tests. Chapter 8 presents the polynomially bounded ellipsoid methods for solving LCPs associated with positive semidefinite matrices, or equivalently convex QPs.

Chapter 9 presents various iterative methods for LCPs. In Chapter 10 we present an extensive survey of various descent methods for unconstrained and linearly constrained minimization problems; these techniques provide alternative methods for solving quadratic programming problems. In Chapter 11 we discuss some of the newer algorithms proposed for solving linear programming problems and their possible extensions to solve LCPs, and we discuss several unsolved research problems in linear complementarity.

To make the book self-contained, in the appendix we provide a complete treatment of theorems of alternatives for linear systems, properties of convex functions and convex sets, and various optimality conditions for nonlinear programming problems.

EXERCISES

Each chapter contains a wealth of various types of exercises. References are provided for theoretical exercises constructed from published literature. A new sequence of exercise numbers begins with each chapter (e.g. Exercise 3.2 refers to Exercise number 2 of Chapter 3).

HOW TO USE THE BOOK IN A COURSE

This book is ideally suited for first year graduate level courses in Mathematical Programming. For teaching a course in nonlinear programming, the best order for presenting the material may be the following: Section 10.1 (formulation example), 10.2 (types of solutions in NLP), 10.3 (types of nonlinear programs and what can and cannot be done efficiently by existing methods), 10.4 (can we at least compute a local minimum efficiently), 10.5 (precision in computation), 10.6 (rates of convergence), Appendix (theorems of alternatives for linear systems of constraints; convex sets and separating hyperplane theorems; convex, concave functions and their properties; optimality conditions), Chapters 1 to 9 in serial order; remaining portions of Chapter 10; and some supplemental material on algorithms for solving nonlinearly constrained problems like the GRG, penalty and barrier methods, and augmented Lagrangian methods. For teaching a course in linear complementarity using the book, it is best to cover the Appendix first, and then go through Chapters 1 to 10 in serial order.

The material contained in Chapters 12, 14, 15, 16 of [2.26] can be combined with that in Appendices 1, 2, Chapter 9 and Section 11.4 of this book to teach an advanced course in linear programming.

Since the book is so complete and comprehensive, it should prove very useful for researchers in LCP, and practitioners using LCP and nonlinear programming in applied work.

ACKNOWLEDGEMENTS

In preparing this book, I have received encouragement, advice, and suggestions from several people. For this I am indebted to Jeff Alden, Ahmet Bolat, R. Chandrasekaran, Akli Gana, Greg Goulding, Philip Jones, Ikuyo Kaneko, Olvi Mangasarian, Chuck Noon, Steve Pollock, Romesh Saigal, Richard Stone, Robert Smith, Vasant Ubhaya, Gary Waissi, Kai Yang and Bill Ziemba; and to David Gale who first introduced me to linear complementarity. Some of the research appearing in this book was supported partially by National Science Foundation under grants ECS-8401081 and ECS-8521183, this support is gratefully acknowledged. I thank Ivajean Marzano and Tana Beard for typing most of the manuscript. My thanks to Bala S. Guthy for drawing many of the figures with curves in them; and to K. T. Anantha for some of the other figures. Finally, I thank my wife Vijaya Katta, to whom this book is dedicated, for her patience and understanding.

CONTENTS

NOTATION

Superscript T	Denotes transposition. A^T is the transpose of the matrix A. If x is a column vector, x^T is the same vector written as a row vector and vice versa. Column vectors are printed as transposes of row vectors to conserve space in the text.
$w,\ z$	$w = (w_1, \ldots, w_n)^T$, $z = (z_1, \ldots, z_n)^T$ are the column vectors of variables in a linear complementarity problem of order n.
(q, M)	A linear complementarity problem in wich the data is the column vector $q = (q_1, \ldots, q_n)^T$, and square matrix $M = (m_{ij})$ of order n.
$\mathbf{R}^n$	Real Euclidean n-dimensional vector space. It is the set of all **ordered** vectors $(x_1, \ldots, x_n)$, where each x_j is a real number, with the usual operations of addition and scalar multiplication defined on it.
$\simeq$	Approximately equal to.
$\lambda \to 0$	λ tends to zero.
$\lambda \to 0^+$	λ tends to zero through positive values.
J, **K**, **H**, **E**, **Z**, **U**, **P**, **A** **Γ**, **I**, **Δ**, **S**, **W**, **D**	These bold face letters usually denote sets that are defined in that section or chapter.
$\sum$	Summation sign.
$\sum(a_j : j \in \mathbf{J})$	Sum of terms a_j over j contained in the set **J**.
$\geqq, \geq, >$	Given two vectors $x = (x_j)$, $y = (y_j)$ in $\mathbf{R}^n$, $x \geqq y$ means that $x_j \geqq y_j$, that is, $x_j - y_j$ is nonnegative, for all j. $x \geq y$ means that $x \geqq y$ but $x \neq y$, that is, $x_j - y_j$ is nonnegative for all j and strictly positive for at least one j. $x > y$ means that $x_j - y_j > 0$, strictly positive, for all j. The vector x is said to be nonnegative if $x \geqq 0$, semipositive if $x \geq 0$, and positive if $x > 0$.
$A_i.$	The ith row vector of the matrix A.

$A_{.j}$ The jth column vector of the matrix A.

Superscripts We use superscripts to enumerate vectors or matrices or elements in any set. When considering a set of vectors, in $\mathbf{R}^n$, x^r may used to denote the rth vector in the set, and it will be the vector $(x_1^r, \ldots, x_n^r)^T$. In a similar manner, while considering a sequence of matrices, the symbol P^r may be used to denote the rth matrix in the sequence. Superscripts should not be confused with exponents and these are distinguished by different type styles.

Exponents In the symbol $\epsilon^{\mathbf{r}}$, r is the exponent. $\epsilon^{\mathbf{r}} = \epsilon \times \epsilon \times \ldots \times \epsilon$, where there are r ϵ's in this product. Notice the difference in type style between superscripts and exponents.

$\log_2 x$ Defined only for positive numbers x. It is the logarithm of the positive real number x, with 2 as the base (or radix).

$||x||$ Euclidean norm of a vector $x \in \mathbf{R}^n$. If $x = (x_1, \ldots, x_n)$, $||x|| = +\sqrt{x_1^{\mathbf{2}} + \ldots + x_n^{\mathbf{2}}}$.

$\lceil \alpha \rceil$ Defined only for real numbers α. It represents the smallest integer that is greater than or equal to α, and is often called the *ceiling* of α. For example $\lceil -4.3 \rceil = -4$, $\lceil 4.3 \rceil = 5$.

$\lfloor \alpha \rfloor$ Defined only for real numbers α. It represents the larg-
est integer less than or equal to α, and is often called the *floor* of α. For example $\lfloor -4.3 \rfloor = -5$, $\lfloor 4.3 \rfloor = 4$.

∞ Infinity.

$\in$ Set inclusion symbol. If $\mathbf{F}$ is a set, "$F_1 \in \mathbf{F}$" means that "F_1 is an element of $\mathbf{F}$". Also "$F_2 \notin \mathbf{F}$" means that "F_2 is not an element of $\mathbf{F}$".

$\subset$ Subset symbol. If $\mathbf{E}$, $\mathbf{\Gamma}$ are two sets, "$\mathbf{E} \subset \mathbf{\Gamma}$" means that "$\mathbf{E}$ is a subset of $\mathbf{\Gamma}$", or that "every element in $\mathbf{E}$ is also an element of $\mathbf{\Gamma}$".

$\cup$ Set union symbol. If **D**, **H** are two sets, $\mathbf{D} \cup \mathbf{H}$ is the set of all elements that are either in **D** or in **H** or in both **D** and **H**.

$\cap$ Set intersection symbol. If **D** and **H** are two sets, $\mathbf{D} \cap \mathbf{H}$ is the set of all elements that are in both **D** and **H**.

$\emptyset$ The empty set. The set containing no elements.

$\setminus$ Set difference symbol. If **D** and **H** are two sets, $\mathbf{D} \setminus \mathbf{H}$ is the set of all elements of **D** that are not in **H**.

$\{ \quad \}$ Set brackets. The notation $\{x$: some property$\}$ represents the set of all elements x, satisfying the property mentioned after the ":".

$|\mathbf{F}|$ If **F** is a set, this symbol denotes its **cardinality**, that is, the number of distinct elements in the set **F**.

e The base of the natural logarithms. $\text{e} = 1 + \sum_{n=1}^{\infty} \frac{1}{n!}$, if is approximately equal to 2.7.

e, e_r The symbol e denotes a column vector, all of whose entries are equal to 1. Its dimension is usually understood from the context. When we want to specify the dimension, e_r denotes the column vector in $\mathbf{R}^r$, all of whose entries are equal to 1.

I, I_r The symbol I denotes the unit matrix, its order understood from the context. When we want to specify the order, I_r denotes the unit matrix of order r.

$|\alpha|$ Absolut value of the real number α.

$\square$ This symbol indicates the end of a proof.

y^+ If $y = (y_j) \in \mathbf{R}^n$, let $y_j^+ = \text{Maximum } \{0, y_j\}$, $j = 1$ to n. Then $y^+ = (y_j^+)$.

$\succ$ Lexicographically greater than. Given two vectors $x = (x_j)$, $y = (y_j)$ in $\mathbf{R}^n$, $x \succ y$ means that for the smallest j for which $x_j - y_j \neq 0$, we have $x_j - y_j > 0$.

Pos$\{A_1, \ldots, A_k\}$ If $A_1, \ldots, A_k$ are vectors in $\mathbf{R}^n$ then Pos$\{A_1, \ldots, A_k\} = \{y : y = \alpha_1 A_1 + \ldots + \alpha_k A_k, \alpha_1 \geqq 0, \ldots, \alpha_k \geqq 0\}$. It is the cone in $\mathbf{R}^n$ which is the nonnegative hull of the set of vectors $\{A_1, \ldots, A_k\}$.

Pos(A) If A is a matrix, Pos$(A) = \{x : x = Ay$ for some $y \geqq 0\}$. It is the cone which is the nonnegative hull of the column vectors of the matrix A.

$n!$ n factorial. Defined only for nonnegative integers. $0! = 1$. And $n!$ is the product of all the positive integers from 1 to n, whenever n is a positive integer.

$\binom{n}{r}$ Defined only for positive integers $n \geqq r$. It is the number of distinct subsets of r objects from a set of n distinct objects. It is equal to $\frac{n!}{r!(n-r)!}$.

$\langle v_1, \ldots, v_r \rangle$ When $v_1, \ldots, v_r$ are all column vectors from the space $\mathbf{R}^n$, say, and satisfy the property that the set of column vectors $\left\{ \begin{pmatrix} 1 \\ v_1 \end{pmatrix}, \ldots, \begin{pmatrix} 1 \\ v_r \end{pmatrix} \right\}$ is linearly independent, then $v_1, \ldots, v_r$ are the vertices of an $(r-1)$-dimensional simplex, which is their convex hull, this simplex is denoted by the symbol $\langle v_1, \ldots, v_r \rangle$. See Section 2.7.8.

$\mathcal{C}(M)$ The class of $2^{\mathbf{n}}$ complementary cones associated with the square matrix M of order n.

K(M) The union of all complementary cones in $\mathcal{C}(M)$. It is the set of all vectors q for which the LCP (q, M) has at least one solution.

Z(y), **W**(y) If $y = (y_1, \ldots, y_n)^T$ is a complementary vector for the LCP (q, M) of order n, then **Z**$(y) = \{j : y_j = z_j\}$ and **W**$(y) = \{j : y_j = w_j\}$. See Section 3.1.

Miminum{ } The minimum number among the set of numbers appearing inside the set brackets. Maximum{ } has a similar meaning. If the set is empty we will adopt the convention that the minimum in it is $+\infty$ and the maximum in it is $-\infty$.

Infimum, minimum; Supremum, maximum Let $\boldsymbol{\Gamma}$ be a subset of $\mathbf{R}^n$ and let $f(x)$ be a real valued function defined on $\boldsymbol{\Gamma}$. The **infimum** for $f(x)$ on $\boldsymbol{\Gamma}$ is defined to be the largest number α satisfying: $f(x) \geqq \alpha$ for all $x \in \boldsymbol{\Gamma}$. If α_0 is the infimum for $f(x)$ on $\boldsymbol{\Gamma}$, and there exists an $\bar{x} \in \boldsymbol{\Gamma}$ satisfying $f(\bar{x}) = \alpha_0$, then α_0 is said to be the **minimum value** of $f(x)$ on $\boldsymbol{\Gamma}$ and $\bar{x}$ is the point which attains it. As an example let $\boldsymbol{\Gamma} \subset \mathbf{R}^1$ be the open interval $0 < x < 1$, and let $f(x) = x$. The infimum of $f(x)$ on $\boldsymbol{\Gamma}$ in this example is 0, it is not a minimum since $0 \notin \boldsymbol{\Gamma}$, and there exists no point x in $\boldsymbol{\Gamma}$ where $f(x) = 0$. As another example let $\boldsymbol{\Gamma} \subset \mathbf{R}^1$ be the unbounded set $1 \leqq x < \infty$ and let $f(x) = \frac{1}{x}$. In this example, the infimum of $f(x)$ on $\boldsymbol{\Gamma}$ is 0, and again this is not a minimum. In the same manner, the **supremum** in $\boldsymbol{\Gamma}$ of a real valued function $f(x)$ defined on $\boldsymbol{\Gamma} \subset \mathbf{R}^n$, is the smallest number γ satisfying: $f(x) \leqq \gamma$ for all $x \in \boldsymbol{\Gamma}$. If γ_0 is the supremum of $f(x)$ on $\boldsymbol{\Gamma}$, and there exists an $\hat{x} \in \boldsymbol{\Gamma}$ satisfying $f(\hat{x}) = \gamma_0$, then γ_0 is said to be the **maximum value** of $f(x)$ on $\boldsymbol{\Gamma}$, and $\hat{x}$ is the point which attains it.

Local minimum, global minimum Consider an optimization problem in which an objective function $\theta(x)$, which is a real valued function defined over $\mathbf{R}^n$, is required to be minimized, subject to possibly some constraints on the decision variables x. Let $\mathbf{K} \subset \mathbf{R}^n$ denote the set of feasible solutions for this problem. A point $\hat{x} \in \mathbf{K}$ is said to be a global minimum for this problem if there exists no $x \in \mathbf{K}$ satisfying $\theta(x) < \theta(\hat{x})$. A point $\bar{x} \in \mathbf{K}$ is said to be a local minimum for this problem if there exists an $\varepsilon > 0$ such that the following system has no feasible solution

$$\begin{aligned} &x \in \mathbf{K} \\ &\theta(x) < \theta(\bar{x}) \\ &||x - \bar{x}|| < \varepsilon \end{aligned}$$

that is, $\bar{x}$ is a local minimum for this problem iff $\bar{x}$ is a global minimum for $\theta(x)$ over $\mathbf{K} \cap \{x : ||x - \bar{x}|| < \varepsilon\}$. See Section 10.2.

Cardinality Defined only for sets. The cardinality of a set is the number of distinct elements in it.

Principal Submatrix $F_{\mathbf{JJ}}$ of square matrix F Let $F = (f_{ij})$ be a given square matrix of order n. Let $\mathbf{J} \subset \{1, \ldots, n\}$. The **principal submatrix** of F determined by the subset $\mathbf{J}$ is the matrix $F_{\mathbf{JJ}} = (f_{ij} : i \in \mathbf{J}, j \in \mathbf{J})$. See Section 1.3.1. The determinant of $F_{\mathbf{JJ}}$ is known as the principal subdeterminant of F corresponding to the subset $\mathbf{J}$.

BFGS updating formula The Broyden-Fletcher-Goldfarb-Shanno formula for updating a positive definite symmetric approximation to the Hessian (or its inverse) of a twice continuously real valued function $\theta(x)$ defined on $\mathbf{R}^n$, as the algorithm moves from one point to next. See Sections 1.3.6 and 10.8.6.

LCP Linear complementarity problem.

NLCP Nonlinear complementarity problem.

LP Linear program.

BFS Basic feasible solution.

NLP Nonlinear program.

PD Positive definite. A square matrix M of order n is said to be PD if $y^T My > 0$ for all $y \in \mathbf{R}^n$, $y \neq 0$.

PSD Positive semidefinite. A square matrix M of order n is said to be PSD if $y^T My \geqq 0$ for all $y \in \mathbf{R}^n$.

ND Negative definite. A square matrix of order n is said to be ND if $y^T My < 0$ for all $y \in \mathbf{R}^n$, $y \neq 0$.

NSD Negative semidefinite. A square matrix of order n is said to be NSD if $y^T My \leqq 0$ for all $y \in \mathbf{R}^n$.

PPT Principal pivot transform. See Section 3.2.

$(i.j)$ This refer to the jth equation in the ith chapter. Equations are numbered serially in each chapter.

Section $i.j$**;** $i.j.k$ The sections are numbered serially in each chapter. "$i.j$" refers to section j in Chapter i. "$i.j.k$" refers to subsection k in section $i.j$.

Figure $i.j$ The jth figure in Chapter i. The figures are numbered serially in this manner in each chapter.

Reference $[i.j]$ The jth reference in the list of references given at the end of the Chapter i. References given at the end of each chapter are numbered serially.

Exercise $i.j$ The jth exercise in Chapter i. Exercises are numbered serially in each chapter.

Figure i**, Exercise** i**, Theorem** i**, Reference** i**, Example** i In the appendices, figures, examples, exercises, theorems, references, etc. are numbered serially using a single number for each. So any figure, example, exercise, theorem or reference with a single number like this must be in the appendix.

Linear Function, affine function

The real valued function $f(x)$ defined over $x \in \mathbf{R}^n$ is called a **linear function** if $f(x) = c_1x_1 + \ldots + c_nx_n$ where $c_1, \ldots, c_n$ are constants, it satisfies the property: $f(\alpha x^1 + \beta x^2) = \alpha f(x^1) + \beta f(x^2)$ for all $x^1, x^2 \in \mathbf{R}^n$ and for all real numbers α, β. The real valued function $g(x)$ defined over $x \in \mathbf{R}^n$ is said to be an **affine function** if $g(x) = \gamma_0 + \gamma_1x_1 + \ldots + \gamma_nx_n$ where $\gamma_0, \gamma_1, \ldots, \gamma_n$ are constants, it satisfies the property: $g(\alpha x^1 + \beta x^2) = \alpha g(x^1) + \beta g(x^2)$ for all $x^1, x^2 \in \mathbf{R}^n$ and for all real numbers α, β satisfying $\alpha + \beta = 1$. Every affine function defined over $\mathbf{R}^n$ in a linear function plus a constant.

Basis, basic vector, basic solution, basic feasible solution

See Section 2.1.

Bounded set

A subset $\mathbf{S} \subset \mathbf{R}^n$ is bounded if there exists a finite real number α such that $||x|| \leqq \alpha$, for all $x \in \mathbf{S}$.

Proper subset

If $\mathbf{E}$ is a subset of a set $\mathbf{\Gamma}$, $\mathbf{E}$ is said to be a proper subset of $\mathbf{\Gamma}$ if $\mathbf{E} \neq \mathbf{\Gamma}$, that is, if $\mathbf{\Gamma} \setminus \mathbf{E} \neq \emptyset$.

Feasible solution

A numerical vector that satisfies all the constraints and restrictions in the problem.

Optimum solution or Optimum feasible solution

A feasible solution that optimizes (i. e., either maximizes or minimizes as required) the objective value among all feasible solutions.

Algorithm

The word from the last name of the Persian scholar Abu Ja'far Mohammed ibn Mûsâ alkhowârizmî whose textbook on arithmetic (about A.D. 825) had a significant influence on the development of these methods. An **algorithm** is a set of rules for getting a required output from a specific input, in which each step is so precisely defined that it can be translated into computer language and executed by machine.

Size The size of an optimization problem is a parameter that measures how large the problem is. Usually it is the number of digits in the data in the optimization problem, when it is encoded in binary form.

$\mathcal{O}(n^{\mathbf{r}})$ A finitely terminating algorithm for solving an optimization problem is said to be of order $n^{\mathbf{r}}$ or $\mathcal{O}(n^{\mathbf{r}})$, if the computational effort required by the algorithm in the worst case, to solve a version of the problem of size n, grows as $\alpha n^{\mathbf{r}}$, where α, r are numbers that are independent of the size n and the data in the problem.

Polynomially bounded algorithm An algorithm is said to be polynomially bounded if it can be proved that the computational efffort required by it is bounded above by a fixed polynomially in the size of the problem.

The class $\mathcal{P}$ of problems This is the class of all problems for solving which there exists a polynomially bounded algorithm.

$\mathcal{NP}$-complete class of problems

A decision problem is one for which the answer is "yes" or "no". For example, given an integer square matrix D of $\mathbf{R}^n$, the problem "is there an $x \in \mathbf{R}^n$ satisfying $x^T Dx < 0$?" is a decision problem. Also, given a square matrix M of order n and a column vector $q \in \mathbf{R}^n$, the problem "does the LCP (q, M) have a solution?" is a decision problem. Often, optimization problems can be handled by studying decision problem versions of them. For example, consider the problem of minimizing $\theta(x)$ over $x \in \mathbf{K}$, where $\mathbf{K}$ represents the set of feasible solutions of this problem. The decision problem version of this optimization problem is "is there an $x \in \mathbf{K}$ satisfying $\theta(x) \leqq \alpha$?" where α is a specified real number. Clearly, by examining this decision problem with varying values of α, we can narrow down the solution of the optimization problem.

The $\mathcal{NP}$-complete class is a class of decision problems in discrete optimization, satisfying the property that if a polynomially bound algorithm exists for any one problem in the classs, then polynomially bounded algorithms exist for every problem in the class. So far no polynomially bounded algorithm is known for any problem in the $\mathcal{NP}$-complete class, and it is believed that all these problems are hard problems (in the worst case, the computational effort required for solving an instance of any problem in the class by any known algorithm, grows asymptotically, faster than any polynomial in the size of the problem). See reference [8.12] for a complete discussion of $\mathcal{NP}$-completeness.

Necessary conditions, sufficient conditions, necessary and sufficient conditions

When studying a property of a system, a condition is said to be a **necessary condition** for that property if that condition is satisfied whenever the property holds. A condition is said to be a **sufficient condition** for the property if the property holds whenever the condition is satisfied. A **necessary and sufficient condition** for the property is a condition that is both necessary condition and a sufficient condition for that property.

Active or tight constraint An inequality constraint $g_p(x) \geqq 0$ is said to be **active** or **tight**, at a point $\bar{x}$ satisfying it, if $g_p(\bar{x}) = 0$. The equality constraint $h_i(x) = 0$ is always an active constraint at any point $\bar{x}$ satifying it.

Infeasible system A system of constraints in the variables $x = (x_j)$ is said to be infeasible, if there exists no vector x satisfying all the constraints.

Complementary pair A pair of variables in an LCP, at least one of which is required to be zero. Each variable in a complementary pair is said to be the complement of the other. A pair of column vectors corresponding to a complementary pair of variables in an LCP is a complementary pair of column vectors. Each column vector in a complementary pair is the complement of the other. In an LCP of order n, there are n complementary pairs, numbered 1 to n.

Complementary set of vectors A vector of n variables in an LCP of order n is a complementary vector if the jth variable in the vector is from the jth complementary pair of variables, for each j. A complementary set of column vectors is an ordered set in which the jth vector is from the jth complementary pair for each j.

Complementary matrix In an LCP of order n, this is a square matrix of order n whose jth column vector is from the jth complementary pair, for each j.

Complementary cone In an LCP of order n, this is $\text{Pos}(A)$ where A is a complementary matrix of this problem.

Complemetary basis It is a complementary matrix which is nonsingular.

Complementary basic vector It is a complementary vector of variables associated with a complementary basis.

Complementary feasible basis It is a complementary basis which is a feasible basis for the problem.

Complementary feasible basic vector — It is a complementary basic vector which is feasible to the problem.

$\bar{z}$ leads to a solution of the LCP (q, M) — We say that the vector $\bar{z}$ leads to a solution of the LCP (q, M) if $(\overline{w} = M\bar{z} + q, \bar{z})$ is a solution of the LCP (q, M).

To process an LCP — When an algorithm for solving LCPs is applied on an LCP, it may either obtain a solution of the LCP, or terminate without obtaining a solution. It is possible that some algorithms may terminate without a solution even though the LCP may have a solution. An algorithm for solving LCPs is said **to process a specified class of LCPs** if, when the algorithm is applied on any LCP from this class and it terminates without obtaining a solution, we can prove that the LCP in fact has no solution. In other words, an algorithm is said to process a class of LCPs iff for every LCP in this class, the algorithm either produces a solution or conclusively establishes that the LCP cannot have a solution.

Secondary ray or terminal ray — This is the half-line or ray obtained at the end of executing the complementary pivot algorithm on an LCP, if the algorithm terminates in ray termination. This secondary ray, if it is obtained, is distinct from the initial ray with which the algorithm is initiated. See Section 2.2.6.

Subcomplementary set, vector — It is a complementary set or vector with one element missing.

Almost complementary vector — It is a vector that is complementary except for one violation which is set up appropriately. See Sections 2.2.4, 2.4.

Copositive matrix — A square matrix M of order n is said to be copositive if $y^T My \geqq 0$ for all $y \geqq 0$ in $\mathbf{R}^n$.

Strictly copositive matrix — A square matrix M of order n is said to be strictly copositive if $y^T My > 0$ for all $y \geq 0$ in $\mathbf{R}^n$.

Copositive plus matrix A square matrix M of order n is said to be copositive plus if it is copositive, and for $y \geqq 0$ in $\mathbf{R}^n$ if $y^T My = 0$ then $(M + M^T)y = 0$.

P_0-matrix A square matrix is a P_0-matrix if all its principal subdeterminants are $\geqq 0$.

P-matrix A square matrix is said to be a P-matrix if all its principal subdeterminants are strictly positive.

Q-matrix A square matrix M of order n is said to be a Q-matrix if the LCP (q, M) has a solution for all $q \in \mathbf{R}^n$.

Z-matrix A square matrix $M = (m_{ij})$ is a Z-matrix if $m_{ij} \leqq 0$ for all $i \neq j$.

Q_0-matrix The square matrix M is said to be a Q_0-matrix if $\mathbf{K}(M)$ is a convex cone.

$\bar{Q}$-matrix, or Completely Q-matrix A square matrix M such that M and all its principal submatrices are Q-matrices.

$\bar{Q}_0$-matrix, or Completely Q_0-matrix A square matrix M such that M and all its principal submatrices are Q_0-matrices.

Faces, Facets

Let $\mathbf{K} \subset \mathbf{R}^n$ be a convex polyhedron. $\mathbf{H} = \{x : ax = a_0\}$ where $a \neq 0$ is a given row vector in $\mathbf{R}^n$. $\mathbf{H}$ is a hyperplane in $\mathbf{R}^n$. $\mathbf{H}$ is said to have $\mathbf{K}$ on one of its sides if either $ax \geqq a_0$ for all $x \in \mathbf{K}$, or $ax \leqq a_0$ for all $x \in \mathbf{K}$. If $\mathbf{H}$ has $\mathbf{K}$ on one of its sides and $\mathbf{H} \cap \mathbf{K} \neq \emptyset$, $\mathbf{H}$ is said to be a **supporting hyperplane** for $\mathbf{K}$. A **face** of $\mathbf{K}$ is either the empty set $\emptyset$, or the set $\mathbf{K}$ itself, or $\mathbf{H} \cap \mathbf{K}$ for some supporting hyperplane $\mathbf{H}$ for $\mathbf{K}$. See reference [2.26]. For example, extreme points of $\mathbf{K}$ are its faces of dimension zero. Edges of $\mathbf{K}$ are its faces of dimension 1, etc.

A face of $\mathbf{K}$ is said to be a **facet** if its dimension is one less than the dimension of $\mathbf{K}$.

For some special convex polyhedra, simplicial cones or simplexes, it is possible to characterize all faces easily. If $\{B_{.1}, \ldots, B_{.n}\}$ is a linearly independent set of column vectors in $\mathbf{R}^n$, then, for the simplicial cone $\text{Pos}\{B_{.1}, \ldots, B_{.n}\}$, the cone $\text{Pos}\{B_{.1}, \ldots, B_{.j-1}, B_{.j+1}, \ldots, B_{.n}\}$ is a facet for any j, and the cone $\text{Pos}\{B_{.j} : j \in \mathbf{J}\}$ is a face for any subset $\mathbf{J} \subset \{1, \ldots, n\}$ (this face is defined to be $\{0\}$, if $\mathbf{J} = \emptyset$). If $\{v_0, \ldots, v_n\}$ are the set of vertices of an n-dimensonal simplex in $\mathbf{R}^n$, the convex hull of $\{v_0, \ldots, v_{j-1}, v_{j+1}, \ldots, v_n\}$ is a facet of this simplex for all j, and the convex hull of $\{v_j : j \in \mathbf{J}\}$ is a face of this simplex for all subsets $\mathbf{J} \subset \{1, \ldots, n\}$ (this face is defined to be the empty set if $\mathbf{J} = \emptyset$).

Principally degenerate, principally nondegenerate, matrices

A square matrix A is said to be **principally nondegenerate** if all its principal subdeterminantes are nonzero; **principally degenerate** if at least one of its principal subdeterminantes has value zero. In this book we are usually concerned only with principal degeneracy or nondegeneracy of square matrices, and hence we usually omit the adjective "principally" and refer to the matrices as being degenerate or nondegenerate.

Degenerate or nondegenerate complementary cone

A complementary cone is nondegenerate if its interior is nonempty, degenerate otherwise.

Strongly degenerate or weakly degenerate complementary cone

A degenerate complementary cone $\text{Pos}(A_{.1}, \ldots, A_{.n})$ is said to be **strongly degenerate** if there exists $(\alpha_1, \ldots, \alpha_n) \geq 0$ such that $0 = \alpha_1 A_{.1} + \ldots + \alpha_n A_{.n}$, that is, if the zero vector can be expressed as a semipositive linear combination of the complementary set of column vectors $\{A_{.1}, \ldots, A_{.n}\}$; **weakly degenerate** otherwise.

Degenerate or nondegenerate basic solutions, vectors, systems of linear equations

Consider the system of linear constraints "$Ax = b$" where A is a matrix of order $m \times n$ and rank m. A basic solution $\bar{x}$ for this system is said to be **nondegenerate** if the number of nonzero variables in $\bar{x}$ is m, **degenerate** if this number is $< m$. The right hand side constants vector b in the system is said to be degenerate if the system has at least one degenerate basic solution, b is said to be nondegenerate if the system has no degenerate basic solution. Thus b is degenerate in the system if it can expressed as a linar combination of $m - 1$ or less column vectors of A, nondegenerate otherwise. The system of constraints is itself said to be degenerate or nondegenerate depending on whether b is degenerate or nondegenerate.

Lipschitz continuous

Let $f(x)$ be a continuous real valued function defined on $\mathbf{K} \subset \mathbf{R}^n$. It is said to be Lipschitz continuous (or Lipschitzian) on $\mathbf{K}$ if there exists a nonnegative number α such that $|f(x) - f(y)| \leqq \alpha ||x - y||$ for all $x, y \in \mathbf{K}$. The number α is known as the Lipschitz constant for this function.

Principal subproblem

Consider the LCP (q, M) with variables $(w_1, \ldots, w_n)^T$, $(z_1, \ldots, z_n)^T$. Let $\mathbf{J} \subset \{1, \ldots, n\}$, $\mathbf{J} \neq \emptyset$. Let $q_{\mathbf{J}} = (q_i : i \in \mathbf{J})^T$, $M_{\mathbf{JJ}} = (m_{ij} : i \in \mathbf{J}, j \in \mathbf{J})$. The LCP $(q_{\mathbf{J}}, M_{\mathbf{JJ}})$ in variables $w_{\mathbf{J}}$, $z_{\mathbf{J}}$ is called the **principal subproblem** of the LCP (q, M) corresponding to the subset $\mathbf{J}$.

Simplex

See Section 2.7.

$\nabla\theta(\bar{x})$

The row vector of partial derivatives $\left(\frac{\partial\theta(x)}{\partial x_1}, \ldots, \frac{\partial\theta(x)}{\partial x_n}\right)$, **gradient vector**, evaluated at $x = \bar{x}$.

$\partial f(x)$ — The subdifferential set of the function $f(x)$ at the point x. See Appendix 3 and Section 2.7.1.

Differentiable function — A real valued function $\theta(x)$ defined on an open subset $\mathbf{\Gamma} \subset \mathbf{R}^n$ is said to be **differentiable** at a point $\bar{x} \in \mathbf{\Gamma}$, if all the partial derivatives $\frac{\partial \theta(\bar{x})}{\partial x_j}$, $j = 1$ to n exist, and for any $y \in \mathbf{R}^n$, $[\theta(\bar{x} + \alpha y) - \theta(\bar{x}) - \alpha \nabla \theta(\bar{x}) y]/\alpha$ tends to zero as α tends to zero. If it is differentiable at every point $\bar{x} \in \mathbf{\Gamma}$, it is said to be differentiable in $\mathbf{\Gamma}$.

Continuously differentiable function — A real-valued function $\theta(x)$ defined on an open subset $\mathbf{\Gamma} \in \mathbf{R}^n$ is said to be **continuously differentiable** at a point $\bar{x} \in \mathbf{\Gamma}$ if it is differentiable at $\mathbf{\Gamma}$ and $\nabla \theta(x)$ is contiuous at $\bar{x}$. If it is continuously differentiable at every point $\bar{x} \in \mathbf{\Gamma}$, it is said to be continuoulsy differentiable in $\mathbf{\Gamma}$.

$H(\theta(\bar{x}))$ — The **Hessian matrix** of $\theta(x)$ at $\bar{x}$. It is the square matrix of second partial derivatives $\left(\frac{\partial^2 \theta(\bar{x})}{\partial x_i \partial x_j}\right)$ evaluated at $\bar{x}$.

Twice differentiable function — A real valued function $\theta(x)$ defined over an open set $\mathbf{\Gamma} \in \mathbf{R}^n$ is said to be **twice differentiable** at $\bar{x} \in \mathbf{\Gamma}$ if $\nabla \theta(\bar{x})$ and $H(\theta(\bar{x}))$ exist, and for all $y \in \mathbf{R}^n$, $[\theta(\bar{x}+\alpha y) - \theta(\bar{x}) - \alpha(\nabla \theta(\bar{x}))y - \frac{\alpha^2}{2} y^T H(\theta(\bar{x}))y]/\alpha^2$ tends to zero as α tends to zero. $\theta(x)$ is said to be twice differentiable in $\mathbf{\Gamma}$ if it is twice differentiable at every point in $\mathbf{\Gamma}$.

Twice continuously differentiable function — A real valued function $\theta(x)$ defined over an open set $\mathbf{\Gamma} \in \mathbf{R}^n$ is said to be **twice continuously differentiable** at $\bar{x} \in \mathbf{\Gamma}$ if it is twice differentiable at $\bar{x}$ and $H(\theta(x))$ is continuous at $\bar{x}$. It is twice continuously differentiable in $\mathbf{\Gamma}$ if it is twice continuously differentiable at every point in $\mathbf{\Gamma}$.

Smooth function

Mathematically, a real valued function defined on $\mathbf{R}^n$ is said to be a smooth function if it has derivatives of all orders. Many of the algorithms discussed in this book use only derivatives of the first or at most second orders. So, for our purpose, we will consider a smooth function to be one which is continuously differentiable, or twice continuously differentiable if the method under consideration uses second order derivatives.

Optimization problems in minimization form

Whenever a function $f(x)$ has to be maximized subject to some conditions, we can look at the equivalent problem of minimizing $-f(x)$ subject to the same conditions. Both problems have the same set of optimum solutions and the maximum value of $f(x) = -$minimum value of $(-f(x))$. Because of this, we discuss only minimization problems.

$\nabla h(x)$ when $h(x) = (h_1(x), \ldots, h_m(x))^T$

Let $h(x)$ denote the column vector of m differentiable functions $h_i(x)$, $i = 1$ to m, defined over $\mathbf{R}^n$. Then $\nabla h(x) = \left(\frac{\partial h_i(x)}{\partial x_j} : i = 1 \text{ to } m,\ j = 1 \text{ to } n\right)$ is the **Jacobian matrix** in which the ith row vector is the gradient vector of $h_i(x)$ written as a row vector.

Nonlinear programming problem

This refers to an optimization problem of the following general form :

$$\begin{array}{lll} \text{minimize} & \theta(x) & \\ \text{subject to} & h_i(x) = 0, & i = 1 \text{ to } m \\ & g_p(x) \geqq 0, & p = 1 \text{ to } t \end{array}$$

where all the functions $\theta(x)$, $h_i(x)$, $g_p(x)$ are real valued continuous functions of $x = (x_1, \ldots, x_n)^T \in \mathbf{R}^n$. The problem is said to be a **smooth nonlinear program** if all the functions are in fact continuously differentiable functions. In this book we only consider smooth nonlinear programs. See Chapter 10.

Quadratic forms in matrix notations

Consider the quadratic form in n variables $x = (x_1, \ldots, x_n)^T$, $f(x) = \sum_{i=1}^{n} g_{ii}x_i^2 + \sum_{i=1}^{n}\sum_{j=i+1}^{n} g_{ij}x_ix_j$. An example for $n = 3$ is $h(x) = 81x_1^2 - 7x_2^2 + 5x_1x_2 - 6x_1x_3 + 18x_2x_3$. Let $F = (f_{ij})$ be a square matrix of order n satisfying

$$f_{ii} = g_{ii}, \quad i = 1 \text{ to } n$$
$$f_{ij} + f_{ji} = g_{ij}, \quad \text{for } i \neq j \text{ and } j > i.$$

Then it can be verified that $f(x) = x^T F x$. In particular, if we define the symmetric matrix $D = (d_{ij})$ of order n, where

$$d_{ii} = g_{ii}, \quad i = 1 \text{ to } n$$
$$d_{ij} = d_{ji} = \frac{1}{2}g_{ij}, \quad \text{for } i \neq j \text{ and } j > i$$

then $f(x) = x^T D x$. For the quadratic form $h(x)$ in 3 variables, $x = (x_1, x_2, x_3)^T$, given above, the matrix D turns out to be

$$D = \begin{pmatrix} 81 & \frac{5}{2} & -3 \\ \frac{5}{2} & -7 & 9 \\ -3 & 9 & 0 \end{pmatrix}$$

and $h(x) = x^T D x$.

Quadratic programming problem; convex or nonconvex quadratic programs

An optimization problem in which a quadratic function of $x = (x_1, \ldots, x_n)^T \in \mathbf{R}^n$ is to be optimized subject to linear constraints on the variables, is called a quadratic programming problem. Its general form is:

$$\begin{array}{ll} \text{minimize} & Q(x) = cx + \frac{1}{2}x^T Dx \\ \text{subject to} & Ax \geqq b \\ & Ex = d \end{array}$$

where D is a square symmetric matrix of order n. The inequality constraints here include any nonnegativity restrictions or the lower or upper bound restrictions on the variables.
This problem is called a **convex quadratic program** if D is a PSD matrix (in this case the objective function to be minimized, $Q(x)$, is convex); a **nonconvex quadratic program** otherwise.

QP Quadratic Programming Problem.

Complemetary basis It is a complementary matrix which is nonsingular.

$\nabla_x(f(\overline{x}, \overline{\mu}))$, $H_x(f(\overline{x}, \overline{\mu}))$ These are respectively the row vector of the partial deri-vates, and the square matrix of the second order partial derivates, of the function $f(x, \mu)$, with respect to the variables in the vector x, at $(\overline{x}, \overline{\mu})$.

Karush-Kuhn-Tucker (or KKT) necessary optimality conditions

Let $\theta(x)$, $h_i(x)$, $g_p(x)$, be real valued continuously differentiable functions defined on $\mathbf{R}^n$ for all i, p. Consider the following mathematical program:

$$\begin{array}{lll} \text{minimize} & \theta(x) & \\ \text{subject to} & h_i(x) = 0, & i = 1 \text{ to } m \\ & g_p(x) \geqq 0, & p = 1 \text{ to } t \end{array}$$

The Karush-Kuhn-Tucker (KKT) Lagrangian for this problem is: $L(x, \mu, \pi) = \theta(x) - \sum_{i=1}^{m} \mu_i h_i(x) - \sum_{p=1}^{t} \pi_p g_p(x)$ where μ_i, π_p are the Lagrange multipliers associated with the constraints. The Karush-Kuhn-Tucker (KKT) necessary optimality condition for this problem are :

$$\frac{\partial}{\partial x} L(x, \mu, \pi) = \nabla\theta(x) - \sum_{i=1}^{m} \mu_i \nabla h_i(x) - \\ - \sum_{p=1}^{t} \pi_p \nabla g_p(x) = 0$$

$$\begin{aligned} h_i(x) &= 0, \quad i = 1 \text{ to } m \\ g_p(x) &\geqq 0, \quad p = 1 \text{ to } t \\ \pi_p &\geqq 0, \quad p = 1 \text{ to } t \\ \pi_p g_p(x) &= 0, \quad p = 1 \text{ to } t \end{aligned}$$

where $\nabla\theta(x)$ etc. are the vectors of partial derivatives. If $\bar{x}$ is a local minimum for this problem, under fairly general conditions (see Appendix 4) it can be shown that there exist multiplier vectors $\bar{\mu}$, $\bar{\pi}$ such that $\bar{x}$, $\bar{\mu}$, $\bar{\pi}$ together satisfy these KKT conditions. In the literature these conditions are usually called **first-order necessary optimality conditions** or Kuhn-Tucker conditions. But it has been found recently that Karush was the first to discuss them. Hence, nowadays, the name Karush-Kuhn-Tucker necessary optimality conditions is coming into Vogue.

A feasible solution $\bar{x}$ satisfying the property that there exist Lagrange multiplier vectors $\bar{\pi}$, $\bar{\mu}$ such that $\bar{x}$, $\bar{\mu}$, $\bar{\pi}$ together satisfy the KKT conditions, is called a KKT point for the problem.

Stationary point for an NLP

Given an NLP, a stationary point for it usually refers to any feasible solution satisfying a necessary optimality condition for it. Every optimum solution is a stationary point, but, in general, there may be stationary points which are not even locally optimal to the problem.

Direction, half-line

Any point $y \in \mathbf{R}^n$, $y \neq 0$ defines a **direction** in $\mathbf{R}^n$. Given $\bar{x} \in \mathbf{R}^n$, points $\bar{x} + \alpha y$, $\alpha \geqq 0$ are obtained when you move from $\bar{x}$ in the direction y. The set of all these points $\{x : x = \bar{x} + \alpha y, \alpha \geqq 0\}$ is the **half-line** or **ray** through $\bar{x}$ in the direction of y. See Section 1.1.1.

Step length

Given $\bar{x} \in \mathbf{R}^n$, $y \in \mathbf{R}^n$, $y \neq 0$; for $\alpha > 0$, the point $\bar{x}+\alpha y$ is obtained by taking a step of length α from $\bar{x}$ in the direction of y. In this process α is the **step length**.

Feasible direction

Given a set $\mathbf{\Gamma} \subset \mathbf{R}^n$, and a point $\bar{x} \in \mathbf{\Gamma}$; the direction $y \in \mathbf{R}^n$, $y \neq 0$, is called a **feasible direction** at $\bar{x}$ for $\mathbf{\Gamma}$ if there exists a positive number $\bar{\alpha}$ such that $\bar{x} + \alpha y \in \mathbf{\Gamma}$ for all $0 \leqq \alpha \leqq \bar{\alpha}$. Thus the direction y is a feasible direction at $\bar{x}$ for $\mathbf{\Gamma}$ iff an initial segment of positive length on the half-line through $\bar{x}$ in the direction y is contained in $\mathbf{\Gamma}$.

Given an optimization problem, and a feasible solution $\overline{x}$ for it, the direction y (in the x-space) is said to be a feasible direction at $\overline{x}$ for this optimization problem if there exists an $\overline{\alpha} > 0$ such that $\overline{x}+\alpha y$ is a feasible solution to the problem for all $0 \leqq \alpha \leqq \overline{\alpha}$.

Descent direction

Let $\theta(x)$ be a real valued function defined over $x \in \mathbf{R}^n$. The direction $y \in \mathbf{R}^n$, $y \neq 0$, is said to be a **descent direction** for $\theta(x)$ at $\bar{x}$ if $\theta(\bar{x}+\alpha y) < \theta(\bar{x})$ whenever α is positive and sufficiently small. So by moving from $\bar{x}$ a small but positive step length in a descent direction, $\theta(x)$ is guaranteed to strictly decrease in value.

A descent direction for a minimization problem at a feasible solution $\overline{x}$, is a feasible direction for the problem at $\overline{x}$, which is a descent direction at $\overline{x}$ for the objective function being minimized.

Line search problem, line search method

Let $\theta(x)$ be a real valued function defined on $\mathbf{R}^n$. Let $\bar{x} \in \mathbf{R}^n$ be a given point and $y \in \mathbf{R}^n$, $y \neq 0$ a given direction. The problem of minimizing $\theta(\bar{x} + \alpha y)$ over $a \leqq \alpha \leqq b$ where a, b are given bounds on α, is called a **line search problem** or a **line minimization problem**; and any method for solving such a problem is called a **line search method**. Since $\bar{x}$, y are given, $\theta(\bar{x} + \alpha y)$ is purely a function of the single variable α, if we denote $\theta(\bar{x} + \alpha y) = f(\alpha)$, the line search problem is the one dimensional minimization problem of finding the minimum of $f(\alpha)$ over $a \leqq \alpha \leqq b$. Typically, in most line search problems encountered in applications, we will have $a = 0$ and b is either a finite positive number, or $+\infty$. When b is finite, the problem is often called a **constrained line search problem**. Several line search methods are discussed in Section 10.7. Many nonlinear programming algorithms use line search methods repeatedly in combination with special subroutines for generating feasible descent directions.

Hereditary symmetry, hereditary PD

Many algorithms for nonlinear programming (for example those discussed in Section 1.3.6 or Chapter 10) are iterative methods which maintain a square matrix B of order n and update it in each step. Let B_t denote this matrix in the tth step. The updating formula in this method provides B_{t+1} as a function of B_t and other quantities which are computed in the tth step or earlier. This updating procedure is said to possess the **hereditary symmetry property** if for any t, the fact that B_t is symmetric implies that B_{t+1} is also symmetric. Similarly, the updating procedure possesses the **hereditary PD property** if for any t the fact that B_t is PD implies that B_{t+1} is also PD. Thus, if the updating procedure has the hereditary symmetry and PD properties, and the initial matrix B used in the method is both symmetric and PD, the matrices B_t obtained in all the steps of the method will also be symmetric and PD.

Active set method

Any method for solving an NLP which partitions the set of inequality constraints into two groups — the active set consisting of those inequalities which are to be treated as active, that is, as equality constraints; and the inactive set. Inequality constraints in the inactive set are presumed to hold as strict inequalities at the optimum solution and are essentially ignored. The remaining problem is solved (treating all the constraints as equality constraints) by any method for solving equality constrained optimization problems. Active set methods also have procedures for revising the active set (either deleting inequality constraints from it, or adding inequality constraints from the inactive set into it) in each step, based on information accumulated in the method so far.

Convex programming problem, nonconvex programming problem

A problem in which a convex objective function is to be minimized over a convex set (usually of the form: minimize $\theta(x)$, subject to $g_i(x) \geqq 0$, $i = 1$ to m and $h_t(x) = 0$, $t = 1$ to p; where all the functions are given and $\theta(x)$ is convex; $g_i(x)$ are concave for all i; and $h_t(x)$ is affine for all t) is said to be a **convex programming problem**. A **nonconvex programming problem** is one which is not convex, that is, does not belong to the above class. For a convex programming problem every local minimum is a global minimum. In general, it is very hard to find the global minimum in a nonconvex programming problem. Necessary and sufficient conditions for optimality are available for convex programming problems. For nonconvex programming problems we have some necessary conditions for a point to be a local minimum, and sufficient conditions for a given point to be a local minimum. No simple set of conditions which are both necessary and sufficient for a given point to be a local minimum, are known for general nonconvex programming problems.

Merit function

In a nonlinear program where an objective function defined on $\mathbf{R}^n$ is to be minimized subject to constraints, a **merit function** is a real valued function defined on $\mathbf{R}^n$, it consists of the objective function plus penalty terms for constraint violations. Usually the penalty terms come from either the absolute-value penalty function (L_1-penalty function) or the quadratic penalty function. Minimizing the merit function balances the two competing goals which result from the desire to decrease the objective function while reducing the amount by which the constraints fail to be satisfied. See Section 1.3.6.

Cauchy-Schwartz inequality

Let x, y be two column vectors in $\mathbf{R}^n$. Then $|x^T y| \leqq ||x||.||y||$, this inequality is known as the **Cauchy-Schwartz inequality**. To prove it consider the quadratic equation in one variable λ, $f(\lambda) = (\lambda x + y)^T(\lambda x + y) = \lambda^{\mathbf{2}}||x||^{\mathbf{2}} + 2\lambda x^T y + ||y||^{\mathbf{2}} = 0$. Since $f(\lambda) = ||\lambda x + y||^{\mathbf{2}}$, it is always $\geqq 0$. This implies that the equation $f(\lambda) = 0$ can have at most one real solution in λ. It is well known that the quadratic equation $a\lambda^{\mathbf{2}} + b\lambda + c = 0$ has at most one real solution iff $b^{\mathbf{2}} - 4ac \leqq 0$, applying this to the equation $f(\lambda) = 0$, we conclude that $(x^T y)^{\mathbf{2}} \leqq ||x||^{\mathbf{2}}.||y||^{\mathbf{2}}$, that is, $|x^T y| \leqq ||x||.||y||$. Also, the quadratic equation $a\lambda^{\mathbf{2}} + b\lambda + c = 0$ has exactly one real solution if $b^{\mathbf{2}} - 4ac = 0$. Applying this to $f(\lambda) = 0$, we conclude that $f(\lambda) = 0$ has a real solution if $|x^T y| = ||x||.||y||$, in this case since $f(\lambda) = ||\lambda x + y||^{\mathbf{2}} = 0$ for some real λ, we must have $\lambda x + y = 0$, or y is scalar multiple of the vector x. Thus, if the Cauchy-Schwartz inequality holds as an equation for two vectors $x, y \in \mathbf{R}^n$, one of these vectors must be a scalar multiple of the other.

Cholesky factor

If M is a square matrix of order n which is symmetric and positive definite, there exists a lower triangular matrix F of order n with positive diagonal elements, satisfying $M = FF^T$. This matrix F is known as the **Cholesky factor** of M. For efficient methods for computing Cholesky factors, see books on computational linear algebra, or [1.28, 2.26].

Homotopy method

To solve a system by a **homotopy method**, we continuously deform a simple system with a known solution, into the system we are trying to solve. For example, consider the problem of solving a smooth system of n equations in n unknowns "$g(x) = 0$". Let a be an initial point from $\mathbf{R}^n$, consider the simple system of equations "$x = a$" with a known solution. Let $F(x, \lambda) = \lambda g(x) + (1 - \lambda)(x - a)$, on $0 \leqq \lambda \leqq 1$, $x \in \mathbf{R}^n$, $F(x, \lambda)$ is continuous in x and λ. The system "$F(x, \lambda) = 0$", treated as a system of equations in x, with λ as a parameter with given value between 0 and 1; is the simple system when $\lambda = 0$, and the system we want to solve when $\lambda = 1$. As the parameter λ varies from 0 to 1, the system "$F(x, \lambda) = 0$" provides a **homotopy** (contiuous deformation) of the simple system "$x = a$" into the system "$g(x) = 0$". The method for solving "$g(x) = 0$" based on the homotopy $F(x, \lambda)$, would follow the curve $x(\lambda)$ (where $x(\lambda)$ is a solution of $F(x, \lambda) = 0$ as a function of the homotopy parameter λ) beginning with $x(0) = a$, until λ assumes the value 1 at which point we have a solution for "$g(x) = 0$".

Principal rearrangement of a square matrix

Let M be a given square matrix of order n. Let $p = (i_1, \ldots, i_n)$ be a permutation of $(1, \ldots, n)$. The square matrix P of order n whose rows are $I_{i_1 .}$, $I_{i_2 .}, \ldots, I_{i_n .}$ in that order is the permutation matrix corresponding to p. P is obtained by essentially permuting the rows of the unit matrix I of order n using the permutation p. The matrix $M' = PMP^T$ is known as the principal rearrangement of M according to the permutation p. Clearly M' is obtained by first rearranging the rows of M according to the permutation p, and in the resulting matrix, rearranging the columns again accordng to the same permutation p. See Section 3.2.1.

Euclidean distance, rectilinear distance

Let $x = (x_j)$, $y = (y_j)$ be two point in $\mathbf{R}^n$. The Euclidean distance between x and y is $||x - y|| = \sqrt{\sum_{j=1}^{n} (x_j - y_j)^{\mathbf{2}}}$. The rectilinear distance between x and y is $\sum_{j=1}^{n} |x_j - y_j|$.

Steepest descent direction at a feasible solution, in a continuous minimization problem.

First, consider an unconstrained minimization problem

$$\text{minimize } \theta(x) \text{ over } x \in \mathbf{R}^n \qquad \text{(i)}$$

where $\theta(x)$ is a real valued continuous function defined over $\mathbf{R}^n$.

Given any direction $y \in \mathbf{R}^n$, $y \neq 0$, the directional derivative of $\theta(x)$ at a point $\overline{x}$ in the direction y is defined to be

$$\text{limit } \frac{\theta(\overline{x} + \alpha y) - \theta(\overline{x})}{\alpha}$$

as $\alpha \to 0^+$, and denoted by $\theta'(\overline{x}; y)$, when it exists. If $\theta(x)$ is differentiable at $\overline{x}$, then $\theta'(\overline{x}; y) = \nabla\theta(\overline{x})y$. In general, $\theta'(\overline{x}, y)$ may exist even if $\theta(x)$ is not differentiable at $\overline{x}$.

$\theta'(\overline{x}; y)$ measures the rate of change in $\theta(x)$ at $x = \overline{x}$, when moving in the direction y.

The direction y is said to be a descent direction at $\overline{x}$ for problem (i), if $\theta'(\overline{x}; y) < 0$.

If $\overline{x}$ is a local minimum for (i), there is no descent direction for (i) at $\overline{x}$, and hence no steepest descent direction. Unfortunately, the converse of this statement may not always be true, that is, the absence of a descent direction at a point $\overline{x}$ does not imply that $\overline{x}$ is a local minimum. See Exercise 20 in Appendix 6. This just means that descent methods are not always guaranteed to find a local minimum.

If $\overline{x}$ is not a local minimum for (i), an optimum solution of

$$\text{minimize } \theta'(\overline{x}; y) \text{ subject to norm } (y) = 1 \qquad \text{(ii)}$$

is called a steepest descent direction at $\overline{x}$ for (i), under the particular norm used, if it is a descent direction at $\overline{x}$ for (i). In (ii), norm (y) is a function which measures the distance between the points 0 and y is $\mathbf{R}^n$. Different norms may lead to different steepest descent directions.

In optimization literature, usually norm (y) is taken as $y^T Ay$ where A is some specified symmetric PD matrix of order n (taking $A = I$, the unit matrix of order n, leads to the Euclidean norm).

Now consider a constrained continuous minimization
problem. Let $\mathbf{K} \subset \mathbf{R}^n$ denote its set of feasible solutions. Then this problem is of the form

$$\text{minimize } \theta(x) \text{ subject to } x \in \mathbf{K} \qquad \text{(iii)}$$

where the objective function $\theta(x)$ is a real valued continuous function defined over $\mathbf{R}^n$. Let $\overline{x} \in \mathbf{K}$ be a given feasible solution.
Again, if $\overline{x}$ is a local minimum for (iii), there is no descent direction and hence no steepest descent direction for (iii) at $\overline{x}$. If $\overline{x}$ is not a local minimum for (iii), any optimum solution of

$$\begin{aligned} &\text{minimize } \theta'(\overline{x}; y) \\ &\text{subject to norm of } (y) = 1, \\ &\quad \text{and } y \text{ is a feasible direction} \\ &\quad \text{at } \overline{x} \text{ for } \mathbf{K}, \text{ and a descent} \\ &\quad \text{direction for } \theta(x) \text{ at } \overline{x} \end{aligned} \qquad \text{(iv)}$$

is known as a steepest descent direction for (iii) at the feasible solution $\overline{x}$.

Descent methods

Descent methods for smooth minimization problems
have the following features. They are initiated with a feasible solution, x^0, for the problem, and generate a sequence $\{x^r : r = 0, 1, 2, \ldots\}$ of feasible points. For each r, the objective value at x^{r+1} is strictly less than the objective value at x^r. For $r \geqq 0$, step $r + 1$ of the method consists of the following two substeps.

1. Generate a feasible direction, y^r, for the problem at the present feasible point x^r, which is a descent direction for the objective function.

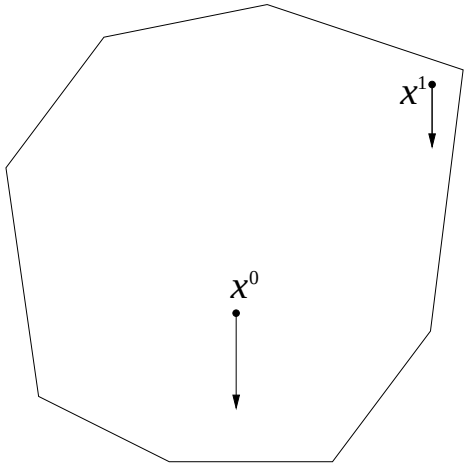
2. Carry out a line search on the half-line $\{x : x = x^r + \lambda y^r, \lambda \geqq 0\}$ for improving the objective value. For this, one has to determine the maximum value of λ, say $\overline{\lambda}$, such that $x^r + \lambda y^r$ remains feasible to the problem for all $0 \leqq \lambda \leqq \overline{\lambda}$ and then solve the line minimization problem of minimizing the objective function over $\{x : x = x^r + \lambda y^r, 0 \leqq \lambda \leqq \overline{\lambda}\}$, the output of which is the next point in the sequence, x^{r+1}.

If there exists no feasible descent direction at x^r, the method terminates with x^r while carrying out substep 1 (unfortunately, this does not guarantee that x^r is even a local minimum for the problem, it just means that we are unable to improve on the point x^r using descent methods.) If subsetp 1 does produce a direction y^r, from the definition of feasible descent directions, $\overline{\lambda}$ is guaranteed to be positive in substep 2 (it may happen that $\overline{\lambda} = \infty$). Different descent methods use different procedures for carrying out substeps 1, 2.

Therefore, the important feature of descent methods is that each move is made along a straight line, and results in a strict improvement in objective value. Since the objective value strictly improves in each step (assuming that the method does not terminate in that step), the sequence of points generated by a descent method is called a **descent sequence**.

Karmarkar's algorithm for LP and an intuitive justification for it

A detailed description of Karmarkar's algorithm, including complete proofs of its polynomial boundedness are provided in Section 11.4. Here we give a statement of this algorithm, with an intuitive justification, for someone interested in an overview without all the technical details and the proofs.
Consider the problem of minimizing a linear function on a convex polytope.



One can improve the current solution substantially by moving in the steepest descent direction, if the current solution is near the center of the feasible region, as in x^0 in the figure given above; but not so if it is near the boundary, as in x^1.
The main ideas behind Karmarkar's algorithm are the following:
i) If the current feasible solution is near the center of the feasible region, it makes sense to move in the steepest descent direction.
ii) If it is possible to transform the problem without changing it in an essential way, that moves the current feasible solution near the center of the feasible region, do it. Karmarkar uses a projective scaling transformation to do exactly this.
A (relative) interior feasible solution to an LP is one which satisfies all inequality constraints as strict inequalities. The basic strategy of Karmarkar's algorithm is to start at a (relative) interior feasible solution, and to carry out a projective scaling transformation to move the current solution to the center.

In the transformed problem, move in the steepest descent direction from this center, but not all the way to the (relative) boundary. Repeat as often as necessary.

Karmarkar considers linear programming problems in the following form

$$\begin{array}{ll} \text{minimize} & cx \\ \text{subject to} & Ax = 0 \\ & e^T x = 1 \\ & x \geqq 0 \end{array} \qquad \text{(P)}$$

where A is a given matrix of order $m \times n$, and e^T is the row vector of all 1's in $\mathbf{R}^n$. The set $\mathbf{S} = \{x : x \in \mathbf{R}^n \text{ and } e^T x = 1, x \geqq 0\}$ is the standard $(n-1)$ dimensional simplex in $\mathbf{R}^n$. The problem (P) is assumed to satisfy the following assumptions.

(1) The point $a^0 = (1/n)e = (1/n, \ldots, 1/n)^T$, the center of $\mathbf{S}$, is feasible to (P).

(2) The problem (P) has an optimum solution, and the optimum objective value in (P) is zero.

Methods for transforming any LP into the form (P) satisfying conditions (1), (2), are discussed in Section 11.4. This is the initialization work before applying Karmarkar's algorithm on an LP. While these initialization methods are simple and mathematically correct, they can ruin the practical efficiency unless done in a clever way. Practically efficient initialization techniques in implementing Karmarkar's algorithm, are the object of intense research investigations at the moment.

Let us now consider the LP (P) satisfying (1) and (2). Karmarkar's method generates a sequence of feasible solutions for (P), $x^0 = a^0, x^1, x^2, \ldots$, all of them in the relative interior of $\mathbf{S}$ (i. e., $x^r > 0$ for all r), with cx^r monotonic decreasing. The method is terminated when we reach a t such that the objective value cx^t is sufficiently close to the optimum objective value of 0. So the terminal solution x^t is a near optimum solution to (P). A pivotal method (needing at most n pivot steps) that leads to an optimum extreme point solution of (P) from a near optimum solution, is discussed in Section 11.4, it

can be used in a final step if necessary. We now provide the general step.

General step $r+1$ in Karmarkar's algorithm: Let $x^r = a = (a_1, \ldots, a_n)^T > 0$ be the current feasible solution of (P). Define D as the $n \times n$ diagonal matrix with diagonal entries $a_1, \ldots, a_n$, that is

$$D = \begin{pmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & a_n \end{pmatrix}.$$

Since the matrix D depends on the current solution, you get a different D in each step. Use the projective transformation $T : \mathbf{S} \to \mathbf{S}$, defining new variables $y = (y_1, \ldots, y_n)$ by

$$y = T(x) = \frac{D^{-1}x}{e^T D^{-1}x}.$$

Since D is a diagonal matrix with positive diagonal entries, D^{-1} is the diagonal matrix whose ith diagonal entry is $(1/a_i)$. For every $x \in \mathbf{S}$, $T(x) \in \mathbf{S}$. Also, points in the relative interior of $\mathbf{S}$ in the x-space map into points in the relative interior of $\mathbf{S}$ in the y-space. The current feasible solution a of (P) in the x-space, maps into the solution $a^0 = (1/n, \ldots, 1/n)$, the center of the simplex $\mathbf{S}$ in the y-space, under this transformation.
To transform the problem (P), we use the inverse transformation

$$x = T^{-1}(y) = \frac{Dy}{e^T Dy}.$$

It can be verified that this transforms the original LP into

$$\begin{aligned} \text{minimize} \quad & \frac{cDy}{e^T Dy} = \theta(y) \\ \text{subject to} \quad & ADy = 0 \\ & e^T y = 1 \\ & y \geqq 0. \end{aligned} \tag{Q}$$

The constraints remain linear and essentially in the

same form as those in (P), but the objective function in (Q) is nonlinear.
Since the current solution for (Q) is a^0, the center of $\mathbf{S}$, it makes sense to move from a^0, in the steepest descent direction in (Q) at a^0. Since $a^0 > 0$, the set of feasible directions for (Q) at a^0 is $\{\xi : \xi \in \mathbf{R}^n, AD\xi = 0, e^T\xi = 0\}$. Let

$$B = \begin{pmatrix} AD \\ \dots \\ e^T \end{pmatrix}.$$

At a^0, the denominator in $\theta(y)$, $e^T Dy$, is equal to $(1/n)$, and it remains quite constant in a small neighborhood of a^0. So, the steepest descent direction for (Q) at the current point a^0 can be approximated by the steepest descent direction for the objective function cDy subject to the same constraints as in (Q), this is the solution of

$$\begin{array}{ll} \text{minimize} & cD\xi \\ \text{subject to} & B\xi = 0 \\ & \|\xi\| = 1. \end{array}$$

The optimum solution of this problem is $\hat{c}_p/\|\hat{c}_p\|$, where

$$\hat{c}_p = cD(I - B^T(BB^T)^{-1}B)$$

$\hat{c}_p$ is the orthogonal projection of cD onto the subspace $\{\xi : B\xi = 0\}$. So, the next point for (Q) is of the form

$$y' = a^0 - \beta\hat{c}_p/\|\hat{c}_p\|$$

where β is a positive step length. β can be chosen as large as possible, but keeping $y' > 0$. This leads to the new solution x^{r+1} for the original problem (P), where

$$x^{r+1} = \frac{Dy'}{e^T Dy'}.$$

If cx^{r+1} is sufficiently close to 0, terminate with cx^{r+1} as a near optimum solution for (P), otherwise, go to the next step with x^{r+1} as the current solution.

Chapter 1

LINEAR COMPLEMENTARITY PROBLEM, ITS GEOMETRY, AND APPLICATIONS

1.1 THE LINEAR COMPLEMENTARITY PROBLEM AND ITS GEOMETRY

The **Linear Complementarity Problem** (abbreviated as LCP) is a general problem which unifies linear and quadratic programs and bimatrix games. The study of LCP has led to many far reaching benefits. For example, an algorithm known as the **complementary pivot algorithm** first developed for solving LCPs, has been generalized in a direct manner to yield efficient algorithms for computing Brouwer and Kakutani fixed points, for computing economic equilibria, and for solving systems of nonlinear equations and nonlinear programming problems. Also, iterative methods developed for solving LCPs hold great promise for handling very large scale linear programs which cannot be tackled with the well known simplex method because of their large size and the consequent numerical difficulties. For these reasons the study of LCP offers rich rewards for people learning or doing research in optimization or engaged in practical applications of optimization. In this book we discuss the LCP in all its depth.

Let M be a given square matrix of order n and q a column vector in $\mathbf{R}^n$. Throughout this book we will use the symbols $w_1, \ldots, w_n$; $z_1, \ldots, z_n$ to denote the variables in the problem. **In an** LCP **there is no objective function to be optimized**. The problem is: find $w = (w_1, \ldots, w_n)^T$, $z = (z_1, \ldots, z_n)^T$ satisfying

$$\begin{aligned} & w - Mz = q \\ & w \geqq 0,\ z \geqq 0 \quad \text{and} \quad w_i z_i = 0 \quad \text{for all } i \end{aligned} \tag{1.1}$$

The only data in the problem is the column vector q and the square matrix M. So we will denote the LCP of finding $w \in \mathbf{R}^n$, $z \in \mathbf{R}^n$ satisfying (1.1) by the symbol (q, M). It is said to be an LCP of **order** n. In an LCP of order n there are $2n$ variables. As a specific example, let $n = 2$, $M = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$, $q = \begin{pmatrix} -5 \\ -6 \end{pmatrix}$. This leads to the LCP

$$\begin{aligned} w_1 \quad & -2z_1 - z_2 = -5 \\ w_2 - z_1 & -2z_2 = -6. \\ w_1, w_2, z_1, z_2 \geqq 0 \quad & \text{and} \quad w_1 z_1 = w_2 z_2 = 0. \end{aligned} \tag{1.2}$$

The problem (1.2) can be expressed in the form of a vector equation as

$$w_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + w_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} + z_1 \begin{pmatrix} -2 \\ -1 \end{pmatrix} + z_2 \begin{pmatrix} -1 \\ -2 \end{pmatrix} = \begin{pmatrix} -5 \\ -6 \end{pmatrix} \tag{1.3}$$

$$w_1, w_2, z_1, z_2 \geqq 0 \quad \text{and} \quad w_1 z_1 = w_2 z_2 = 0 \tag{1.4}$$

In any solution satisfying (1.4), at least one of the variables in each pair (w_j, z_j), has to equal zero. One approach for solving this problem is to pick one variable from each of the pairs (w_1, z_1), (w_2, z_2) and to fix them at zero value in (1.3). The remaining variables in the system may be called **usable variables**. After eliminating the zero variables from (1.3), if the remaining system has a solution in which the usable variables are nonnegative, that would provide a solution to (1.3) and (1.4).

Pick w_1, w_2 as the zero-valued variables. After setting w_1, w_2 equal to 0 in (1.3), the remaining system is

$$\begin{gathered} z_1 \begin{pmatrix} -2 \\ -1 \end{pmatrix} + z_2 \begin{pmatrix} -1 \\ -2 \end{pmatrix} = \begin{pmatrix} -5 \\ -6 \end{pmatrix} = \begin{pmatrix} q_1 \\ q_2 \end{pmatrix} = q \\ z_1 \geqq 0, \quad z_2 \geqq 0 \end{gathered} \tag{1.5}$$

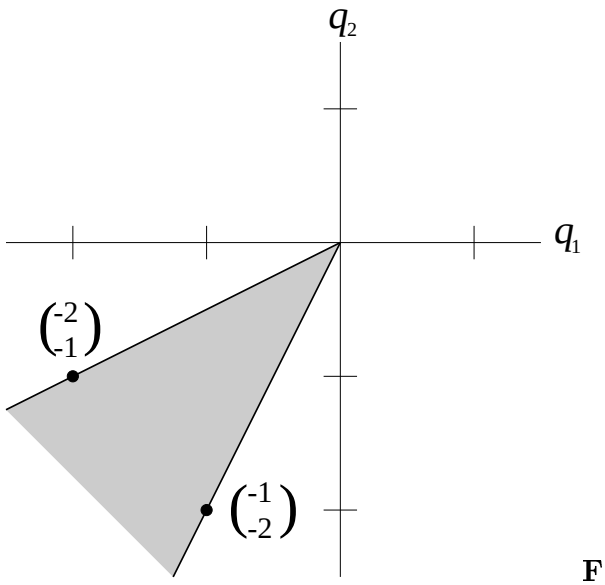


Figure 1.1 A Complementary Cone

Equation (1.5) has a solution iff the vector q can be expressed as a nonnegative linear combination of the vectors $(-2, -1)^T$ and $(-1, -2)^T$. The set of all nonnegative

linear combinations of $(-2,-1)^T$ and $(-1,-2)^T$ is a cone in the q_1, q_2-space as in Figure 1.1. Only if the given vector $q = (-5,-6)^T$ lies in this cone, does the LCP (1.2) have a solution in which the usable variables are z_1, z_2. We verify that the point $(-5,-6)^T$ does lie in the cone, that the solution of (1.5) is $(z_1, z_2) = (4/3, 7/3)$ and, hence, a solution for (1.2) is $(w_1, w_2; z_1, z_2) = (0, 0; 4/3, 7/3)$. The cone in Figure 1.1 is known as a **complementary cone** associated with the LCP (1.2). Complementary cones are generalizations of the well-known class of quadrants or orthants.

1.1.1 Notation

The symbol I usually denotes the unit matrix. If we want to emphasize its order, we denote the unit matrix of order n by the symbol I_n.

We will use the abbreviation LP for "Linear Program" and BFS for "Basic Feasible Solution". See [1.28, 2.26]. LCP is the abbreviation for "Linear Complementarity Problem" and NLP is the abbreviation for "Nonlinear Program".

Column and Row Vectors of a Matrix

If $A = (a_{ij})$ is a matrix of order $m \times n$ say, we will denote its jth column vector $(a_{1j}, \ldots, a_{mj})^T$ by the symbol $A_{.j}$, and its ith row vector $(a_{i1}, \ldots, a_{in})$ by $A_{i.}$.

Nonnegative, Semipositive, Positive Vectors

Let $x = (x_1, \ldots, x_n)^T \in \mathbf{R}^n$. $x \geqq 0$, that is nonnegative, if $x_j \geqq 0$ for all j. Clearly, $0 \geqq 0$. x is said to be semipositive, denoted by $x \geq 0$, if $x_j \geqq 0$ for all j and at least one $x_j > 0$. Notice the distinction in the symbols for denoting nonnegative ($\geqq$ with two lines under the $>$) and semipositive ($\geq$ with only a single line under the $>$). $0 \ngeq 0$, the zero vector is the only nonnegative vector which is not semipositive. Also, if $x \geq 0$, $\sum_{j=1}^{n} x_j > 0$. The vector $x > 0$, strictly positive, if $x_j > 0$ for all j. Given two vectors $x, y \in \mathbf{R}^n$; we write $x \geqq y$, if $x - y \geqq 0$, $x \geq y$ if $x - y \geq 0$, and $x > y$ if $x - y > 0$.

Pos Cones

If $\{x^1, \ldots, x^r\} \subset \mathbf{R}^n$, the cone $\{x : x = \alpha_1 x^1 + \ldots + \alpha_r x^r, \alpha_1, \ldots, \alpha_r \geqq 0\}$ is denoted by $\text{Pos}\{x^1, \ldots, x^r\}$. Given the matrix A of order $m \times n$, $\text{Pos}(A)$ denotes the cone $\text{Pos}\{A_{.1}, \ldots, A_{.n}\} = \{x : x = A\alpha \text{ for } \alpha = (\alpha_1, \ldots, \alpha_n)^T \geqq 0\}$.

Directions, Rays, Half-Lines, and Step Length

Any point $y \in \mathbf{R}^n$, $y \neq 0$, defines a direction in $\mathbf{R}^n$. Given the direction y, it's ray is the half-line obtained by joining the origin 0 to y and continuing indefinitely in the

same direction, it is the set of points $\{\alpha y : \alpha \geqq 0\}$. Given $\overline{x} \in \mathbf{R}^n$, by moving from $\overline{x}$ in the direction y we get points of the form $\overline{x} + \alpha y$ where $\alpha \geqq 0$, and the set of all such points $\{\overline{x} + \alpha y : \alpha \geqq 0\}$ is the **halfline** or **ray** through $\overline{x}$ in the direction y. The point $\overline{x} + \alpha y$ for $\alpha > 0$ is said to have been obtained by moving from $\overline{x}$ in the direction y a step length of α. As an example, if $y = (1, 1)^T \in \mathbf{R}^n$, the ray of y is the set of all points of the form $\{(\alpha, \alpha)^T : \alpha \geqq 0\}$. In addition, if, $\overline{x} = (1, -1)^T$, the halfline through $\overline{x}$ in the direction y is the set of all points of the form $\{(1 + \alpha, -1 + \alpha)^T : \alpha \geqq 0\}$. See Figure 1.2. In this half-line, letting $\alpha = 9$, we get the point $(10, 8)^T$, and this point is obtained by taking a step of length 9 from $\overline{x} = (1, -1)^T$ in the direction $y = (1, 1)^T$.

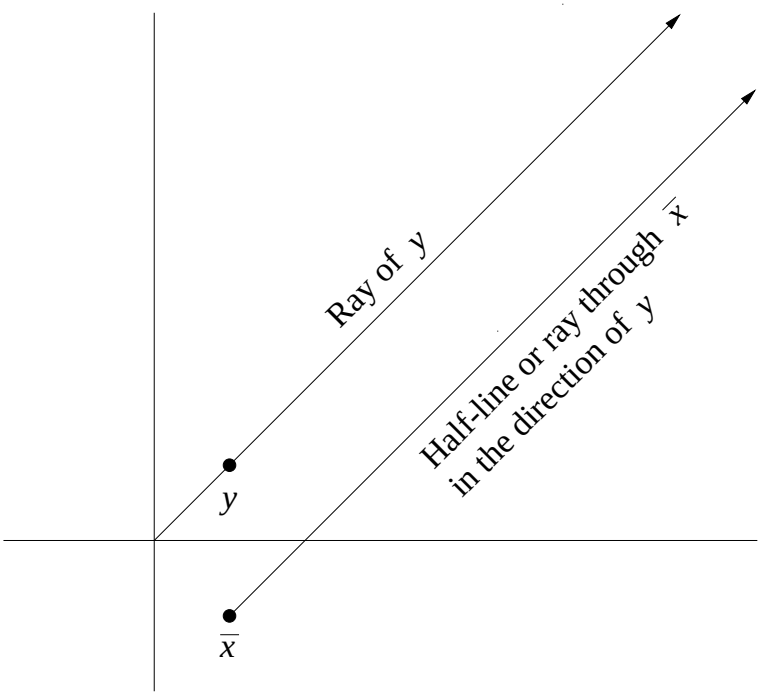


Figure 1.2 Rays and Half-Lines

1.1.2 Complementary Cones

In the LCP (q, M), the complementary cones are defined by the matrix M. The point q does not play any role in the definition of complementary cones.

Let M be a given square matrix of order n. For obtaining $\mathcal{C}(M)$, the class of complementary cones corresponding to M, the pair of column vectors $(I_{.j}, -M_{.j})$ is

known as the jth **complementary pair of vectors**, $1 \leqq j \leqq n$. Pick a vector from the pair $(I_{.j}, -M_{.j})$ and denote it by $A_{.j}$. The ordered set of vectors $(A_{.1}, \ldots, A_{.n})$ is known as a **complementary set of vectors**. The cone $\text{Pos}(A_{.1}, \ldots, A_{.n}) = \{y : y = \alpha_1 A_{.1} + \ldots + \alpha_n A_{.n}; \alpha_1 \geqq 0, \ldots, \alpha_n \geqq 0\}$ is known as a **complementary cone** in the class $\mathcal{C}(M)$. Clearly there are 2^n complementary cones.

Example 1.1

Let $n = 2$ and $M = I$. In this case, the class $\mathcal{C}(I)$ is just the class of orthants in $\mathbf{R}^2$. In general for any n, $\mathcal{C}(I)$ is the class of orthants in $\mathbf{R}^n$. Thus the class of complementary cones is a generalization of the class of orthants. See Figure 1.3. Figures 1.4 and 1.5 provide some more examples of complementary cones. In the example in Figure 1.5 since $\{I_{.1}, -M_{.2}\}$ is a linearly dependent set, the cone $\text{Pos}(I_{.1}, -M_{.2})$ has an empty interior. It consists of all the points on the horizontal axis in Figure 1.6 (the thick axis). The remaining three complementary cones have nonempty interiors.

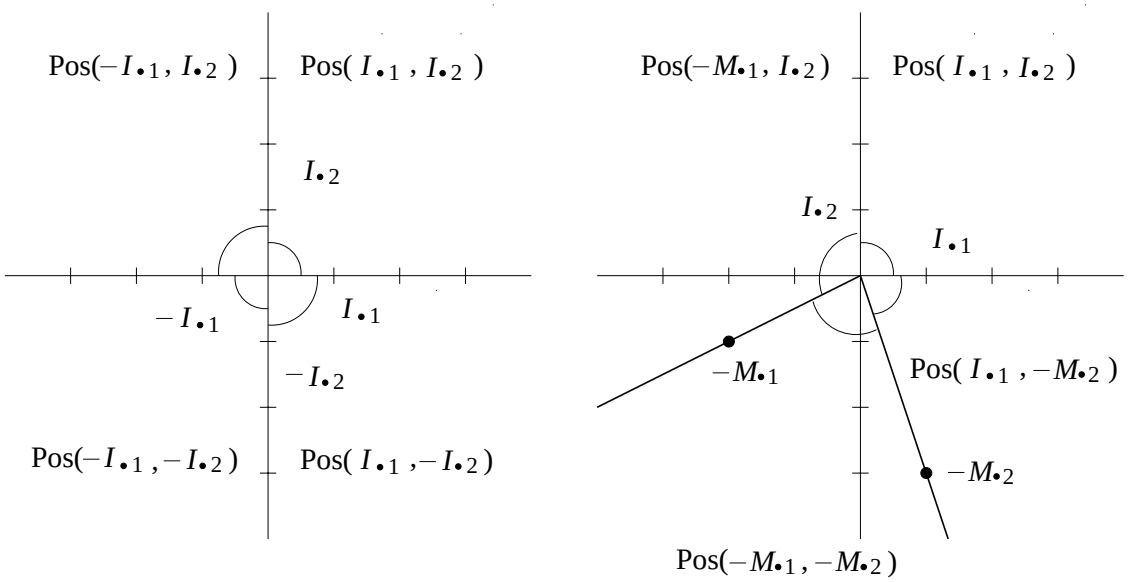


Figure 1.3 When $M = I$, the Complementarity Cones are the Orthants.

Figure 1.4 Complementary Cones when $M = \begin{pmatrix} 2 & -1 \\ 1 & 3 \end{pmatrix}$.

Degenerate, Nondegenerate Complementary Cones

Let $\text{Pos}(A_{.1}, \ldots, A_{.n})$ be a complementary cone in $\mathcal{C}(M)$. This cone is said to be a nondegenerate complementary cone if it has a nonempty interior, that is if $\{A_{.1}, \ldots, A_{.n}\}$ is a linearly independent set; degenerate complementary cone if its interior is empty, which happens when $\{A_{.1}, \ldots, A_{.n}\}$ is a linearly dependent set. As examples, all the complementary cones in Figures 1.3, 1.4, 1.5, are nondegenerate. In Figure 1.6 the complementary cone $\text{Pos}(I_{.1}, -M_{.2})$ is degenerate, the remaining three complementary cones are nondegenerate.

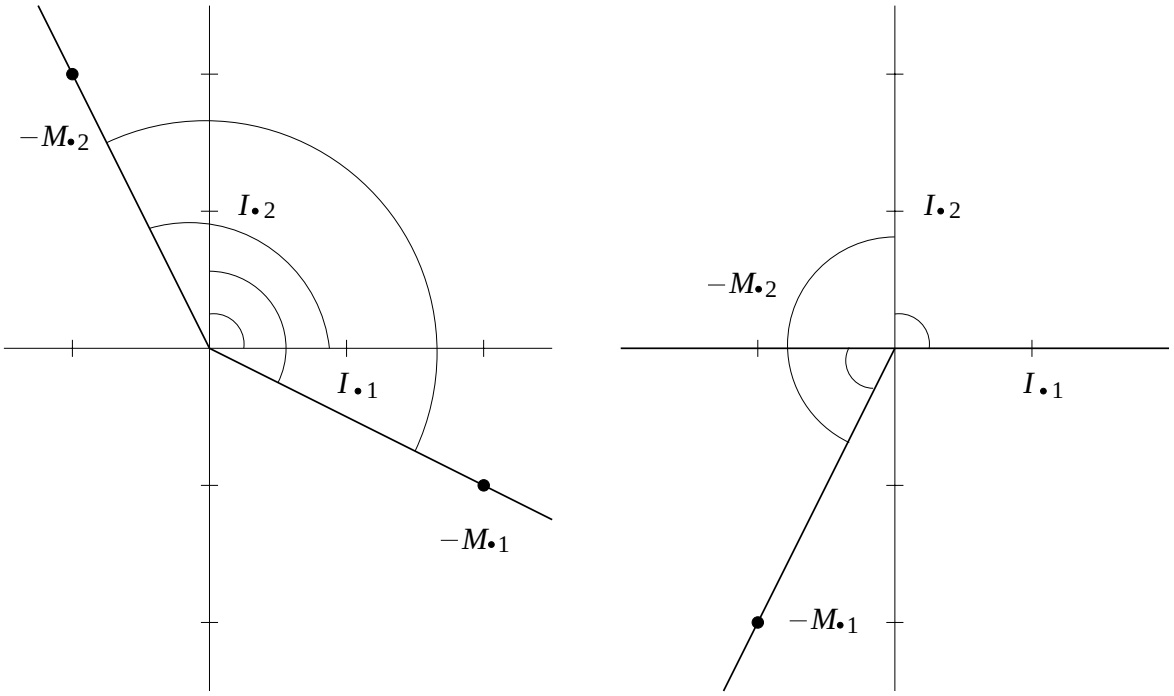


Figure 1.5 Complementary Cones when $M = \begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix}$.

Figure 1.6 Complementary Cones when $M = \begin{pmatrix} 1 & 1 \\ 2 & 0 \end{pmatrix}$.

1.1.3 The Linear Complementary Problem

Given the square matrix M of order n and the column vector $q \in \mathbf{R}^n$, the LCP (q, M), is equivalent to the problem of finding a complementary cone in $\mathcal{C}(M)$ that contains the point q, that is, to find a complementary set of column vectors $(A_{.1}, \ldots, A_{.n})$ such that

(i) $A_{.j} \in \{I_{.j}, -M_{.j}\}$ for $1 \leqq j \leqq n$

(ii) q can be expressed as a nonnegative linear combination of $(A_{.1}, \ldots, A_{.n})$

where I is the identity matrix of order n and $I_{.j}$ is its jth column vector. This is equivalent to finding $w \in \mathbf{R}^n$, $z \in \mathbf{R}^n$ satisfying $\sum_{j=1}^{n} I_{.j} w_j - \sum_{j=1}^{n} M_{.j} z_j = q, w_j \geqq 0$, $z_j \geqq 0$ for all j, and either $w_j = 0$ or $z_j = 0$ for all j. In matrix notation this is

$$w - Mz = q \tag{1.6}$$

$$w \geqq 0 \quad z \geqq 0 \tag{1.7}$$

$$w_j z_j = 0 \quad \text{for all } j. \tag{1.8}$$

Because of (1.7), the condition (1.8) is equivalent to $\sum_{j=1}^{n} w_j z_j = w^T z = 0$; this condition is known as the **complementarity constraint**. In any solution of the LCP (q, M), if one of the variables in the pair (w_j, z_j) is positive, the other should be zero. Hence, the pair (w_j, z_j) is known as the jth **complementary pair of variables** and each variable in this pair is the **complement** of the other. In (1.6) the column vector corresponding to w_j is $I_{.j}$, and the column vector corresponding to z_j is $-M_{.j}$. For $j = 1$ to n, the pair $(I_{.j}, -M_{.j})$ is the jth complementary pair of column vectors in the LCP (q, M). For $j = 1$ to n, let $y_j \in \{w_j, z_j\}$ and let $A_{.j}$ be the column vector corresponding to y_j in (1.6). So $A_{.j} \in \{I_{.j} - M_{.j}\}$. Then $y = (y_1, \ldots, y_n)$ is a **complementary vector of variables** in this LCP, the ordered set $(A_{.1}, \ldots, A_{.n})$ is the **complementary set of column vectors corresponding to it** and the matrix A with its column vectors as $A_{.1}, \ldots, A_{.n}$ in that order is known as the **complementary matrix** corresponding to it. If $\{A_{.1}, \ldots, A_{.n}\}$ is linearly independent, y is a **complementary basic vector of variables** in this LCP, and the complementary matrix A whose column vectors are $A_{.1}, \ldots, A_{.n}$ in that order, is known as the **complementary basis** for (1.6) corresponding to the complementary basic vector y. The cone $\text{Pos}(A_{.1}, \ldots, A_{.n}) = \{x : x = \alpha_1 A_{.1} + \ldots + \alpha_n A_{.n},\ \alpha_1 \geqq 0, \ldots, \alpha_n \geqq 0\}$ is the complementary cone in the class $\mathcal{C}(M)$ corresponding to the complementary set of column vectors $(A_{.1}, \ldots, A_{.n})$, or the associated complementary vector of variables y. A **solution of the LCP** (q, M), always means a $(w; z)$ satisfying all the constraints (1.6), (1.7), (1.8).

A **complementary feasible basic vector** for this LCP is a complementary basic vector satisfying the property that q can be expressed as a nonnegative combination of column vectors in the corresponding complementary basis. Thus each complementary feasible basic vector leads to a solution of the LCP.

The union of all the complementary cones associated with the square matrix M is denoted by the symbol $\mathbf{K}(M)$. $\mathbf{K}(M)$ is clearly the set of all vectors q for which the LCP (q, M) has at least one solution.

We will say that the vector $\overline{z}$ **leads to a solution** of the LCP (q, M) iff $(\overline{w} = M\overline{z} + q, \overline{z})$ is a solution of this LCP.

As an illustration, here are all the complementary vectors of variables and the corresponding complementary matrices for (1.2), an LCP of order 2.

Complementary vector of variables	The corresponding complementary matrix
(w_1, w_2)	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$
(w_1, z_2)	$\begin{pmatrix} 1 & -1 \\ 0 & -2 \end{pmatrix}$
(z_1, w_2)	$\begin{pmatrix} -2 & 0 \\ -1 & 1 \end{pmatrix}$
(z_1, z_2)	$\begin{pmatrix} -2 & -1 \\ -1 & -2 \end{pmatrix}$

Since each of these complementary matrices is nonsingular, all the complementary vectors are complementary basic vectors, and all the complementary matrices are complementary bases, in this LCP. Since $q = (-5, -6)^T$ in (1.2) can be expressed as a nonnegative combination of the complementary matrix corresponding to (z_1, z_2); (z_1, z_2) is a complementary feasible basic vector for this LCP. The reader should draw all the complementary cones corresponding to this LCP on the two dimensional Cartesian plane, and verify that for this LCP, their union, the set $\mathbf{K}(M) = \mathbf{R}^2$.

The Total Enumeration Method for the LCP

Consider the LCP (q, M) of order n. The complementarity constraint (1.8) implies that in any solution (w, z) of this LCP, for each $j = 1$ to n, we must have

$$\text{either } w_j = 0$$
$$\text{or } \ z_j = 0.$$

This gives the LCP a combinatorial, rather than nonlinear flavour. It automatically leads to an enumeration method for the LCP.

There are exactly $2^{\mathbf{n}}$ complementary vectors of variables. Let

$$y^r = (y_1^r, \ldots, y_n^r), \quad r = 1 \ \text{ to } \ 2^{\mathbf{n}}$$

where $y_j^r \in \{w_j, z_j\}$ for each $j = 1$ to n, be all the complementary vectors of variables. Let A_r be the complementary matrix corresponding to y^r, $r = 1$ to $2^{\mathbf{n}}$. Solve the following system (P_r).

$$\begin{aligned} A_r y^r &= q \\ y^r &\geqq 0 \ . \end{aligned} \tag{P_r}$$

This system can be solved by Phase I of the simplex method for LP, or by other methods for solving linear equality and inequality systems. If this system has a feasible solution, $\overline{y}^r$, say, then

$$y^r = \overline{y}^r$$
$$\text{all variables not in } y^r, \text{ equal to zero}$$

is a solution of LCP (q, M). If the complementary matrix A_r is singular, the system (P_r) may have no feasible solution, or have one or an infinite number of feasible solutions. Each feasible solution of (P_r) leads to a solution of the LCP (q, M) as discussed above. When this is repeated for $r = 1$ to 2^n, all solutions of the LCP (q, M) can be obtained. The method discussed at the beginning of Section 1.1 to solve an LCP of order 2 is exactly this enumeration method.

This enumeration method is convenient to use only when $n = 2$, since $2^2 = 4$ is small; and to check whether the system (P_r) has a solution for any r, we can draw the corresponding complementary cone in the two dimensional Cartesian plane and check whether it contains q. When $n > 2$, particularly for large n, this enumeration method becomes impractical since 2^n grows very rapidly. In Chapter 2 and later chapters we discuss efficient pivotal and other methods for solving special classes of LCPs that arise in several practical applications. In Section 8.7 we show that the general LCP is a hard problem. At the moment, the only known algorithms which are guaranteed to solve the general LCP are enumerative methods, see Section 11.3.

1.2 APPLICATION TO LINEAR PROGRAMMING

In a general LP there may be some inequality constraints, equality constraints, sign restricted variables and unrestricted variables. Transform each lower bounded variable, say $x_j \geqq l_j$, into a nonnegative variable by substituting $x_j = l_j + y_j$ where $y_j \geqq 0$. Transform each sign restricted variable of the form $x_j \leqq 0$ into a nonnegative variable by substituting $x_j = -y_j$ where $y_j \geqq 0$. Eliminate the unrestricted variables one after the other, using the equality constraints (see Chapter 2 of [1.28 or 2.26]). In the resulting system, if there is still an equality constraint left, eliminate a nonnegative variable from the system using it, thereby transforming the constraint into an inequality constraint in the remaining variables. Repeat this process until there are no more equality constraints. In the resulting system, transform any inequality constraint of the "$\leqq$" form into one of "$\geqq$" form, by multiplying both sides of it by '-1'. If the objetive function is to be maximized, replace it by its negative which should be minimized, and eliminate any constant terms in it. When all this work is completed, the original LP is transformed into:

$$\begin{array}{lll} \text{Minimize} & cx & \\ \text{Subject to} & Ax & \geqq b \\ & x & \geqq 0 \end{array} \tag{1.9}$$

which is in **symmetric form**. Here, suppose A is of order $m \times N$. If x is an optimum feasible solution of (1.9), by the results of the duality theory of linear programming (see [1.28, 2.26]) there exists a dual vector $y \in \mathbf{R}^m$, primal slack vector $v \in \mathbf{R}^m$, and dual slack vector $u \in \mathbf{R}^N$ which together satisfy

$$\begin{pmatrix} u \\ v \end{pmatrix} - \begin{pmatrix} 0 & -A^T \\ A & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} c^T \\ -b \end{pmatrix}$$
$$\begin{pmatrix} u \\ v \end{pmatrix} \geqq 0 \quad \begin{pmatrix} x \\ y \end{pmatrix} \geqq 0 \quad \text{and} \quad \begin{pmatrix} u \\ v \end{pmatrix}^T \begin{pmatrix} x \\ y \end{pmatrix} = 0 \ . \tag{1.10}$$

Conversely, if u, v, x, y together satisfy all the conditions in (1.10), x is an optimum solution of (1.9). In (1.10) all the vectors and matrices are written in **partitioned form**. For example, $\binom{u}{v}$ is the vector $(u_1, \ldots, u_N, v_1, \ldots, v_m)^T$. If $n = m + N$,

$$w = \begin{pmatrix} u \\ v \end{pmatrix} , \ z = \begin{pmatrix} x \\ y \end{pmatrix} , \ M = \begin{pmatrix} 0 & -A^T \\ A & 0 \end{pmatrix} , \ q = \begin{pmatrix} c^T \\ -b \end{pmatrix} ,$$

(1.10) is seen to be an LCP of order n of the type (1.6) to (1.8). Solving the LP (1.9) can be achieved by solving the LCP (1.10).

Also, the various complementary pairs of variables in the LCP (1.10) are exactly those in the pair of primal, dual LPs (1.9) and its dual. As an example consider the following LP.

$$\begin{array}{lrcl} \text{Minimize} & -13x_1 + 42x_2 & & \\ \text{Subject to} & 8x_1 - x_2 + 3x_2 & \geqq & -16 \\ & -3x_1 + 2x_2 - 13x_3 & \geqq & 12 \\ & x_j & \geqq & 0, \ j = 1, 2, 3. \end{array}$$

Let (v_1, y_1), (v_2, y_2) denote the nonnegative slack variable, dual variable respectively, associated with the two primal constraints in that order. Let u_1, u_2, u_3 denote the nonnegative dual slack variable associated with the dual constraint corresponding to the primal variable x_1, x_2, x_3, in that order. Then the primal and dual systems together with the complementary slackness conditions for optimality are

$$\begin{array}{rrrrrrr} 8x_1 - & x_2 + & 3x_3 - & v_1 & & = & -16 \\ -3x_1 + & 2x_2 - & 13x_3 & & - v_2 & = & 12 \\ 8y_1 - & 3y_2 & & + u_1 & & = & -13 \\ -y_1 + & 2y_2 & & & + u_2 & = & 42 \\ 3y_1 - & 13y_2 & & & + u_3 & = & 0. \end{array}$$

$$x_j, u_j, y_i, v_i \geqq 0 \ \text{ for all } i, j.$$
$$x_j u_j = y_i v_i = 0 \ \text{ for all } i, j.$$

This is exactly the following LCP.

u_1	u_2	u_3	v_1	v_2	x_1	x_2	x_3	y_1	y_2	
1	0	0	0	0	0	0	0	8	−3	−13
0	1	0	0	0	0	0	0	−1	2	42
0	0	1	0	0	0	0	0	3	−13	0
0	0	0	1	0	−8	1	−3	0	0	16
0	0	0	0	1	3	−2	13	0	0	−12

All variables $\geqq 0$. $u_1x_1 = u_2x_2 = u_3x_3 = v_1y_1 = v_2y_2 = 0$.

1.3 QUADRATIC PROGRAMMING

Using the methods discussed in Section 1.2 any problem in which a quadratic objective function has to be optimized subject to linear equality and inequality constraints can be transformed into a problem of the form

$$\begin{array}{lll} \text{Minimize} & Q(x) = cx + \frac{1}{2}x^T Dx & \\ \text{Subject to} & Ax \geqq b & \\ & x \geqq 0 & \end{array} \tag{1.11}$$

where A is a matrix of order $m \times N$, and D is a **square symmetric matrix of order** N. There is no loss of generality in assuming that D is a symmetric matrix, because if it is not symmetric replacing D by $(D + D^T)/2$ (which is a symmetric matrix) leaves $Q(x)$ unchanged. **We assume that D is symmetric**.

1.3.1 Review on Positive Semidefinite Matrices

A square matrix $F = (f_{ij})$ of order n, whether it is symmetric or not, is said to be a **positive semidefinite matrix** if $y^T Fy \geqq 0$ for all $y \in \mathbf{R}^n$. It is said to be a **positive definite matrix** if $y^T Fy > 0$ for all $y \neq 0$. We will use the abbreviations PSD, PD for "positive semidefinite" and "positive definite", respectively.

Principal Submatrices, Principal Subdeterminants

Let $F = (f_{ij})$ be a square matrix of order n. Let $\{i_1, \ldots, i_r\} \subset \{1, \ldots, n\}$ with its elements arranged in increasing order. Erase all the entries in F in row i and column i for each $i \notin \{i_1, \ldots, i_r\}$. What remains is a square submatrix of F of order r:

$$\begin{pmatrix} f_{i_1,i_1} & \dots & f_{i_1,i_r} \\ \vdots & & \vdots \\ f_{i_r,i_1} & \dots & f_{i_r,i_r} \end{pmatrix} .$$

This submatrix is known as the **principal submatrix** of F determined by the subset $\{i_1, \dots, i_r\}$. Denoting the subset $\{i_1, \dots, i_r\}$ by $\mathbf{J}$, we denote this principal submatrix by the symbol $F_{\mathbf{JJ}}$. It is $(f_{ij} : i \in \mathbf{J}, j \in \mathbf{J})$. The determinant of this principal submatrix is called the principal subdeterminant of F determined by the subset $\mathbf{J}$. The principal submatrix of F determined by ϕ, the empty set, is the empty matrix which has no entries. Its determinant is defined by convention to be equal to 1. The principal submatrix of F determined by $\{1, \dots, n\}$ is F itself. The principal submatrices of F determined by nonempty subsets of $\{1, \dots, n\}$ are **nonempty principal submatrices** of F. Since the number of distinct nonempty subsets of $\{1, \dots, n\}$ is $2^{\mathbf{n}} - 1$, there are $2^{\mathbf{n}} - 1$ nonempty principal submatrices of F. The principal submatrices of F determined by proper subsets of $\{1, \dots, n\}$ are known as **proper principal submatrices** of F. So each proper principal submatrix of F is of order $\leqq n - 1$.

Example 1.2

Let

$$F = \begin{pmatrix} 0 & -1 & 2 \\ 1 & 3 & 4 \\ 1 & 5 & -3 \end{pmatrix} .$$

The principal submatrix corresponding to the subset $\{1, 3\}$ is $\begin{pmatrix} 0 & 2 \\ 1 & -3 \end{pmatrix}$. The principal submatrix corresponding to the subset $\{2\}$ is 3, the second element in the principal diagonal of F.

Several results useful in studying P(S)D matrices will now be discussed.

Results on P(S)D *Matrices*

Result 1.1 If $B = (b_{11})$ is a matrix of order 1×1, it is PD iff $b_{11} > 0$, and it is PSD iff $b_{11} \geqq 0$.

Proof. Let $y = (y_1) \in \mathbf{R}^1$. Then $y^T B y = b_{11} y_1^{\mathbf{2}}$. So $y^T B y > 0$ for all $y \in \mathbf{R}^1$, $y \neq 0$, iff $b_{11} > 0$, and hence B is PD iff $b_{11} > 0$. Also $y^T B y \geqq 0$ for all $y \in \mathbf{R}^1$, iff $b_{11} \geqq 0$, and hence B is PSD iff $b_{11} \geqq 0$.

□

Result 1.2 If F is a PD matrix all its principal submatrices must also be PD.

Proof. Consider the principal submatrix, G, generated by the subset $\{1, 2\}$.

$$G = \begin{pmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{pmatrix} . \quad \text{Let } t = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} .$$

Pick $y = (y_1, y_2, 0, 0, \ldots, 0)^T$. Then $y^T F y = t^T G t$. However, since F is PD, $y^T F y > 0$ for all $y \neq 0$. So $t^T G t > 0$ for all $t \neq 0$. Hence, G is PD too. A similar argument can be used to prove that every principal submatrix of F is also PD.

□

Result 1.3 If F is PD, $f_{ii} > 0$ for all i. This follows as a corollary of Result 1.2.

Result 1.4 If F is a PSD matrix, all principal submatrices of F are also PSD. This is proved using arguments similar to those in Result 1.2.

Result 1.5 If F is PSD matrix, $f_{ii} \geqq 0$ for all i. This follows from Result 1.4.

Result 1.6 Suppose F is a PSD matrix. If $f_{ii} = 0$, then $f_{ij} + f_{ji} = 0$ for all j.

Proof. To be specific let f_{11} be 0 and suppose that $f_{12} + f_{21} \neq 0$. By Result 1.4 the principal submatrix

$$\begin{pmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{pmatrix} = \begin{pmatrix} 0 & f_{12} \\ f_{21} & f_{22} \end{pmatrix}$$

must be PSD. Hence $f_{22}y_2^2 + (f_{12} + f_{21})y_1 y_2 \geqq 0$ for all y_1, y_2. Since $f_{12} + f_{21} \neq 0$, take $y_1 = (-f_{22} - 1)/(f_{12} + f_{21})$ and $y_2 = 1$. The above inequality is violated since the left-hand side becomes equal to -1, leading to a contradiction.

□

Result 1.7 If D is a symmetric PSD matrix and $d_{ii} = 0$, then $D_{.i} = D_{i.} = 0$. This follows from Result 1.6.

Definition: The Gaussian Pivot Step
Let $A = (a_{ij})$ be a matrix of order $m \times n$. A Gaussian pivot step on A, with row r as the pivot row and column s as the pivot column can only be carried out if the element lying in both of them, a_{rs}, is nonzero. This element a_{rs} is known as the pivot element for this pivot step. The pivot step subtracts suitable multiples of the pivot row from each row i for $i > r$ so as to transform the entry in this row and the pivot column into zero. Thus this pivot step transforms

$$A = \begin{pmatrix} a_{11} & \ldots & a_{1s} & \ldots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{r1} & \ldots & a_{rs} & \ldots & a_{rn} \\ a_{r+1,1} & \ldots & a_{r+1,s} & \ldots & a_{r+1,n} \\ \vdots & & \vdots & & \vdots \\ a_{m1} & \ldots & a_{ms} & \ldots & a_{mn} \end{pmatrix}$$

$$\text{into} \quad \begin{pmatrix} a_{11} & \ldots & a_{1s} & \ldots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{r1} & \ldots & a_{rs} & \ldots & a_{rn} \\ a'_{r+1,1} & \ldots & 0 & \ldots & a'_{r+1,n} \\ \vdots & & \vdots & & \vdots \\ a'_{m1} & \ldots & 0 & \ldots & a'_{mn} \end{pmatrix}$$

where $a'_{ij} = a_{ij} - (a_{rj}a_{is})/a_{rs}$, for $i = r+1$ to m, $j = 1$ to n. As an example consider the Gaussian pivot step in the following matrix with row 2 as the pivot row and column 3 as the pivot column. The pivot element is inside a box.

$$\begin{pmatrix} 1 & -2 & 10 & -4 & -1 \\ 4 & 6 & \boxed{2} & -8 & -4 \\ -3 & 1 & -1 & 2 & 3 \\ 1 & -4 & 2 & 3 & 0 \end{pmatrix}$$

This Gaussian pivot step transforms this matrix into

$$\begin{pmatrix} 1 & -2 & 10 & -4 & -1 \\ 4 & 6 & 2 & -8 & -4 \\ -1 & 4 & 0 & -2 & 1 \\ -3 & -10 & 0 & 11 & 4 \end{pmatrix}$$

Result 1.8 Let D be a square symmetric matrix of order $n \geqq 2$. Suppose D is PD. Subtract suitable multiples of row 1 from each of the other rows so that all the entries in column 1 except the first is transformed into zero. That is, transform

$$D = \begin{bmatrix} d_{11} & \dots & d_{1n} \\ d_{21} & \dots & d_{2n} \\ \vdots & & \vdots \\ d_{n1} & \dots & d_{nn} \end{bmatrix} \quad \text{into} \quad D_1 = \begin{bmatrix} d_{11} & & \dots & d_{1n} \\ 0 & \tilde{d}_{22} & \dots & \tilde{d}_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & \tilde{d}_{n2} & \dots & \tilde{d}_{nn} \end{bmatrix}$$

by a Gaussian pivot step with row 1 as pivot row and column 1 as pivot column, clearly $\tilde{d}_{ij} = d_{ij} - d_{1j}d_{i1}/d_{11}$ for all $i, j \geqq 2$. E_1, the matrix obtained by striking off the first row and the first column from D_1, is also symmetric and PD.

Also, if D is an arbitrary square symmetric matrix, it is PD iff $d_{11} > 0$ and the matrix E_1 obtained as above is PD.

Proof. Since D is symmetric $d_{ij} = d_{ji}$ for all i, j. Therefore,

$$y^T D y = \sum_{i=1}^{n}\sum_{j=1}^{n} y_i y_j d_{ij} = d_{11}y_1^2 + 2y_1\sum_{j=2}^{n} d_{1j}y_j + \sum_{i,j\geqq 2} y_i y_j d_{ij}$$

$$= d_{11}\Big(y_1 + \Big(\sum_{j=2}^{n} d_{1j}y_j\Big)/d_{11}\Big)^2 + \sum_{i,j\geqq 2} y_i \tilde{d}_{ij} y_j \ .$$

Letting $y_1 = -(\sum_{j=2}^{n} d_{1j}y_j)/d_{11}$, we verify that if D is PD, then $\sum_{i,j\geqq 2} y_i \tilde{d}_{ij} y_j > 0$ for all $(y_2, \dots, y_n) \neq 0$, which implies that E_1 is PD. The fact that E_1 is also symmetric is clear since $\tilde{d}_{ij} = d_{ij} - d_{1j}d_{i1}/d_{11} = \tilde{d}_{ji}$ by the symmetry of D. If D is an arbitrary symmetric matrix, the above equation clearly implies that D is PD iff $d_{11} > 0$ and E_1 is PD.

□

Result 1.9 A square matrix F is PD (or PSD) iff $F + F^T$ is PD (or PSD).

Proof. This follows because $x^T(F + F^T)x = 2x^TFx$.

□

Result 1.10 Let F be a square matrix of order n and E a matrix of order $m \times n$. The square matrix $A = \begin{pmatrix} F & -E^T \\ E & 0 \end{pmatrix}$ of order $(m+n)$ is PSD iff F is PSD.

Proof. Let $\xi = (y_1, \ldots, y_n, t_1, \ldots, t_m)^T \in \mathbf{R}^{n+m}$ and $y = (y_1, \ldots, y_n)^T$. For all ξ, we have $\xi^T A \xi = y^T F y$. So $\xi^T A \xi \geqq 0$ for all $\xi \in \mathbf{R}^{n+m}$ iff $y^T F y \geqq 0$ for all $y \in \mathbf{R}^n$. That is, A is PSD iff F is PSD.

□

Result 1.11 If B is a square nonsingular matrix of order n, $D = B^TB$ is PD and symmetric.

Proof. The symmetry follows because $D^T = D$. For any $y \in \mathbf{R}^n$, $y \neq 0$, $y^TDy = y^TB^TBy = \|yB\|^2 > 0$ since $yB \neq 0$ (because B is nonsingular, $y \neq 0$ implies $yB \neq 0$). So D is PD.

□

Result 1.12 If A is any matrix of order $m \times n$, A^TA is PSD and symmetric.

Proof. Similar to the proof of Result 1.11.

□

Principal Subdeterminants of PD, PSD *Matrices*

We will need the following theorem from elementary calculus.

Theorem 1.1 Intermediate value theorem: *Let $f(\lambda)$ be a continuous real valued function defined on the closed interval $\lambda_0 \leqq \lambda \leqq \lambda_1$ where $\lambda_0 < \lambda_1$. Let $\overline{f}$ be a real number strictly between $f(\lambda_0)$ and $f(\lambda_1)$. Then there exists a $\overline{\lambda}$ satisfying $\lambda_0 < \overline{\lambda} < \lambda_1$, and $f(\overline{\lambda}) = \overline{f}$.*

□

For a proof of Theorem 1.1 see books on calculus, for example, W. Rudin, *Principles of Mathematical Analysis*, McGraw-Hill, second edition, 1964, p. 81. Theorem 1.1 states that a continuous real valued function defined on a closed interval, assumes all intermediate values between its initial and final values in this interval.

Now we will resume our discussion of PD, PSD matrices.

Theorem 1.2 *If F is a* PD *matrix, whether it is symmetric or not, the determinant of F is strictly positive.*

Proof. Let F be of order n. Let I be the identity matrix of order n. If the determinant of F is zero, F is singular, and hence there exists a nonzero column vector $x \in \mathbf{R}^n$ such that $x^TF = 0$, which implies that $x^TFx = 0$, a contradiction to the hypothesis that F

is PD. So the determinant of F is nonzero. In a similar manner we conclude that the determinant of any PD-matrix is nonzero. For $0 < \lambda < 1$, define $F(\lambda) = \lambda F + (1-\lambda)I$, and $f(\lambda) =$ determinant of $F(\lambda)$.

Obviously $f(\lambda)$ is a polynomial in λ, and hence $f(\lambda)$ is a real valued continuous function defined on the interval $0 \leqq \lambda \leqq 1$. Given a column vector $x \in \mathbf{R}^n$, $x \neq 0$, $x^T F(\lambda) x = \lambda x^T F x + (1-\lambda) x^T x > 0$ for all $0 \leqq \lambda \leqq 1$ because F is PD. So $F(\lambda)$ is a PD matrix for all $0 \leqq \lambda \leqq 1$. So from the above argument $f(\lambda) \neq 0$ for any λ satisfying $0 \leqq \lambda \leqq 1$. Clearly, $f(0) = 1$, and $f(1) =$ determinant of F. If $f(1) < 0$ by Theorem 1.1 there exists a $\overline{\lambda}$ satisfying $0 < \overline{\lambda} < 1$ and $f(\overline{\lambda}) = 0$, a contradiction. Hence $f(1) \not< 0$. Hence the determinant of F cannot be negative. Also it is nonzero. Hence the determinant of F is strictly positive.

□

Theorem 1.3 *If F is a* PD *matrix, whether it is symmetric or not, all principal subdeterminants of F are strictly positive.*

Proof. This follows from Result 1.2 and Theorem 1.2.

□

Theorem 1.4 *If F is a* PSD *matrix, whether it is symmetric or not, its determinant is nonnegative.*

Proof. For $0 \leqq \lambda \leqq 1$, define $F(\lambda)$, $f(\lambda)$ as in the proof of Theorem 1.2. Since I is PD, and F is PSD; $F(\lambda)$ is a PD matrix for $0 \leqq \lambda < 1$. $f(0) = 1$, and $f(1)$ is the determinant of F. If $f(1) < 0$, there exists a $\overline{\lambda}$ satisfying $0 < \overline{\lambda} < 1$, and $f(\overline{\lambda}) = 0$, a contradiction since $F(\overline{\lambda})$ is a PD matrix. Hence $f(1) \not< 0$. So the determinant of F is nonnegative.

□

Theorem 1.5 *If F is a* PSD *matrix, whether it is symmetric or not, all its principal subdeterminants are nonnegative.*

Proof. Follows from Result 1.4 and Theorem 1.4.

□

Theorem 1.6 *Let*

$$H = \begin{pmatrix} d_{11} & \dots & d_{1n} & d_{1,n+1} \\ \vdots & & \vdots & \vdots \\ d_{n1} & \dots & d_{nn} & d_{n,n+1} \\ d_{n+1,1} & \dots & d_{n+1,n} & d_{n+1,n+1} \end{pmatrix}, \quad D = \begin{pmatrix} d_{11} & \dots & d_{1n} \\ \vdots & & \vdots \\ d_{n1} & \dots & d_{nn} \end{pmatrix}$$

be symmetric matrices. H is of order $n+1$ and D is a principal submatrix of H. So $d_{ij} = d_{ji}$ for all i, $j = 1$ to $n+1$. Let $x \in \mathbf{R}^n$, $d = (d_{1,n+1}, \dots, d_{n,n+1})^T$, and $Q(x) = x^T D x + 2d^T x + d_{n+1,n+1}$. Suppose D is a PD *matrix. Let $x^* = -D^{-1}d$. Then x^* is the point which minimizes $Q(x)$ over $x \in \mathbf{R}^n$, and*

$$Q(x^*) = \text{ (determinant of } H) \text{ / (determinant of } D). \tag{1.12}$$

Also for any $x \in \mathbf{R}^n$

$$Q(x) = Q(x^*) + (x - x^*)^T D(x - x^*). \tag{1.13}$$

Proof. Since H is symmetric $\frac{\partial Q(x)}{\partial x} = 2(Dx + d)$. Hence x^* is the only point in $\mathbf{R}^n$ which satisfies $\frac{\partial Q(x)}{\partial x} = 0$. Also $Dx^* = -d$ implies

$$\begin{aligned} Q(x^*) &= {x^*}^T Dx^* + 2d^T x^* + d_{n+1,n+1} \\ &= d^T x^* + d_{n+1,n+1} \ . \end{aligned} \tag{1.14}$$

For $i = 1$ to $n+1$, if $g_{i,n+1} = d_{i,n+1} + \sum_{j=1}^{n} d_{ij} x_j^*$, and if $g = (g_{1,n+1}, \ldots, g_{n,n+1})^T$, then $g = d + Dx^* = 0$. Also $g_{n+1,n+1} = d_{n+1,n+1} + d^T x^* = Q(x^*)$ from (1.14). Now, from the properties of determinants, it is well known that the value of a determinant is unaltered if a constant multiple of one of its columns is added to another. For $j = 1$ to n, multiply the jth column of H by x_j^* and add the result to column $n+1$ of H. This leads to

$$\begin{aligned} \text{Determinant of } H \ &= \ \text{determinant of } \begin{pmatrix} d_{11} & \ldots & d_{1n} & g_{1,n+1} \\ \vdots & & \vdots & \vdots \\ d_{n1} & \ldots & d_{nn} & g_{n,n+1} \\ d_{n+1,1} & \ldots & d_{n+1,n} & g_{n+1,n+1} \end{pmatrix} \\ &= \ \text{determinant of } \begin{pmatrix} d_{11} & \ldots & d_{1n} & 0 \\ \vdots & & \vdots & \vdots \\ d_{n1} & \ldots & d_{nn} & 0 \\ d_{n+1,1} & \ldots & d_{n+1,n} & Q(x^*) \end{pmatrix} \\ &= \ (Q(x^*)) \ (\text{determinant of } D) \end{aligned}$$

which yields (1.12). (1.13) can be verified by straight forward expansion of its right hand side, or it also follows from Taylor expansion of $Q(x)$ around x^*, since $\frac{\partial^2 Q(x)}{\partial x^2} = 2D$ and x^* satisfies $\frac{\partial Q(x)}{\partial x} = 0$. Since D is a PD matrix, we have $(x - x^*)^T D(x - x^*) > 0$, for all $x \in \mathbf{R}^n$, $x \neq x^*$. This and (1.13) together imply that: $Q(x) > Q(x^*)$, for all $x \in \mathbf{R}^n$, $x \neq x^*$. Hence x^* is the point which minimizes $Q(x)$ over $x \in \mathbf{R}^n$.

□

Theorem 1.7 *Let H, D be square, symmetric matrices defined as in Theorem 1.6. H is* PD *iff D is* PD *and the determinant of H is strictly positive.*

Proof. Suppose H is PD. By Theorem 1.2 the determinant of H is strictly positive, and by Result 1.2 its principal submatrix D is also PD.

Suppose that D is PD and the determinant of H is strictly positive. Let $x = (x_1, \ldots, x_n)^T$ and $\xi = (x_1, \ldots, x_n, x_{n+1})^T$. Define d, $Q(x)$ as in Theorem 1.6. If $x_{n+1} = 0$, but $\xi \neq 0$ (i. e., $x \neq 0$), $\xi^T H \xi = x^T Dx > 0$, since D is PD. Now suppose $x_{n+1} \neq 0$. Let $\eta = (1/x_{n+1})x$. Then $\xi^T H \xi = x^T Dx + 2x_{n+1} d^T x + d_{n+1,n+1} x_{n+1}^{\mathbf{2}} = x_{n+1}^{\mathbf{2}} Q(\eta)$.

So, when $x_{n+1} \neq 0$, $\xi^T H \xi = x_{n+1}^2 Q(\eta) \geqq x_{n+1}^2$ (minimum value of $Q(\eta)$ over $\eta \in \mathbf{R}^n$) $= x_{n+1}^2$ ((determinant of H)/determinant of D)) > 0. So under our hypothesis that D is PD and the determinant of H is strictly positive, we have $\xi^T H \xi > 0$ for all $\xi \in \mathbf{R}^{n+1}$, $\xi \neq 0$, that is H is PD.

□

Theorem 1.8 *Let H be the square symmetric matrix defined in Theorem 1.6. H is* PD *iff the determinants of these $n+1$ principal submatrices of H,*

$$(d_{11}),\ \begin{pmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{pmatrix},\ \begin{pmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & d_{22} & d_{23} \\ d_{31} & d_{32} & d_{33} \end{pmatrix}, \ldots, D, H$$

are strictly positive.

Proof. Proof is by induction on the order of the matrix. Clearly, the statement of the theorem is true if H is of order 1. Now suppose the statement of the theorem is true for all square symmetric matrices of order n. By this and the hypothesis, we know that the matrix D is PD. So D is PD and the determinant of H is strictly positive by the hypothesis. By Theorem 1.7 these facts imply that H is PD too. Hence, by induction, the statement of the theorem is true in general.

□

Theorem 1.9 *A square symmetric matrix is* PD *iff all its principal subdeterminants are strictly positive.*

Proof. Let the matrix be H defined as in Theorem 1.6. If H is PD, all its principal subdeterminants are strictly positive by Theorem 1.3. On the other hand, if all the principal subdeterminants of H are strictly positive, the $n+1$ principal subdeterminants of H discussed in Theorem 1.8 are strictly positive, and by Theorem 1.8 this implies that H is PD.

□

Definition: P-matrix

A square matrix, whether symmetric or not, is said to be a P-matrix iff all its principal subdeterminants are strictly positive.

As examples, the matrices I, $\begin{pmatrix} 2 & 24 \\ 0 & 2 \end{pmatrix}$, $\begin{pmatrix} 2 & 1 \\ 2 & 2 \end{pmatrix}$ are P-matrices. The matrices $\begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}$, $\begin{pmatrix} -1 & 0 \\ 0 & 10 \end{pmatrix}$, $\begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}$ are not P-matrices.

Theorem 1.10 *A symmetric P-matrix is always* PD. *If a P-matrix is not symmetric, it may not be* PD.

Proof. By Theorem 1.9 B, a symmetric matrix is PD iff it is a P-matrix. Consider the matrix B,

$$B = \begin{pmatrix} 1 & 0 \\ 6 & 1 \end{pmatrix}, \quad B + B^T = \begin{pmatrix} 2 & 6 \\ 6 & 2 \end{pmatrix}.$$

Since all its principal subdeterminants are 1, B is a P-matrix. However, the determinant of $(B + B^T)$ is strictly negative, and hence it is not a PD matrix by Theorem 1.9, and by Result 1.9 this implies that B is not PD. Actually, it can be verified that, $(1, -1)B(1, -1)^T = -4 < 0$.

□

Note 1.1 The interesting thing to note is that if H is a symmetric matrix, and if the $n+1$ principal subdeterminants of H discussed in Theorem 1.8 are strictly positive, by Theorems 1.10 and 1.8 all principal subdeterminants of H are positive. This result may not be true if H is not symmetric.

Exercises

1.1 If H is a square symmetric PSD matrix, and its determinant is strictly positive, then prove that H is a PD matrix. Construct a numerical example to show that this result is not necessarily true if H is not symmetric.

1.2 Is the following statement true? "H is PSD iff its $(n+1)$ principal subdeterminants discussed in Theorem 1.8 are all nonnegative." Why? Illustrate with a numerical example.

By Theorem 1.9 the class of PD matrices is a subset of the class of P-matrices. By Theorem 1.10 when restricted to symmetric matrices, the property of being a PD matrix is the same as the property of being a P-matrix. An asymmetric P-matrix may not be PD, it may be a PSD matrix as the matrix $\widetilde{M}(n)$ below is, or it may not even be a PSD matrix. Let

$$\widetilde{M}(n) = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ 2 & 1 & 0 & \dots & 0 & 0 \\ 2 & 2 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 2 & 2 & 2 & \dots & 1 & 0 \\ 2 & 2 & 2 & \dots & 2 & 1 \end{pmatrix} . \tag{1.15}$$

$\widetilde{M}(n)$ is a lower triangular matrix in which all the diagonal entries are 1, and all entries below the diagonal are 2. All the principal subdeterminants of $\widetilde{M}(n)$ are clearly equal to 1, and hence $\widetilde{M}(n)$ is a P-matrix. However, $\widetilde{M}(n) + (\widetilde{M}(n))^T$ is the matrix in which all the entries are 2, and it can be verified that it is a PSD matrix and not a PD matrix.

Theorem 1.11 *Let F be a square* PSD *matrix of order n, whether it is symmetric or not. If $\overline{x} \in \mathbf{R}^n$ is such that $\overline{x}^T F \overline{x} = 0$, then $(F + F^T)\overline{x} = 0$.*

Proof. Let $D = F + F^T$. D is symmetric and by Result 1.9, D is PSD. For all $x \in \mathbf{R}^n$, $x^T Dx = 2x^T Fx$. So $\overline{x}^T D\overline{x} = 0$ too. We wish to prove that $D\overline{x} = 0$. Let $x \in \mathbf{R}^n$. For all real numbers λ, $(\overline{x} + \lambda x)^T D(\overline{x} + \lambda x) \geqq 0$, that is

$$\lambda^{\mathbf{2}} x^T Dx + 2\lambda \overline{x}^T Dx \geqq 0 \tag{1.16}$$

since $\overline{x}^T D\overline{x} = 0$. If $x^T Dx = 0$, by taking $\lambda = 1$ and then -1 in (1.16), we conclude that $\overline{x}^T Dx = 0$. If $x^T Dx \neq 0$, since D is PSD, $x^T Dx > 0$. In this case, from (1.16) we conclude that $2\overline{x}^T Dx \geqq -\lambda x^T Dx$ for $\lambda > 0$, and $2\overline{x}^T Dx \leqq -\lambda x^T Dx$ for $\lambda < 0$. Taking λ to be a real number of very small absolute value, from these we conclude that $\overline{x}^T Dx$ must be equal to zero in this case. Thus whether $x^T Dx = 0$, or $x^T Dx > 0$, we have $\overline{x}^T Dx = 0$. Since this holds for all $x \in \mathbf{R}^n$, we must have $\overline{x}^T D = 0$, that is $D\overline{x} = 0$.

□

Algorithm for Testing Positive Definiteness

Let $F = (f_{ij})$ be a given square matrix of order n. Find $D = F + F^T$. F is PD iff D is. To test whether F is PD, we can compute the n principal subdeterminants of D determined by the subsets $\{1\}, \{1,2\}, \ldots, \{1,2,\ldots,n\}$. F is PD iff each of these n determinants are positive, by Theorem 1.8. However, this is not an efficient method unless n is very small, since the computation of these separate determinants is time consuming.

We now describe a method for testing positive definiteness of F which requires at most n Gaussian pivot steps on D along its main diagonal; hence the computational effort required by this method is $O(n^{\mathbf{3}})$. This method is based on Result 1.8.

(i) If any of the principal diagonal elements in D are nonpositive, D is not PD. Terminate.

(ii) Subtract suitable multiples of row 1 from all the other rows, so that all the entries in column 1 and rows 2 to n of D are transformed into zero. That is, transform D into D_1 as in Result 1.8. If any diagonal element in the transformed matrix, D_1, is nonpositive, D is not PD. Terminate.

(iii) In general, after r steps we will have a matrix D_r of the form:

$$\begin{bmatrix} d_{11} & d_{12} & & & & \ldots & d_{1n} \\ 0 & \tilde{d}_{22} & & & & \ldots & \tilde{d}_{1n} \\ & 0 & \ddots & & & & \vdots \\ & & & \overline{d}_{rr} & & \ldots & \overline{d}_{rn} \\ & & & 0 & \hat{d}_{r+1,r+1} & \ldots & \hat{d}_{r+1,n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & & 0 & \hat{d}_{n,r+1} & \ldots & \hat{d}_{nn} \end{bmatrix}.$$

Subtract suitable multiples of row $r+1$ in D_r from rows i for $i > r+1$, so that all the entries in column $r+1$ and rows i for $i > r+1$ are transformed into 0.

This transforms D_r into D_{r+1}. If any element in the principle diagonal of D_{r+1} is nonpositive, D is not PD. Terminate. Otherwise continue the algorithm in the same manner for $n-1$ steps, until D_{n-1} is obtained, which is of the form

$$\begin{bmatrix} d_{11} & d_{12} & \dots & d_{1n} \\ 0 & \tilde{d}_{22} & \dots & \tilde{d}_{2n} \\ & 0 & & \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \overline{d}_{nn} \end{bmatrix}.$$

D_{n-1} is upper triangular. That's why this algorithm is called the **superdiagonalization algorithm**. If no termination has occured earlier and all the diagonal elements of D_{n-1} are positive, D, and hence, F is PD.

Example 1.3

Test whether

$$F = \begin{bmatrix} 3 & 1 & 2 & 2 \\ -1 & 2 & 0 & 2 \\ 0 & 4 & 4 & \frac{5}{3} \\ 0 & -2 & -\frac{13}{3} & 6 \end{bmatrix} \text{ is PD, } \quad D = F + F^T = \begin{bmatrix} 6 & 0 & 2 & 2 \\ 0 & 4 & 4 & 0 \\ 2 & 4 & 8 & -\frac{8}{3} \\ 2 & 0 & -\frac{8}{3} & 12 \end{bmatrix}.$$

All the entries in the principal diagonal of D (i. e., the entries d_{ii} for all i) are strictly positive. So apply the first step in superdiagonalization getting D_1. Since all elements in the principal diagonal of D_1 are strictly positive, continue. The matrices obtained in the order are:

$$D_1 = \begin{bmatrix} 6 & 0 & 2 & 2 \\ 0 & 4 & 4 & 0 \\ 0 & 4 & \frac{22}{3} & -\frac{10}{3} \\ 0 & 0 & -\frac{10}{3} & \frac{34}{3} \end{bmatrix}, \quad D_2 = \begin{bmatrix} 6 & 0 & 2 & 2 \\ 0 & 4 & 4 & 0 \\ 0 & 0 & \frac{10}{3} & -\frac{10}{3} \\ 0 & 0 & -\frac{10}{3} & \frac{34}{3} \end{bmatrix},$$

$$D_3 = \begin{bmatrix} 6 & 0 & 2 & 2 \\ 0 & 4 & 4 & 0 \\ 0 & 0 & \frac{10}{3} & -\frac{10}{3} \\ 0 & 0 & 0 & 8 \end{bmatrix}.$$

The algorithm terminates now. Since all diagonal entries in D_3 are strictly positive, conclude that D and, hence, F is PD.

Example 1.4

Test whether $D = \begin{pmatrix} 1 & 0 & 2 & 0 \\ 0 & 2 & 4 & 0 \\ 2 & 4 & 4 & 5 \\ 0 & 0 & 5 & 3 \end{pmatrix}$ is PD.

D is already symmetric, and all its diagonal elements are positive. The first step of the algorithm requires performing the operation: (row 3) – 2(row 1) on D. This leads to

$$D_1 = \begin{pmatrix} 1 & 0 & 2 & 0 \\ 0 & 2 & 4 & 0 \\ 0 & 4 & 0 & 5 \\ 0 & 0 & 5 & 3 \end{pmatrix} .$$

Since the third diagonal element in D_1 is not strictly positive, D is not PD.

Algorithm for Testing Positive Semidefiniteness

Let $F = (f_{ij})$ be the given square matrix. Obtain $D = F + F^T$. If any diagonal element of D is 0, all the entries in the row and column of the zero diagonal entry must be zero. Otherwise D (and hence F) is not PSD and we terminate. Also, if any diagonal entries in D are negative, D cannot be PSD and we terminate. If termination has not occurred, reduce the matrix D by striking off the rows and columns of zero diagonal entries.

Start off by performing the row operations as in (ii) above, that is, transform D into D_1. If any diagonal element in D_1 is negative, D is not PSD. Let E_1 be the submatrix of D_1 obtained by striking off the first row and column of D_1. Also, if a diagonal element in E_1 is zero, all entries in its row and column in E_1 must be zero. Otherwise D is not PSD. Terminate. Continue if termination does not occur.

In general, after r steps we will have a matrix D_r as in (iii) above. Let E_r be the square submatrix of D_r obtained by striking off the first r rows and columns of D_r. If any diagonal element in E_r is negative, D cannot be PSD. If any diagonal element of E_r is zero, all the entries in its row and column in E_r must be zero; otherwise D is not PSD. Terminate. If termination does not occur, continue.

Let d_{ss} be the first nonzero (and, hence, positive) diagonal element in E_r. Subtract suitable multiples of row s in D_r from rows i, $i > s$, so that all the entries in column s and rows i, $i > s$ in D_r, are transformed into 0. This transforms D_r into D_s and we repeat the same operations with D_s. If termination does not occur until D_{n-1} is obtained and, if the diagonal entries in D_{n-1} are nonnegative, D and hence F are PSD.

In the process of obtaining D_{n-1}, if all the diagonal elements in all the matrices obtained during the algorithm are strictly positive, D and hence F is not only PSD but actually PD.

Example 1.5

Is the matrix

$$F = \begin{bmatrix} 0 & -2 & -3 & -4 & 5 \\ 2 & 3 & 3 & 0 & 0 \\ 3 & 3 & 3 & 0 & 0 \\ 4 & 0 & 0 & 8 & 4 \\ -5 & 0 & 0 & 4 & 2 \end{bmatrix} \text{ PSD? } \quad D = F + F^T = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 6 & 6 & 0 & 0 \\ 0 & 6 & 6 & 0 & 0 \\ 0 & 0 & 0 & 16 & 8 \\ 0 & 0 & 0 & 8 & 4 \end{bmatrix} .$$

$D_{.1}$ and $D_{1.}$ are both zero vectors. So we eliminate them, but we will call the remaining matrix by the same name D. All the diagonal entries in D are nonnegative. Thus we apply the first step in superdiagonalization. This leads to

$$D_1 = \begin{pmatrix} 6 & 6 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 16 & 8 \\ 0 & 0 & 8 & 4 \end{pmatrix} \qquad E_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 16 & 8 \\ 0 & 8 & 4 \end{pmatrix} .$$

The first diagonal entry in E_1 is 0, but the first column and row of E_1 are both zero vectors. Also all the remaining diagonal entries in D_1 are strictly positive. So continue with superdiagonalization. Since the second diagonal element in D_1 is zero, move to the third diagonal element of D_1. This step leads to

$$D_3 = \begin{pmatrix} 6 & 6 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 16 & 8 \\ 0 & 0 & 0 & 0 \end{pmatrix} .$$

All the diagonal entries in D_3 are nonnegative. D and hence F is PSD but not PD.

Example 1.6

Is the matrix D in Example 1.4 PSD? Referring to Example 1.4 after the first step in superdiagonalization, we have

$$E_1 = \begin{pmatrix} 2 & 4 & 0 \\ 4 & 0 & 5 \\ 0 & 5 & 3 \end{pmatrix} .$$

The second diagonal entry in E_1 is 0, but the second row and column of E_1 are not zero vectors. So D is not PSD.

1.3.2 Relationship of Positive Semidefiniteness to the Convexity of Quadratic Functions

Let $\mathbf{\Gamma}$ be a convex subset of $\mathbf{R}^n$, and let $g(x)$ be a real valued function defined on $\mathbf{\Gamma}$. $g(x)$ is said to be a **convex function** on $\mathbf{\Gamma}$, if

$$g(\alpha x^1 + (1-\alpha)x^2) \leqq \alpha g(x^1) + (1-\alpha)g(x^2) \tag{1.17}$$

for every pair of points x^1, x^2 in $\mathbf{\Gamma}$, and for all $0 \leqq \alpha \leqq 1$. $g(x)$ is said to be a **strictly convex function** on $\mathbf{\Gamma}$ if (1.17) holds as a strict inequality for every pair of distinct points x^1, x^2 in $\mathbf{\Gamma}$ (i. e., $x^1 \neq x^2$) and for all $0 < \alpha < 1$. See Appendix 3.

Let F be a given square matrix of order n and c a row vector in $\mathbf{R}^n$. Let $f(x) = cx + x^T Fx$. Here we discuss conditions under which $f(x)$ is convex, or strictly convex. Let $D = (1/2)(F + F^T)$. If F is symmetric then $F = D$, otherwise D is the symmetrized form of F. Clearly $f(x) = cx + x^T Dx$. It can be verified that $\frac{\partial f(x)}{\partial x} = \left(\frac{\partial f(x)}{\partial x_1}, \ldots, \frac{\partial f(x)}{\partial x_n}\right)^T = c^T + (F + F^T)x = c^T + 2Dx$, and that $\frac{\partial^2 f(x)}{\partial x^2} =$ the Hessian of $f(x) = F + F^T = 2D$. Let x^1, x^2 be two arbitrary column vectors in $\mathbf{R}^n$ and let $\xi = x^1 - x^2$. Let α be a number between 0 and 1. By expanding both sides it can be verified that $\alpha f(x^1) + (1-\alpha) f(x^2) - f(\alpha x^1 + (1-\alpha)x^2) = \alpha(1-\alpha)\xi^T D\xi$ where $\xi = x^1 - x^2$. So $\alpha f(x^1) + (1-\alpha) f(x^2) - f(\alpha x^1 + (1-\alpha)x^2) \geqq 0$ for all $x^1, x^2 \in \mathbf{R}^n$ and $0 \leqq \alpha \leqq 1$, iff $\xi^T D\xi \geqq 0$ for all $\xi \in \mathbf{R}^n$, that is iff D (or equivalently F) is PSD. Hence $f(x)$ is convex on $\mathbf{R}^n$ iff F (or equivalently D) is PSD.

Also by the above argument we see that $\alpha f(x^1) + (1-\alpha) f(x^2) - f(\alpha x^1 + (1-\alpha)x^2) > 0$ for all $x^1 \neq x^2$ in $\mathbf{R}^n$ and $0 < \alpha < 1$, iff $\xi^T D\xi > 0$ for all $\xi \in \mathbf{R}^n$, $\xi \neq 0$. Hence $f(x)$ is strictly convex on $\mathbf{R}^n$ iff $\xi^T D\xi > 0$ for all $\xi \neq 0$, that is iff D (or equivalently F) is PD. These are the conditions for the convexity or strict convexity of the quadratic function $f(x)$ **over the whole space** $\mathbf{R}^n$. It is possible for $f(x)$ to be convex on a lower dimensional convex subset of $\mathbf{R}^n$ (for example, a subspace of $\mathbf{R}^n$) even though the matrix F is not PSD. For example, the quadratic form $f(x) = (x_1, x_2) \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} (x_1, x_2)^T$ is convex over the subspace $\{(x_1, x_2) : x_1 = 0\}$ but not over the whole of $\mathbf{R}^2$.

Exercise

1.3 Let $\mathbf{K} \subset \mathbf{R}^n$ be a convex set and $Q(x) = cx + \frac{1}{2}x^T Dx$. If $Q(x)$ is convex over $\mathbf{K}$ and $\mathbf{K}$ has a nonempty interior, prove that $Q(x)$ is convex over the whole space $\mathbf{R}^n$.

1.3.3 Necessary Optimality Conditions for Quadratic Programming

We will now resume our discussion of the quadratic program (1.11).

Theorem 1.12 *If $\overline{x}$ is an optimum solution of (1.11), $\overline{x}$ is also an optimum solution of the LP*

$$\begin{array}{lll} \textit{minimize} & (c + \overline{x}^T D)x & \\ \textit{subject to} & Ax \geqq b & \\ & x \geqq 0\,. & \end{array} \tag{1.18}$$

Proof. Notice that the vector of decision variables in (1.18) is x; $\overline{x}$ is a given point and the cost coefficients in the LP (1.18) depend on $\overline{x}$. The constraints in both (1.11) and (1.18) are the same. The set of feasible solutions is a convex polyhedron. Let $\hat{x}$ be any feasible solution. By convexity of the set of feasible solutions $x_\lambda = \lambda\hat{x} + (1-\lambda)\overline{x} = \overline{x} + \lambda(\hat{x} - \overline{x})$ is also a feasible solution for any $0 < \lambda < 1$. Since $\overline{x}$ is an optimum feasible solution of (1.11), $Q(x_\lambda) - Q(\overline{x}) \geqq 0$, that is $\lambda(c + \overline{x}^T D)(\hat{x} - \overline{x}) + (1/2)\lambda^2(\hat{x}-\overline{x})^T D(\hat{x}-\overline{x}) \geqq 0$ for all $0 < \lambda < 1$. Dividing both sides by λ leads to $(c+\overline{x}^T D)(\hat{x}-\overline{x}) \geqq (-\lambda/2)(\hat{x}-\overline{x})^T D(\hat{x}-\overline{x})$ for all $0 < \lambda < 1$. This obviously implies $(c+\overline{x}^T D)(\hat{x} - \overline{x}) \geqq 0$, that is, $(c + \overline{x}^T D)\hat{x} \geqq (c + \overline{x}^T D)\overline{x}$. Since this must hold for an arbitrary feasible solution $\hat{x}$, $\overline{x}$ must be an optimum feasible solution of (1.18).

□

Corollary 1.1 *If $\overline{x}$ is an optimum feasible solution of (1.11), there exist vectors $\overline{y} \in \mathbf{R}^m$ and slack vectors $\overline{u} \in \mathbf{R}^N$, $\overline{v} \in \mathbf{R}^m$ such that $\overline{x}$, $\overline{y}$, $\overline{u}$, $\overline{v}$ together satisfy*

$$\begin{pmatrix} \overline{u} \\ \overline{v} \end{pmatrix} - \begin{pmatrix} D & -A^T \\ A & 0 \end{pmatrix} \begin{pmatrix} \overline{x} \\ \overline{y} \end{pmatrix} = \begin{pmatrix} c^T \\ -b \end{pmatrix}$$
$$\begin{pmatrix} \overline{u} \\ \overline{v} \end{pmatrix} \geqq 0 \quad \begin{pmatrix} \overline{x} \\ \overline{y} \end{pmatrix} \geqq 0 \quad \text{and} \quad \begin{pmatrix} \overline{u} \\ \overline{v} \end{pmatrix}^T \begin{pmatrix} \overline{x} \\ \overline{y} \end{pmatrix} = 0 \,. \tag{1.19}$$

Proof. ¿From the above theorem $\overline{x}$ must be an optimum solution of the LP (1.18). The corollary follows by using the results of Section 1.2 on this fact.

□

Necessary and Sufficient Optimality Conditions for Convex Quadratic Programs

The quadratic minimization problem (1.11) is said to be a **convex quadratic program** if $Q(x)$ is convex, that is, if D is a PSD matrix (by the results in Section 1.3.2, or Theorem 17 of Appendix 3). If D is not PSD, (1.11) is said to be a **nonconvex quadratic program**. Associate a Lagrange multiplier y_i to the ith constraint "$A_i.x \geqq b_i$" $i = 1$ to m; and a Lagrange multiplier u_j to the sign restriction on x_j in (1.11), $j = 1$ to N. Let $y = (y_1, \ldots, y_m)^T$, $u = (u_1, \ldots, u_N)^T$. Then the **Lagrangian** corresponding to the quadratic program (1.11) is $L(x, y, u) = Q(x) - y^T(Ax - b) - u^T x$. The Karush-Kuhn-Tucker necessary optimality conditions for (1.11) are

$$\begin{aligned} \frac{\partial L}{\partial x}(x, y, u) = c^T + Dx - A^T y - u = 0 \\ y \geqq 0,\ u \geqq 0 \\ y^T(Ax - b) = 0,\ u^T x = 0 \\ Ax - b \geqq 0,\ x \geqq 0\,. \end{aligned} \tag{1.20}$$

Denoting the slack variables $Ax - b$ by v, the conditions (1.20) can be verified to be exactly those in (1.19), written out in the form of an LCP. A feasible solution x

for (1.11), is said to be a **Karush-Kuhn-Tucker point** (or abbreviated as a **KKT point**) if there exist Lagrange multiplier vectors y, u, such that x, y, u together satisfy (1.20) or the equivalent (1.19). So the LCP (1.19) is the problem of finding a KKT point for (1.11). We now have the following results.

Theorem 1.13 *If $\overline{x}$ is an optimum solution for (1.11), $\overline{x}$ must be a KKT point for it, whether $Q(x)$ is convex or not.*

Proof. Follows from Theorem 1.12 and Corollary 1.1.

□

Thus (1.20) or equivalently (1.19) provide the necessary optimality conditions for a feasible solution x of (1.11) to be optimal. Or, in other words, every optimum solution for (1.11) must be a KKT point for it. However, given a KKT point for (1.11) we cannot guarantee that it is optimal to (1.11) in general. In the special case when D is PSD, every KKT point for (1.11) is optimal to (1.11), this is proved in Theorem 1.14 below. Thus for convex quadratic programs, (1.20) or equivalently (1.19) provide necessary and sufficient optimality conditions.

Theorem 1.14 *If D is* PSD *and $\overline{x}$ is a KKT point of (1.11), $\overline{x}$ is an optimum feasible solution of (1.11).*

Proof. ¿From the definition of a KKT point and the results in Section 1.2, if $\overline{x}$ is a KKT point for (1.11), it must be an optimum feasible solution of the LP (1.18). Let x be any feasible solution of (1.11).

$$Q(x) - Q(\overline{x}) = (c + \overline{x}^T D)(x - \overline{x}) + \frac{1}{2}(x - \overline{x})^T D(x - \overline{x}) .$$

The first term on the right-hand side expression is nonnegative since $\overline{x}$ is an optimal feasible solution of (1.18). The second term in that expression is also nonnegative since D is PSD. Hence, $Q(x) - Q(\overline{x}) \geqq 0$ for all feasible solutions, x, of (1.11). This implies that $\overline{x}$ is an optimum feasible solution of (1.11).

□

Clearly (1.19) is an LCP. An optimum solution of (1.11) must be a KKT point for it. Solving (1.19) provides a KKT point for (1.11) and if D is PSD, this KKT point is an optimum solution of (1.11). [If D is not PSD and if a KKT point is obtained when (1.19) is solved, it may not be an optimum solution of (1.11).]

Example 1.7 Minimum Distance Problem.

Let **K** denote the shaded convex polyhedral region in Figure 1.7. Let P_0 be the point $(-2, -1)$. Find the point in **K** that is closest to P_0 (in terms of the usual Euclidean distance). Such problems appear very often in operations research applications.

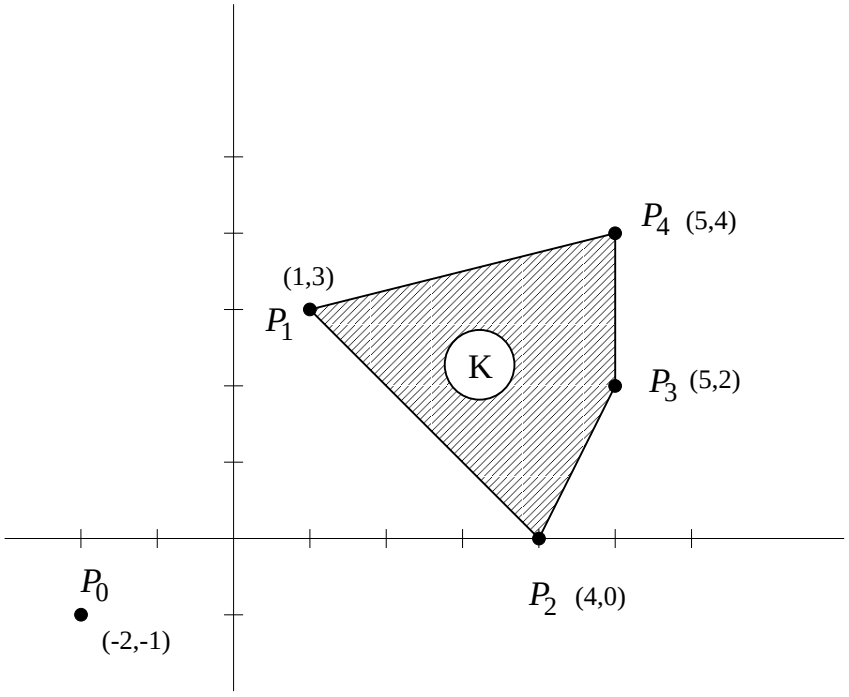


Figure 1.7

Every point in **K** can be expressed as a convex combination of its **extreme points** (or **corner points**) P_1, P_2, P_3, P_4. That is, the coordinates of a general point in **K** are: $(\lambda_1+4\lambda_2+5\lambda_3+5\lambda_4, 3\lambda_1+0\lambda_2+2\lambda_3+4\lambda_4)$ where the λ_i satisfy $\lambda_1+\lambda_2+\lambda_3+\lambda_4=1$ and $\lambda_i \geqq 0$ for all i. Hence, the problem of finding the point in **K** closest to P_0 is equivalent to solving:

$$\begin{array}{ll} \text{Minimize} & (\lambda_1+4\lambda_2+5\lambda_3+5\lambda_4-(-2))^{\mathbf{2}}+(3\lambda_1+2\lambda_3+4\lambda_4-(-1))^{\mathbf{2}} \\ \text{Subject to} & \lambda_1+\lambda_2+\lambda_3+\lambda_4=1 \\ & \lambda_i \geqq 0 \quad \text{for all } i\,. \end{array}$$

λ_4 can be eliminated from this problem by substituting the expression $\lambda_4 = 1-\lambda_1-\lambda_2-\lambda_3$ for it. Doing this and simplifying, leads to the quadratic program

$$\begin{array}{ll} \text{Minimize} & (-66,-54,-20)\lambda+\left(\frac{1}{2}\right)\lambda^T \begin{pmatrix} 34 & 16 & 4 \\ 16 & 34 & 16 \\ 4 & 16 & 8 \end{pmatrix} \lambda \\ \text{Subject to} & -\lambda_1-\lambda_2-\lambda_3 \geqq -1 \\ & \lambda \geqq 0 \end{array}$$

where $\lambda=(\lambda_1,\lambda_2,\lambda_3)^T$. Solving this quadratic program is equivalent to solving the LCP

$$\begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ v_1 \end{pmatrix} - \begin{pmatrix} 34 & 16 & 4 & 1 \\ 16 & 34 & 16 & 1 \\ 4 & 16 & 8 & 1 \\ -1 & -1 & -1 & 0 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ y_1 \end{pmatrix} = \begin{pmatrix} -66 \\ -54 \\ -20 \\ 1 \end{pmatrix} .$$

$$\text{All variables } u_1, u_2, u_3, v_1, \lambda_1, \lambda_2, \lambda_3, y_1 \geqq 0$$

$$\text{and } u_1\lambda_1 = u_2\lambda_2 = u_3\lambda_3 = v_1 y_1 = 0\,.$$

Let $(\tilde{u}_1, \tilde{u}_2, \tilde{u}_3, \tilde{v}_1, \tilde{\lambda}_1, \tilde{\lambda}_2, \tilde{\lambda}_3, \tilde{y}_1)$ be a solution to this LCP. Let $\tilde{\lambda}_4 = 1 - \tilde{\lambda}_1 - \tilde{\lambda}_2 - \tilde{\lambda}_3$. Then $\tilde{x} = (\tilde{\lambda}_1 + 4\tilde{\lambda}_2 + 5\tilde{\lambda}_3 + 5\tilde{\lambda}_4, 3\tilde{\lambda}_1 + 2\tilde{\lambda}_3 + 4\tilde{\lambda}_4)$ is the point in K that is closest to P_0.

1.3.4 Convex Quadratic Programs and LCPs Associated with PSD Matrices

Consider the LCP (q, M), which is (1.6) – (1.8), in which the matrix M is PSD. Consider also the quadratic program

$$\begin{array}{lll} \text{Minimize} & z^T(Mz+q) \\ \text{Subject to} & Mz+q \geqq 0 \\ & z \geqq 0\,. \end{array}$$

This is a convex quadratic programming problem since M is PSD. If the optimum objective value in this quadratic program is > 0, clearly the LCP (q, M) has no solution. If the optimum objective value in this quadratic program is zero, and $\overline{z}$ is any optimum solution for it, then $(\overline{w} = M\overline{z} + q, \overline{z})$ is a solution of the LCP. Conversely if $(\tilde{w}, \tilde{z})$ is any solution of the LCP (q, M), the optimum objective value in the above quadratic program must be zero, and $\tilde{z}$ is an optimum solution for it. Thus every LCP associated with a PSD matrix can be posed as a convex quadratic program.

Now, consider a convex quadratic program in which $Q(x) = cx + \frac{1}{2}x^T Dx$ (where D is a symmetric PSD matrix) has to be minimized subject to linear constraints. Replace each equality constraint by a pair of opposing inequality constraints (for example, $Ax = b$ is replaced by $Ax \leqq b$ and $Ax \geqq b$). Now the problem is one of minimizing $Q(x)$ subject to a system of linear inequality constraints. This can be transformed into an LCP as discussed in Section 1.3.3. The matrix M in the corresponding LCP will be PSD by Result 1.10, since D is PSD. Thus every convex quadratic programming problem can be posed as an LCP associated with a PSD matrix.

1.3.5 Applications of Quadratic Programming

The Portfolio Problem

A big investment firm has a total of \$ a to invest. It has a list of n stocks in which this money can be invested. The problem is to determine how much of the available money should be invested in each stock. The solution of this problem is called a **portfolio**. In this problem, it is well known that "one should never put all of their eggs in one basket". So after a thorough study, the manager of the company has determined a lower bound \$ l_j and an upper bound \$ k_j for the amount to be invested in stock j, $j = 1$ to n. The yield from each stock varies randomly from year to year. By the analysis of past data, μ_j, the expected (or average) yield per dollar invested in stock j per year has been estimated. The yields from various stocks are not mutually independent, and the analysis of past data has also provided an estimate of the variance-covariance matrix, D, for the annual yields from the various stocks per dollar invested. D is a symmetric positive definite matrix of order n. If \$ x_j is the amount invested in stock j, $j = 1$ to n, the portfolio is $x = (x_1, \ldots, x_n)^T$, the expected annual yield from it is $\sum_{j=1}^{n} \mu_j x_j$ and the variance of the yield is $x^T D x$. The variance is a measure of the random fluctuation in the annual yield and hence it should be minimized. The company would, of course, like to see its expected yield maximized. One way of achieving both of these objectives is to specify a target or lower bound, say μ, on the expected yield and to minimize the variance subject to this constraint. This leads to the problem:

$$\begin{array}{ll} \text{Minimize} & x^T D x \\ \text{Subject to} & \sum_{j=1}^{n} \mu_j x_j \geqq \mu \\ & \sum x_j \leqq a \\ & l_j \leqq x_j \leqq k_j \,,\ j = 1 \text{ to } n \end{array}$$

which is a quadratic programming problem.

Constrained Linear Regression

We will illustrate this application with an example of eggs and chickens due to C. Marmolinero [1.22]. The first step in chicken farming is hatching, carried out by specialized hatcheries. When hatched, a **day-old-chicken** is born. It needs no food for the first two days, at the end of which it is called a **growing pullet** and moved out of the hatchery. Pullets have to grow over a period of approximately 19 weeks before they start producing eggs, and this is done by specialized growing units under optimum conditions of diet, heating, lighting etc. After 19 weeks of age, pullets are moved into the laying flock and are then called **hens**. Consider a geographical region, say a State. Data on the number of chickens hatched by hatcheries in the state during each month is available from published state government statistics. But, day-old-chickens may be

bought from, or sold to firms outside the state, statistics on which are not available. Define

y_t = number (in millions) of growing pullets in the state, on the first day of month t.

d_t = number (in millions) of day-old-chickens hatched by hatcheries in the state in month t (from government statistics).

Here d_t are not variables, but are the given data. People in the business of producing chicken feed are very much interested in getting estimates of y_t from d_t. This provides useful information to them in their production planning, etc. Not all the day-old-chickens placed by hatcheries in a month may be alive in a future month. Also, after five months of age, they are recorded as hens and do not form part of the population of growing pullets. So the appropriate linear regression model for y_t as a function of the d_t's seems to be $y_t = \beta_0 + \sum_{i=1}^{5} \beta_i d_{t-i}$, where β_0 is the number of pullets in census, which are not registered as being hatched (pullets imported into the State, or chickens exported from the State), and β_i is a survival rate (the proportion of chickens placed in month $t-i$ that are alive in month t, $i = 1$ to 5). We, of course, expect the parameters β_i to satisfy the constraints

$$0 \leqq \beta_5 \leqq \beta_4 \leqq \beta_3 \leqq \beta_2 \leqq \beta_1 \leqq 1\,. \tag{1.21}$$

To get the best estimates for the parameters $\beta = (\beta_0, \beta_1, \beta_2, \beta_3, \beta_4, \beta_5)^T$ from past data, the least squares method could be used. Given data on y_t, d_t over a period of time (say for the last 10 years), define $L_2(\beta) = \sum_t (y_t - \beta_0 - \sum_{i=1}^{5} \beta_i d_{t-i})^{\mathbf{2}}$. Under the least squares method the best values for β are taken to be those that minimize $L_2(\beta)$ subject to the constraints (1.21). This is clearly a quadratic programming problem.

One may be tempted to simplify this problem by ignoring the constraints (1.21) on the parameters β. The unconstrained minimum of $L_2(\beta)$ can be found very easily by solving the system of equations $\frac{\partial L_2(\beta)}{\partial \beta} = 0$.

There are two main difficulties with this approach. The first is that the solution of this system of equations requires the handling of a square matrix (a_{ij}) with $a_{ij} = 1/(i+j-1)$, known as the **Hilbert matrix**, which is difficult to use in actual computation because of ill-conditioning. It magnifies the uncertainty in the data by very large factors. We will illustrate this using the Hilbert matrix of order 2. This matrix is

$$H_2 = \begin{pmatrix} 1 & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} \end{pmatrix}.$$

Consider the following system of linear equations with H_2 as the coefficient matrix.

x_1	x_2	
1	$\frac{1}{2}$	b_1
$\frac{1}{2}$	$\frac{1}{3}$	b_2

It can be verified that the solution of this system of linear equations is $\overline{x} = (4b_1 - 6b_2, -6b_1 + 12b_2)^T$. Suppose we have the exact value for b_1 but only an approximate value for b_2. In the solution $\overline{x}$, errors in b_2 are magnified by 12 times in $\overline{x}_2$, and 6 times in $\overline{x}_1$. This is only in a small system involving the Hilbert matrix of order 2. The error magnification grows very rapidly in systems of linear equations involving Hilbert matrices of higher orders. In real world applications, the coefficients in the system of linear equations (constants corresponding to b_1, b_2 in the above system) are constructed using observed data, which are always likely to have small errors. These errors are magnified in the solution obtained by solving the system of equations, making that solution very unreliable. See reference [1.36]. The second difficulty is that even if we are able to obtain a reasonable accurate solution $\hat{\beta}$ for the system of equations $\frac{\partial L_2(\beta)}{\partial \beta} = 0$, $\hat{\beta}$ may violate the constraints (1.21) that the parameter vector β is required to satisfy. For example, when this approach was applied on our problem with actual data over a 10-year horizon from a State, it led to the estimated parameter vector $\hat{\beta} = (4, .22, 1.24, .70 - .13, .80)^T$. We have $\hat{\beta}_4 < 0$ and $\hat{\beta}_2 > 1$, these values are not admissible for survival rates. So $\beta = \hat{\beta}$ does not make any sense in the problem. For the same problem, when $L_2(\beta)$ was minimized subject to the constraints (1.21), using a quadratic programming algorithm it gave an estimate for the parameter vector which was quite good.

Parameter estimation in linear regression using the least squares method is a very common problem in many statistical applications, and in almost all branches of scientific research. In a large proportion of these applications, the parameter values are known to satisfy one or more constraints (which are usually linear). The parameter estimation problem in constrained linear regression is a quadratic programming problem when the constraints on the parameters are linear.

1.3.6 Application of Quadratic Programming in Algorithms for NLP, Recursive Quadratic Programming Methods for NLP

Recently, algorithms for solving general nonlinear programs, through the solution of a series of quadratic subproblems have been developed [1.41 to 1.54]. These methods are called **recursive quadratic programming methods**, or **sequential quadratic programming methods**, or **successive quadratic programming methods** in the literature. Computational tests have shown that these methods are especially efficient in terms of the number of function and gradient evaluations required. Implementation of these methods requires efficient algorithms for quadratic programming. We provide here a brief description of this approach for nonlinear programming. Consider the nonlinear program:

$$\begin{array}{ll} \text{Minimize} & \theta(x) \\ \text{Subject to} & g_i(x) = 0,\ i = 1 \text{ to } k \\ & g_i(x) \geqq 0,\ i = k+1 \text{ to } m \end{array} \tag{1.22}$$

where $\theta(x)$ and $g_i(x)$ are real valued twice continuously differentiable functions defined over $\mathbf{R}^n$. Let $g(x) = (g_1(x), \ldots, g_m(x))^T$. Given the Lagrange multiplier vector $\pi = (\pi_1, \ldots, \pi_k, \pi_{k+1}, \ldots, \pi_m)$, the Lagrangian corresponding to (1.22) is $L(x, \pi) = \theta(x) - \pi g(x)$. The first order (or Karush-Kuhn-Tucker) necessary optimality conditions for this problem are

$$\begin{aligned} \nabla_x L(x, \pi) = \nabla\theta(x) - \pi\nabla g(x) &= 0 \\ \pi_i &\geqq 0 \qquad i = k+1 \text{ to } m \\ \pi_i g_i(x) &= 0 \qquad i = k+1 \text{ to } m \\ g_i(x) &= 0 \qquad i = 1 \text{ to } k \\ g_i(x) &\geqq 0 \qquad i = k+1 \text{ to } m\,. \end{aligned} \tag{1.23}$$

The methods described here for tackling (1.22) try to obtain a solution $\overline{x}$ and a Lagrange multiplier vector $\overline{\pi}$, which together satisfy (1.23), through an iterative process. In each iteration, a quadratic programming problem is solved, the solution of which provides revised estimates of the Lagrange multipliers and also determines a search direction for a **merit function**. The merit function is an absolute value penalty function (L_1-penalty function) that balances the two competing goals of decreasing $\theta(x)$ and reducing the amounts by which the constraints are violated. The merit function is then minimized in the descent direction by using a line minimization procedure. The solution of this line minimization problem produces a revised point x. With the revised x and π, the method goes to the next iteration. The first iteration begins with an initial point x and Lagrange multiplier vector π satisfying $\pi_i \geqq 0$, $i = k+1$ to m.

At the beginning of an iteration, let $\hat{x}$, $\hat{\pi}$ be the current vectors. Define

$$Q(d) = L(\hat{x}, \hat{\pi}) + \big(\nabla_x L(\hat{x}, \hat{\pi})\big)d + \frac{1}{2}d^T \frac{\partial^2 L(\hat{x}, \hat{\pi})}{\partial x^2} d \tag{1.24}$$

where $d = x - \hat{x}$. $Q(d)$ is the Taylor series approximation for $L(x, \hat{\pi})$ around the current point $\hat{x}$ up to the second order. Clearly $\frac{\partial^2 L(\hat{x}, \hat{\pi})}{\partial x^2}$ changes in each iteration. Since this is an $n \times n$ matrix, recomputing it in each iteration can be very expensive computationally. So in computer implementations of this method, $\frac{\partial^2 L(\hat{x}, \hat{\pi})}{\partial x^2}$ is approximated by a matrix B which is revised from iteration to iteration using the BFGS Quasi-Newton update formula that is widely used for unconstrained minimization. In the initial step, approximate $\frac{\partial^2 L}{\partial x^2}$ by $B_0 = I$, the unit matrix of order n. Let x^t, π^t, B_t, denote the initial point, the initial Lagrange multiplier vector, and the approximation for $\frac{\partial^2 L}{\partial x^2}$ in the t-th iteration. Let x^{t+1} be the point and π^{t+1} the Lagrange multiplier vector at the end of this iteration. Define

$$\begin{aligned} \xi^{t+1} &= x^{t+1} - x^t \\ q^{t+1} &= \big(\nabla_x L(x^{t+1}, \pi^{t+1}) - \nabla_x L(x^t, \pi^{t+1})\big)^T \\ p^{t+1} &= r_{t+1} q^{t+1} + (1 - r_{t+1}) B_t \xi^{t+1} \end{aligned}$$

where

$$r_{t+1} \begin{cases} = 1 & \text{if } (\xi^{t+1})^T q^{t+1} \geqq (0.2)(\xi^{t+1})^T B_t \xi^{t+1} \\ = \dfrac{(0.8)((\xi^{t+1})^T B_t \xi^{t+1})}{(\xi^{t+1})^T B_t \xi^{t+1} - (\xi^{t+1})^T q^{t+1}}, & \text{if } (\xi^{t+1})^T q^{t+1} < (0.2)(\xi^{t+1})^T B_t \xi^{t+1}. \end{cases}$$

Then update $\frac{\partial^2 L}{\partial x^2}$ by the formula

$$B_{t+1} = B_t + \frac{p^{t+1}(p^{t+1})^T}{(\xi^{t+1})^T p^{t+1}} - \frac{(B_t \xi^{t+1})(B_t \xi^{t+1})^T}{(\xi^{t+1})^T B_t \xi^{t+1}}. \tag{1.25}$$

This updating formula is a slight modification of the BFGS (Broyden-Fletcher-Goldfarb-Shanno) formula for updating the Hessian (the BFGS updating formula discussed in Section 10.8.6 is for updating the inverse of the Hessian, the one given here is for updating the actual Hessian itself).

If $r_{t+1} = 1$, then $p^{t+1} = q^{t+1}$ and the updating formula reduces to the standard BFGS formula for the approximation of the Hessian. The definition of p^{t+1} using r_{t+1} is introduced to assure that $(\xi^{t+1})^T p^{t+1} > 0$, which guarantees the hereditary positive definiteness of the updates B_t. The quantities 0.2, 0.8 are choosen from numerical experiments, they can be changed. This updating formula provides a symmetric positive definite approximation for $\frac{\partial^2 L}{\partial x^2}$. Also, in actual implementation, the second term in $Q(d)$ in (1.24) is replaced by $(\nabla\theta(\hat{x}))d$.
Therefore, the quadratic program solved in this iteration is: find d that

$$\begin{array}{ll} \text{minimizes} & (\nabla\theta(\hat{x}))d + (1/2)d^T \hat{B} d \\ \text{subject to} & g_i(\hat{x}) + (\nabla g_i(\hat{x}))d \begin{cases} = 0, & i = 1 \text{ to } k \\ \geqq 0, & i = k+1 \text{ to } m \end{cases} \end{array} \tag{1.26}$$

where $\hat{B}$ is the current approximation for $\frac{\partial^2 L}{\partial x^2}$.

Let $\tilde{d}$ denote the optimum solution of the quadratic program (1.26), and let $\tilde{\pi} = (\tilde{\pi}_1, \ldots, \tilde{\pi}_m)$ denote the associated Lagrange multiplier vector corresponding to the constraints in (1.26). If $\tilde{d} = 0$, from the optimality conditions for the quadratic program (1.26), it can be verified that $(\hat{x}, \tilde{\pi})$ together satisfy (1.23) and we terminate. If $\tilde{d} \neq 0$, it will be a descent direction for the merit function at $\hat{x}$. In the quadratic programm (1.26), to make sure that the Taylor series approximations remain reasonable, one could add additional bound conditions of the form $-\delta_j \leqq d_j \leqq \delta_j$, $j = 1$ to n, where δ_j are suitably choosen small positive numbers.

The form of the function that is minimized in the line search in this iteration is the merit function which is a L_1-penalty function

$$S(x) = \theta(x) + \sum_{i=1}^{k} \hat{\mu}_i |g_i(x)| + \sum_{i=k+1}^{m} \hat{\mu}_i |\text{minimum } \{0, g_i(x)\}| \tag{1.27}$$

where the last two terms are weighted sums of the absolute constraint violations. The weights $\hat{\mu}_i$ used in (1.27) satisfy $\mu_i > |\tilde{\pi}_i|$, they are usually obtained from

$$\hat{\mu}_i = \text{ maximum } \{|\tilde{\pi}_i|, (1/2)(\overline{\mu}_i + |\tilde{\pi}_i|)\}, \; i = 1 \text{ to } m,$$

where $\overline{\mu}_i$ are the weights used in the previous iteration. In Theorem 1.15 given below we prove that if $\tilde{d} \neq 0$, it is a descent direction at the current point $\hat{x}$, for the specially choosen merit functions $S(x)$ defined in (1.27) (this means that for $\alpha > 0$ and small $S(\hat{x}+\alpha\tilde{d}) < S(\hat{x})$, i. e., that $S(x)$ strictly decreases as we move from $\hat{x}$ in the direction $\tilde{d}$). The merit function $S(x)$ is minimized on the half-line $\{x : x = \hat{x} + \alpha\tilde{d},\ \alpha \geqq 0\}$. For this we define $f(\alpha) = S(\hat{x} + \alpha\tilde{d})$ and minimize $f(\alpha)$ over $\alpha \geqq 0$ by using some one dimensional line search algorithm (see Chapter 10). If $\tilde{\alpha}$ is the value of α that minimizes $f(\alpha)$ over $\alpha \geqq 0$, let $\tilde{x} = \hat{x}+\tilde{\alpha}\tilde{d}$. The point $\tilde{x}$ is the new point, it is obtained by moving a step length of $\tilde{\alpha}$ from $\hat{x}$ in the direction $\tilde{d}$.

If $\tilde{x}$, $\tilde{\pi}$ satisfy (1.23) to a reasonalbe degree of approximation, the method terminates, otherwise it moves to the next iteration with them.

The Descent Property

Theorem 1.15 *Suppose $\hat{B}$ is symmetric and* PD. *Let $\tilde{d}$, $\tilde{\pi}$ be the optimum solution and the associated Lagrange multiplier vector for the quadratic program (1.26). If $\tilde{d} \neq 0$, it is a descent direction for the merit function $S(x)$ at $\hat{x}$.*

Proof. By the first order necessary optimality conditions for the quadratic program (1.26) we have

$$\begin{aligned} \nabla\theta(\hat{x}) + (\hat{B}\tilde{d})^T - \tilde{\pi}\nabla g(\hat{x}) = 0 \\ \tilde{\pi}_i(g_i(\hat{x}) + (\nabla g_i(\hat{x}))\tilde{d}) = 0\,,\ i = 1 \text{ to } m\,. \end{aligned} \tag{1.28}$$

So, for α positive and sufficiently small, since all the functions are continuously differentiable, we have

$$\begin{aligned} f(\alpha) = \ & S(\hat{x} + \alpha\tilde{d}) = \theta(\hat{x}) + \alpha(\nabla\theta(\hat{x}))\tilde{d} + \\ & \sum_{i=1}^{k} \hat{\mu}_i |g_i(\hat{x}) + \alpha(\nabla g_i(\hat{x}))\tilde{d}| \\ & - \sum_{i=k+1}^{m} \hat{\mu}_i(\min\{0, g_i(\hat{x}) + \alpha(\nabla g_i(\hat{x}))\tilde{d}\}) + o(\alpha) \end{aligned} \tag{1.29}$$

where $o(\alpha)$ is a function of α satisfying the property that the limit $(o(\alpha)/\alpha)$ as $\alpha \to 0$ is 0 (the reason for the minus sign in the last line of (1.29) is the following. Since $\min\{0, g_i(x)\} \leqq 0$, $|\min\{0, g_i(x)\}| = -\min\{0, g_i(x)\}$).

Let $\mathbf{J} = \{i : k+1 \leqq i \leqq m, g_i(\hat{x}) < 0\}$, the index set of inequality constraints in the original problem (1.22) violated by the current point $\hat{x}$. For $k+1 \leqq i \leqq m$, $i \notin \mathbf{J}$, if $g_i(\hat{x}) = 0$, then $(\nabla g_i(\hat{x}))\tilde{d} \geqq 0$, from the constraints in (1.26). So, when α is positive but sufficiently small, for $k+1 \leqq i \leqq m$, $i \notin \mathbf{J}$, $\min\{0, g_i(\hat{x}) + \alpha(\nabla g_i(\hat{x}))\tilde{d}\} = 0$. Therefore,

$$\sum_{i=k+1}^{m} \hat{\mu}_i(\min\{0, g_i(\hat{x}) + \alpha(\nabla g_i(\hat{x}))\tilde{d}\}) = \sum_{i\in\mathbf{J}} \hat{\mu}_i(g_i(\hat{x}) + \alpha(\nabla g_i(\hat{x}))\tilde{d})\,. \tag{1.30}$$

Also, for $1 \leqq i \leqq k$, $(\nabla g_i(\hat{x}))\tilde{d} = -g_i(\hat{x})$ by the constraints in (1.26). Therefore

$$\sum_{i=1}^{k} \hat{\mu}_i |g_i(\hat{x}) + \alpha(\nabla g_i(\hat{x}))\tilde{d}| = (1-\alpha) \sum_{i=1}^{k} \hat{\mu}_i |g_i(\hat{x})| \,. \tag{1.31}$$

¿From (1.28) we have $(\nabla\theta(\hat{x}))\tilde{d} = -\tilde{d}^T\hat{B}\tilde{d} + (\tilde{\pi}\nabla g(\hat{x}))\tilde{d} = -\tilde{d}^T\hat{B}\tilde{d} + \sum_{i=1}^{m} \tilde{\pi}_i(\nabla g_i(\hat{x}))\tilde{d} = -\tilde{d}^T\hat{B}\tilde{d} - \sum_{i=1}^{m} \tilde{\pi}_i g_i(\hat{x})$. Using this and (1.30), (1.31) in (1.29), we get

$$\begin{aligned} f(\alpha) =& \theta(\hat{x}) + \sum_{i=1}^{k} \hat{\mu}_i |g_i(\hat{x})| - \sum_{i\in\mathbf{J}} \hat{\mu}_i g_i(\hat{x}) \\ &+ \alpha[-\tilde{d}^T\hat{B}\tilde{d} - \sum_{i=1}^{k} \hat{\mu}_i |g_i(\hat{x})| - \sum_{i=1}^{m} \tilde{\pi}_i g_i(\hat{x}) - \sum_{i\in\mathbf{J}} \hat{\mu}_i(\nabla g_i(\hat{x}))\tilde{d}] + o(\alpha) \\ =& f(0) + \alpha[-\tilde{d}^T\hat{B}\tilde{d} - \sum_{i=1}^{k} (\hat{\mu}_i |g_i(\hat{x})| + \hat{\pi}_i g_i(\hat{x})) \\ &- \sum_{i\in\overline{\mathbf{J}}} \tilde{\pi}_i g_i(\hat{x}) - \sum_{i\in\mathbf{J}} (\hat{\mu}_i(\nabla g_i(\hat{x}))\tilde{d} + \tilde{\pi}_i g_i(\hat{x}))] + o(\alpha) \,, \end{aligned} \tag{1.32}$$

where $\overline{\mathbf{J}} = \{k+1, \ldots, m\} \setminus \mathbf{J}$. Now $\tilde{d}^T\hat{B}\tilde{d} > 0$ since $\hat{B}$ is PD and $\tilde{d} \neq 0$. Also, $\sum_{i=1}^{k} (\hat{\mu}_i |g_i(\hat{x})| + \tilde{\pi}_i g_i(\hat{x})) \geqq 0$, since $\hat{\mu}_i \geqq |\tilde{\pi}_i|$ for all $i = 1$ to k. Again $\sum_{i\in\overline{\mathbf{J}}} \tilde{\pi}_i g_i(\hat{x}) \geqq 0$ since $\tilde{\pi}_i \geqq 0$ and $g_i(\hat{x}) \geqq 0$ for all $i \in \overline{\mathbf{J}} = \{k+1, \ldots, m\} \setminus \mathbf{J}$. Further, for $i \in \mathbf{J}$, $g_i(\hat{x}) < 0$, the constraints in the quadratic program imply $(\nabla g_i(\hat{x}))\tilde{d} \geqq -g_i(\hat{x}) > 0$; therefore, $\sum_{i\in\mathbf{J}} (\hat{\mu}_i(\nabla g_i(\hat{x}))\tilde{d} + \tilde{\pi}_i g_i(\hat{x})) \geqq \sum_{i\in\mathbf{J}} |g_i(\hat{x})|(\hat{\mu}_i - \tilde{\pi}_i) \geqq 0$. All this implies that the coefficient of α on the right hand side of (1.32) is strictly negative, that is, $f(\alpha) - f(0) < 0$ when α is sufficiently small and positive.

□

It is possible that even though the original problem is feasible and has a KKT point, the quadratic program (1.26) may be infeasible in some steps. See Example 1.8. In such steps, it is possible to define an alternate quadratic program of higher dimension which is always feasible, whose solution again provides a descent direction for the merit function $S(x)$. One such modification is given by the following quadratic programming problem

$$\begin{aligned} \text{minimize} \quad & (\nabla\theta(\hat{x}))d + (1/2)d^T\hat{B}d + \rho\Big(\sum_{i=1}^{m} u_i + \sum_{i=1}^{k} v_i\Big) \\ \text{subject to} \quad & g_i(\hat{x}) + (\nabla g_i(\hat{x}))d + u_i - v_i = 0 \,, \; i = 1 \text{ to } k \\ & g_i(\hat{x}) + (\nabla g_i(\hat{x}))d + u_i \geqq 0 \,, \; i = k+1 \text{ to } m \\ & u_i, v_i \geqq 0, \text{ for all } i \end{aligned} \tag{1.33}$$

where ρ is a positive penalty parameter.

The quadratic program (1.33) is always feasible, since, $d = 0$ leads to a feasible solution to it. Let $\tilde{d}$, $\tilde{\pi}$ be an optimum solution and the associated Lagrange multiplier vector for (1.33). If $\tilde{d} \neq 0$, it can be shown that it provides a descent direction for the merit function $S(x)$ at the current point $\hat{x}$ using arguments similar to those in the proof of Theorem 1.15, and the method proceeds as usual. If (1.26) is infeasible and $\tilde{d} = 0$ is an optimum solution of (1.33), we cannot conclude that $\hat{x}$ is a KKT point for the original problem (1.22), and the method breaks down; however, the possibility of this occurrence can be discounted in practice.

Example 1.8

Consider the following nonlinear program from the paper of K. Tone [1.53].

$$\begin{array}{lll} \text{Minimize} & \theta(x) = x_1^{\mathbf{3}} + x_2^{\mathbf{2}} & \\ \text{Subject to} & g_1(x) = x_1^{\mathbf{2}} + x_2^{\mathbf{2}} - 10 = 0 & \\ & g_2(x) = x_1 \quad\quad - \ 1 \geqq 0 & \\ & g_3(x) = \quad\quad x_2 - \ 1 \geqq 0\,. & \end{array} \tag{1.34}$$

The set of feasible solutions for this problem is the thick chord of the circle in $\mathbf{R}^2$ in Figure 1.8. It can be verified that $x = (1, 3)^T$ is an optimum solution of this problem.

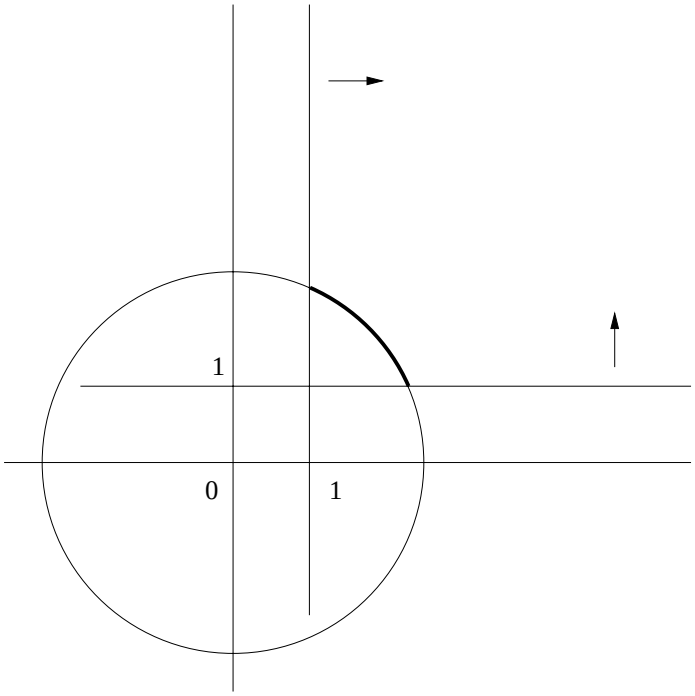


Figure 1.8

We have:

$$\nabla\theta(x) = (3x_1^2, 2x_2)$$
$$\nabla g_1(x) = (2x_1, 2x_2)$$
$$\nabla g_2(x) = (1, 0)$$
$$\nabla g_3(x) = (0, 1)\,.$$

We try to solve this problem by the recursive quadratic programming method using $x^\circ = (-10, -10)^T$ as the initial point. The constraints for the initial quadratic programming subproblem are

$$\begin{aligned} g_1(x^\circ) + (\nabla g_1(x^\circ))d = &\quad 190 - 20d_1 - 20d_2 = 0 \\ g_2(x^\circ) + (\nabla g_2(x^\circ))d = -\ &11 + \quad d_1 \qquad\quad \geqq 0 \\ g_3(x^\circ) + (\nabla g_3(x^\circ))d = -\ &11 \qquad\quad + \quad d_2 \geqq 0\,. \end{aligned}$$

Even though the original NLP is feasible and has an optimum solution, it can be verified that this quadratic subproblem is infeasible. So, we use the quadratic programming subproblem as in (1.33). Taking the initial approximation to the Hessian of the Lagrangian to be $B_\circ = I_2$, this leads to the following quadratic programming problem.

$$\begin{aligned} \text{minimize} \quad & 300d_1 - 20d_2 + (1/2)(d_1^2 + d_2^2) \\ & \qquad + \rho(u_1 + u_2 + u_3 + v_1) \\ \text{subject to} \quad & 20d_1 + 20d_2 \qquad\quad + u_1 - v_1 = 190 \\ & d_1 \qquad\qquad\quad + u_2 \qquad \geqq 11 \\ & \qquad d_2 \qquad\qquad\quad + u_3 \geqq 11 \\ & \qquad\qquad u_1, v_1, u_2, u_3 \qquad \geqq 0\,. \end{aligned} \tag{1.35}$$

Taking the penalty parameter $\rho = 1000$, this quadratic program has $\tilde{d} = (-1.5, 11)^T$ as the optimum solution with $\tilde{\pi} = (-35 \cdot 75, 1000, 692 \cdot 5)$ as the associated Lagrange multiplier vector corresponding to the constraints.

If we take penalty parameter vector $\mu = (1100, 1100, 1100)$ for constructing the merit function, we get the merit function

$$S(x) = x_1^3 + x_1^2 + 1100|x_1^2 + x_2^2 - 10| + 1100|\min\{0, x_1 - 1\}| + 1100|\min\{0, x_2 - 1\}|\,.$$

We minimize $S(x)$ on the half-line $\{x^\circ + \lambda\tilde{d} = (-10 - 1 \cdot 5\lambda, -10 + 11\lambda)^T, \lambda \geqq 0\}$. This problem can be solved using some of the line minimization algorithms discussed in Chapter 10. If the output of this problem is x^1, we update the Hessian approximation B, and with x^1, $\tilde{\pi}$ move over to the next quadratic programming subproblem and continue in the same way.

Under the assumptions:

(i) the quadratic program has an optimum solution in each step,

(ii) if $(\overline{x}, \overline{\pi})$ satisfies the KKT optimality conditions (1.23), then letting $\mathbf{J}(\overline{x}) = \{i : 1 \leqq i \leqq m, g_i(\overline{x}) = 0\}$, we have $\{\nabla g_i(\overline{x}) : i \in \mathbf{J}(\overline{x})\}$ is linearly independent; $\overline{\pi}_i > 0$

for all $i \in \mathbf{J}(\overline{x}) \cap \{k+1, \ldots, m\}$; and for any $y \neq 0$, $y \in \{y : (\nabla g_i(\overline{x}))\ y = 0,\ i \in \mathbf{J}(\overline{x})\}$, $y^T \left(\frac{\partial^2 L(\overline{x}, \overline{\pi})}{\partial x^2}\right) y > 0$,

(iii) the initial point x^0 is sufficiently close to a KKT point for (1.22);

it has been proved (see references [1.44, 1.45]) that the sequence (x^r, π^r) generated by the algorithm converges superlinearly to $(\overline{x}, \overline{\pi})$ which together satisfy (1.23).

These recursive quadratic programming methods have given outstanding numerical performance and thereby attracted a lot of attention. However, as pointed out above, one difficulty with this approach is that the quadratic programming problem (1.26) may be infeasible in some steps, even if the original nonlinear program has an optimum solution, in addition the modified quadratic program (1.33) may have the optimum solution $\tilde{d} = 0$, in which case the method breaks down. Another difficulty is that constraint gradients need to be computed for each constraint in each step, even for constraints which are inactive. Yet another difficulty is the function $f(\alpha)$ minimized in the line search routine in each step, which is a non-differentiable L_1-penalty function. To avoid these and other difficulties, the following modified sequential quadratic programming method has been proposed for solving (1.22) by K. Schittkowski [1.50, 1.51].

Choose the initial point x^0, multiplier vector π^0, $B_0 = I$ or some PD symmetric approximation for $\frac{\partial^2 L(x^0, \pi^0)}{\partial x^2}$, $\rho_0 \in R^1$, $\gamma^0 \in \mathbf{R}^m$ ($\rho_0 > 0$, $\gamma^0 > 0$) and constants $\varepsilon > 0$, $\overline{\rho} > 1$, $0 < \overline{\delta} < 1$. The choice of $\varepsilon = 10^{-7}$, $\overline{\delta} = 0.9$, $\overline{\rho} = 100$, and suitable positive values for ρ_0, γ^0 is reported to work well by K. Schittkowski [1.51]. Evaluate $\theta(x^0)$, $g_i(x^0)$, $\nabla g_i(x^0)$, $i = 1$ to m and go to stage 1.

General Stage $r+1$: Let x^r, π^r denote the current solution and Lagrange multiplier vector. Define

$$\mathbf{J}_1 = \{1, \ldots, k\} \cup \{i : k+1 \leqq i \leqq m, \text{ and either } g_i(x^r) \leqq \varepsilon \text{ or } \pi_i^r > 0\}$$
$$\mathbf{J}_2 = \{1, \ldots, m\} \setminus \mathbf{J}_1 .$$

The constraints in (1.22) corresponding to $i \in \mathbf{J}_1$ are treated as the active set of constraints at this stage, constraints in (1.22) corresponding to $i \in \mathbf{J}_2$ are the current inactive constraints.

Let B_r be the present matrix which is a PD symmetric approximation for $\frac{\partial^2 L(x^r, \pi^r)}{\partial x^2}$, this matrix is updated from step to step using the BFGS quasi-Newton update formula discussed earlier. The quadratic programming subproblem to be solved at this stage contains an additional variable, x_{n+1}, to make sure it is feasible. It is the following

$$\begin{aligned}
\text{minimize} \quad & P(d) = \frac{1}{2} d^T B_r d + (\nabla \theta(x^r)) d + \frac{1}{2}\left(\rho_r x_{n+1}^{\mathbf{2}}\right) \\
\text{subject to} \quad & (\nabla g_i(x^r)) d + (1 - x_{n+1}) g_i(x^r) \begin{cases} = 0, & i = 1 \text{ to } k \\ \geqq 0, & i \in \mathbf{J}_1 \cap \{k+1, \ldots, m\} \end{cases} \qquad (1.36) \\
& (\nabla g_i(x^{s_i})) d + g_i(x^r) \geqq 0,\ i \in \mathbf{J}_2 \\
& 0 \leqq x_{n+1} \leqq 1
\end{aligned}$$

where, for each $i \in \mathbf{J}_2$, x^{s_i} denotes the most recent point in the sequence of points obtained under the method, at which $\nabla g_i(x)$ was evaluated; and ρ_r is a positive penalty parameter which is updated in each step using the formula

$$\rho_r = \text{ maximum } \left\{ \rho_0 \, , \ \frac{\rho^* ((d^{r-1})^T A_{r-1} u^{r-1})^{\mathbf{2}}}{\left(1 - x_{n+1}^{r-1}\right)^{\mathbf{2}} (d^{r-1})^T B_{r-1} d^{r-1}} \right\} \tag{1.37}$$

where x_{n+1}^{r-1} u^{r-1}, d^{r-1} are the value of x_{n+1} in the optimum solution, the optimum Lagrange multiplier vector, and the optimum d-vector, associated with the quadratic programming problem in the previous stage; $\rho^* > 1$ is a constant; and A_{r-1} is the $n \times m$ matrix, whose jth column is the gradient vector of $g_i(x)$ computed at the most recent point, written as a column vector.

By definition of the set $\mathbf{J}_2$, the vector $(d = 0, x_{n+1} = 1)$ is feasible to this quadratic program, and hence, when B_r is PD, this quadratic program (1.34) has a finite unique optimum solution. One could also add additional bound constraints on the variables of the form $\delta_j \leqq d_j \leqq \delta_j$, $j = 1$ to n, where δ_j are suitable chosen positive numbers, to the quadratic programming subproblem (1.34), as discussed earlier.

Let (d^r, x_{n+1}^r), u^r, be the optimum solution and the optimum Lagrange multiplier vector, for the quadratic program (1.36). The solution of the quadratic programming subproblem (1.36) gives us the search direction d^r, for conducting a line search for a merit function or line search function corresponding to the original nonlinear program (1.22). If $x_{n+1}^r > \overline{\delta}$, change ρ_r into $\overline{\rho}\rho_r$ in (1.36) and solve (1.36) after this change. If this fails to lead to a solution with the value of x_{n+1} within the upper bound, define

$$\begin{aligned} d^r &= -B_r^{-1} \left(\nabla_x (\phi_{\gamma^r}(x^r, \pi^r)) \right)^T \\ u^r &= \pi^r - \nabla_\pi (\phi_{\gamma^r}(x^r, \pi^r)) \end{aligned} \tag{1.38}$$

where $\phi_{\gamma^r}(x^r, \pi^r)$ is the line search function or the merit function defined later on in (1.39).

The new point in this stage is of the form

$$\begin{aligned} x^{r+1} &= x^r + \alpha_r d^r \\ \pi^{r+1} &= \pi^r + \alpha_r (u^r - \pi^r) \end{aligned}$$

where α_r is a step length obtained by solving the line search problem

$$\text{minimize } \ h(\alpha) = \phi_{\gamma^{r+1}}(x^r + \alpha d^r, \pi^r + \alpha(u^r - \pi^r))$$

over $\alpha \in \mathbf{R}^1$, where

$$\phi_\gamma(x, \pi) = \theta(x) - \sum_{i \in \Gamma} (\pi_i g_i(x) - \frac{1}{2} \gamma_i (g_i(x))^{\mathbf{2}}) - \frac{1}{2} \sum_{i \in \Delta} \pi_i^{\mathbf{2}} / \gamma_i \tag{1.39}$$

where $\mathbf{\Gamma} = \{1, \ldots, k\} \cup \{i : k < i \leqq m, \ g_i(x) \leqq \pi_i/\gamma_i\}$, $\mathbf{\Delta} = \{1, \ldots, m\} \setminus \mathbf{\Gamma}$, and the penalty parameters γ_i are updated using the formula

$$\gamma_i^{r+1} = \text{ maximum } \left\{ \sigma_i^r \gamma_i^r \, , \ \frac{2m(u_i^r - \pi_i^r)^{\mathbf{2}}}{(1 - x_{n+1}^r)(d^r)^T B_r d^r} \right\} \, , \quad i = 1 \text{ to } m \, . \tag{1.40}$$

The sequence $\{\sigma_i^r : r = 0, 1, \ldots\}$ is a bounded sequence with $\sigma_i^r \leqq 1$ for all r, and it allows the possibility of decreasing the penalty parameters γ_i. A possible choice for updating these parameters σ^r from stage to stage is by the formula

$$\sigma_i^r = \text{ minimum } \left\{1, \frac{r}{\sqrt{\gamma_i^r}}\right\}, \; r = 1, 2, \ldots, \; i = 1 \text{ to } m \, .$$

The function $\phi_\gamma(x, \pi)$ is a differentiable augmented Lagrangian function. If (d^r, u^r) are obtained from the solution of the quadratic program (1.36), let γ^{r+1} be obtained using (1.40). On the other hand, if (d^r, u^r) are obtained from (1.38), let $\gamma^{r+1} = \gamma^r$.

If $\frac{d\,h(0)}{d\alpha} \geqq 0$, replace ρ_r by $\overline{\rho}\rho_r$, and go back to solving the modified quadratic subproblem (1.36). Otherwise, perform a line search to minimize $h(\alpha)$ with respect to α, over $\alpha \geqq 0$, and let α_r be the optimum value of α for this line minimization problem. Define

$$x^{r+1} = x^r + \alpha_r d^r$$
$$\pi^{r+1} = \pi^r + \alpha_r (u^r - \pi^r)$$

update the matrix B_r by the BFGS updating formula (1.25) and go to the next stage with these new quantities.

The algorithm can be terminated in the rth stage, if the following conditions are satisfied

$$(d^r)^T B_r d^r \leqq \varepsilon^{\mathbf{2}}$$
$$\sum_{i=1}^{m} |u_i^r g_i(x^r)| \leqq \varepsilon$$
$$\|\nabla_x L(x^r, u^r))\|^{\mathbf{2}} \leqq \varepsilon$$
$$\sum_{i=1}^{k} |g_i(x^r)| + \sum_{i=k+1}^{m} | \text{ minimum } (0, g_i(x^r))| \leqq \sqrt{\varepsilon} \, .$$

For a global convergence analysis of this algorithm under suitable constraint qualification assumptions, see [1.51].

Algorithms for Quadratic Programming Problems

In this book we will discuss algorithms for quadratic programming problems which are based on its transformation to an LCP as discussed above. Since the quadratic program is a special case of a nonlinear program, it can also be solved by the reduced gradient methods, linearly constrained nonlinear programming algorithms, and various other methods for solving nonlinear programs. For a survey of all these nonlinear programming algorithms, see Chapter 10.

1.4 TWO PERSON GAMES

Consider a game where in each play of the game, player I picks one out of a possible set of his m choices and independently player II picks one out of a possible set of his

N choices. In a play, if player I has picked his choice, i, and player II has picked his choice j, then player I loses an amount a'_{ij} dollars and player II loses an amount b'_{ij} dollars, where $A' = (a'_{ij})$ and $B' = (b'_{ij})$ are given **loss matrices**.

If $a'_{ij} + b'_{ij} = 0$ for all i and j, the game is known as a **zero sum game**; in this case it is possible to develop the concept of an **optimum strategy** for playing the game using Von Neumann's Minimax theorem. Games that are not zero sum games are called **nonzero sum games** or **bimatrix games**. In bimatrix games it is difficult to define an optimum strategy. However, in this case, an **equilibrium pair of strategies** can be defined (see next paragraph) and the problem of computing an equilibrium pair of strategies can be transformed into an LCP.

Suppose player I picks his choice i with a probability of x_i. The column vector $x = (x_i) \in \mathbf{R}^m$ completely defines player I's strategy. Similarly let the probability vector $y = (y_j) \in \mathbf{R}^N$ be player II's strategy. If player I adopts strategy x and player II adopts strategy y, the expected loss of player I is obviously $x^T A' y$ and that of player II is $x^T B' y$.

The strategy pair $(\overline{x}, \overline{y})$ is said to be an **equilibrium pair** if no player benefits by unilaterally changing his own strategy while the other player keeps his strategy in the pair $(\overline{x}, \overline{y})$ unchanged, that is, if

$$\overline{x}^T A' \overline{y} \leqq x^T A' \overline{y} \quad \text{for all probability vectors } x \in \mathbf{R}^m$$

and

$$\overline{x}^T B' \overline{y} \leqq \overline{x}^T B' y \quad \text{for all probability vectors } y \in \mathbf{R}^N .$$

Let α, β be arbitrary positive numbers such that $a_{ij} = a'_{ij} + \alpha > 0$ and $b_{ij} = b'_{ij} + \beta > 0$ for all i, j. Let $A = (a_{ij})$, $B = (b_{ij})$. Since $x^T A' y = x^T A y - \alpha$ and $x^T B' y = x^T B y - \beta$ for all probability vectors $x \in \mathbf{R}^m$ and $y \in \mathbf{R}^N$, if $(\overline{x}, \overline{y})$ is an equilibrium pair of strategies for the game with loss matrices A', B', then $(\overline{x}, \overline{y})$ is an equilibrium pair of strategies for the game with loss matrices A, B, and vice versa. So without any loss of generality, consider the game in which the loss matrices are A, B.

Since x is a probability vector, the condition $\overline{x}^T A \overline{y} \leqq x^T A \overline{y}$ for all probability vectors $x \in \mathbf{R}^m$ is equivalent to the system of constraints

$$\overline{x}^T A \overline{y} \leqq A_{i.} \overline{y} \quad \text{for all } i = 1 \text{ to } m .$$

Let e_r denote the column vector in R^r in which all the elements are equal to 1. In matrix notation the above system of constraints can be written as $(\overline{x}^T A \overline{y}) e_m \leqq A \overline{y}$. In a similar way the condition $\overline{x}^T B \overline{y} \leqq \overline{x}^T B y$ for all probability vectors $y \in \mathbf{R}^N$ is equivalent to $(\overline{x}^T B \overline{y}) e_N \leqq B^T \overline{x}$. Hence the strategy pair $(\overline{x}, \overline{y})$ is an equilibrium pair of strategies for the game with loss matrices A, B iff

$$\begin{aligned} A \overline{y} &\geqq (\overline{x}^T A \overline{y}) e_m \\ B^T \overline{x} &\geqq (\overline{x}^T B \overline{y}) e_N . \end{aligned} \tag{1.41}$$

Since A, B are strictly positive matrices, $\overline{x}^T A\overline{y}$ and $\overline{x}^T B\overline{y}$ are strictly positive numbers. Let $\overline{\xi} = \overline{x}/(\overline{x}^T B\overline{y})$ and $\overline{\eta} = \overline{y}/(\overline{x}^T A\overline{y})$. Introducing slack variables corresponding to the inequality constraints, (1.41) is equivalent to

$$\begin{pmatrix} \overline{u} \\ \overline{v} \end{pmatrix} - \begin{pmatrix} 0 & A \\ B^T & 0 \end{pmatrix} \begin{pmatrix} \overline{\xi} \\ \overline{\eta} \end{pmatrix} = \begin{pmatrix} -e_m \\ -e_N \end{pmatrix}$$
$$\begin{pmatrix} \overline{u} \\ \overline{v} \end{pmatrix} \geqq 0, \quad \begin{pmatrix} \overline{\xi} \\ \overline{\eta} \end{pmatrix} \geqq 0, \quad \begin{pmatrix} \overline{u} \\ \overline{v} \end{pmatrix}^T \begin{pmatrix} \overline{\xi} \\ \overline{\eta} \end{pmatrix} = 0 \, . \tag{1.42}$$

Conversely, it can easily be shown that if $(\overline{u}, \overline{v}, \overline{\xi}, \overline{\eta})$ is a solution of the LCP (1.42) then an equilibrium pair of strategies for the original game is $(\overline{x}, \overline{y})$ where $\overline{x} = \overline{\xi}/(\sum \overline{\xi}_i)$ and $\overline{y} = \overline{\eta}/(\sum \overline{\eta}_j)$. Thus an equilibrium pair of strategies can be computed by solving the LCP (1.42).

Example 1.9

Consider the game in which the loss matrices are

$$A' = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \quad B' = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix} \, .$$

Player I's strategy is a probability vector $x = (x_1, x_2)^T$ and player II's strategy is a probability vector $y = (y_1, y_2, y_3)^T$. Add 1 to all the elements in A' and 2 to all the elements in B', to make all the elements in the loss matrices strictly positive. This leads to

$$A = \begin{pmatrix} 2 & 2 & 1 \\ 1 & 2 & 2 \end{pmatrix} \quad B = \begin{pmatrix} 1 & 3 & 2 \\ 2 & 1 & 3 \end{pmatrix} \, .$$

The LCP corresponding to this game problem is

$$\begin{bmatrix} u_1 \\ u_2 \\ v_1 \\ v_2 \\ v_3 \end{bmatrix} - \begin{bmatrix} 0 & 0 & 2 & 2 & 1 \\ 0 & 0 & 1 & 2 & 2 \\ 1 & 2 & 0 & 0 & 0 \\ 3 & 1 & 0 & 0 & 0 \\ 2 & 3 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \eta_1 \\ \eta_2 \\ \eta_3 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \\ -1 \\ -1 \\ -1 \end{bmatrix} \tag{1.43}$$

$$u, v, \xi, \eta \geqq 0 \text{ and } u_1\xi_1 = u_2\xi_2 = v_1\eta_1 = v_2\eta_2 = v_3\eta_3 = 0 \, .$$

Example 1.10

The Prisoner's Dilemma:
Here is an illustration of a bimatrix game problem from [1.31]. Two well known criminals were caught. During plea bargaining their Judge urged them both to confess and plead guilty. He explained that if one of them confesses and the other does not, the one who confesses will be acquitted and the other one given a sentence of 10 years

in prison. If both of them confess, each will get a 5 year prison sentence. Both of them know very well that the prosecution's case against them is not strong, and the established evidence against them rather weak. However, the Judge said that if both of them decide not to confess, he will book both of them on some traffic violations for a year's prison term each. For each prisoner, let 1 refer to his choice of confessing and 2 to the choice of pleading not guilty. Measuring the loss in years in prison, their loss matrices are:

	A		B	
Player II's Choice $\rightarrow$	1	2	1	2
1	5	0	5	10
Player I's Choice				
2	10	1	0	1

In this game it can be verified that the probability vectors ($\overline{x} = (1,0)^T$, $\overline{y} = (1,0)^T$) provide the unique equilibrium pair for this game, resulting in a loss of a five year prison term for each player. But if both player's collude and agree to use the probability vectors ($\hat{x} = (0,1)^T$, $\hat{y} = (0,1)^T$), the result, loss of a year's prison term for each player, is much better for both. The trouble with the strategy $(\hat{x}, \hat{y})$ is that each can gain by double-crossing the other.

Example 1.11

The Battle of the Sexes:

Here is another illustration of a bimatrix game from [1.31]. A newly married couple have to decide how they will spend Friday evening. The husband (player II) proposes to go to a boxing match and the wife (player I) proposes to go to a musical concert. The man rates the pleasure (or gain, or negative loss) he derives by going to the concert and the boxing match to be 1 and 4 units respectively on a scale from 0 to 5; and the corresponding figure for the woman are 4 and 1 units respectively. For each player let 1, 2 refer to the choices of insisting on going to concert, boxing match respectively. If their choices disagree, there is a fight, and neither gains any pleasure from going out that evening. Treating loss as negative pleasure, here are the loss matrices.

	A		B	
Player II's Choice $\rightarrow$	1	2	1	2
1	-4	0	-1	0
Player I's Choice				
2	0	-1	0	-4

For this game, it can be verified that the probability vectors $(\overline{x} = (1,0)^T,\ \overline{y} = (1,0)^T)$. $(\hat{x} = (0,1)^T,\ \hat{y} = (0,1)^T)$ are both equilibrium pairs. The losses from the two equilibrium pairs $(\overline{x}, \overline{y})$, $(\hat{x}, \hat{y})$ are distinct, $(\overline{x}, \overline{y})$ will be preferred by player I, whereas II will prefer $(\hat{x}, \hat{y})$. Because of this, these equilibrium pairs are unstable. Even if player I knows that II will use the strategy $\hat{y}$, she may insist on using strategy $\overline{x}$ rather than $\hat{x}$, hoping that this will induce II to switch to $\overline{y}$. So, in this game, it is difficult to foresee what will happen. The probability vectors $(\tilde{x} = (4/5, 1/5),\ \tilde{y} = (1/5, 4/5)^T)$ is another equilibrium pair. In this problem, knowledge of these equilibrium pairs seems to have contributed very little towards the development of any "optimum" strategy.

Even though the theory of equilibrium strategies is mathematically elegant, and algorithms for computing them (through the LCP formulation) are practically efficient, they have not found many real world applications because of the problems with them illustrated in the above examples.

1.5 OTHER APPLICATIONS

Besides these applications, LCP has important applications in the nonlinear analysis of certain elastic-plastic structures such as reinforced concrete beams, in the free boundary problems for journal bearings, in the study of finance models, and in several other areas. See references [1.1 to 1.5, 1.8, 1.12, 1.13, 1.19, 1.21, 1.29, 1.32, 1.35].

1.6 THE NONLINEAR COMPLEMENTARITY PROBLEM

For each $j = 1$ to n, let $f_j(z)$ be a real valued function defined on $\mathbf{R}^n$. Let $f(z) = (f_1(z), \ldots, f_n(z))^T$. The problem of finding $z \in \mathbf{R}^n$ satisfying

$$\begin{aligned} z \geqq 0, \quad & f(z) \geqq 0 \\ z_j f_j(z) = 0, \quad & \text{for each } j = 1 \text{ to } n \end{aligned} \tag{1.44}$$

is known as a nonlinear complementarity problem (abbreviated as NLCP). If we define $f_j(z) = M_j.z + q_j$ for $j = 1$ to n, it can be verified that (1.44) becomes the LCP (1.1). Thus the LCP is a special case of the NLCP. Often, it is possible to transform the necessary optimality conditions for a nonlinear program into that of an NLCP and thereby solve the nonlinear program using algorithms for NLCP. The NLCP can be transformed into a fixed point computing problem, as discussed in Section 2.7.7, and solved by the piecewise linear simplicial methods presented in Section 2.7. Other than this, we will not discuss any detailed results on NLCP, but the references [1.14 to 1.16, 1.24, 1.25, 1.39] can be consulted by the interested reader.

1.7 Exercises

1.4 Consider the two person game with loss matrices A, B. Suppose $A + B = 0$. Then the game is said to be a **zero sum game** (see references [1.28, 1.31]). In this case prove that every equilibrium pair of strategies for this game is an optimal pair of strategies in the minimax sense (that is, it minimizes the maximum loss that each player may incur. See references [1.28, 1.31]). Show that the same results continue to hold as long as $a_{ij} + b_{ij}$ is a constant independent of i and j.

1.5 Consider the bimatrix game problem with given loss matrices A, B. Let $x = (x_1, \ldots, x_m)^T$ and $y = (y_1, \ldots, y_n)^T$ be the probability vectors of the two players. Let $X = (x_1, \ldots, x_m, x_{m+1})^T$ and $Y = (y_1, \ldots, y_n, y_{n+1})^T$. Let e_r be the column vector in $\mathbf{R}^r$ all of whose entries are 1. Let $\mathbf{S} = \{X : B^T x - e_n^T x_{m+1} \geqq 0, e_m^T x = 1, x \geqq 0\}$ and $\mathbf{T} = \{Y : Ay - e_m^T y_{n+1} \geqq 0, e_n^T y = 1, y \geqq 0\}$. Let $Q(X, Y) = x^T(A+B)y - x_{m+1} - y_{n+1}$. If $(\overline{x}, \overline{y})$ is an equilibrium pair of strategies for the game and $\overline{x}_{m+1} = \overline{x}^T B \overline{y}$, $\overline{y}_{n+1} = \overline{x}^T A \overline{y}$, prove that $(\overline{X}, \overline{Y})$ minimizes $Q(X, Y)$ over $\mathbf{S} \times \mathbf{T} = \{(X, Y) : X \in \mathbf{S},\ Y \in \mathbf{T}\}$. (O. L. Mangasarian)

1.6 Consider the quadratic program:

$$\begin{array}{ll} \text{Minimize} & Q(x) = cx + \frac{1}{2}x^T D x \\ \text{Subject to} & Ax \geqq b \\ & x \geqq 0 \end{array}$$

where D is a symmetric matrix. $\mathbf{K}$ is the set of feasible solutions for this problem. $\overline{x}$ is an interior point of $\mathbf{K}$ (i. e., $A\overline{x} > b$ and $\overline{x} > 0$).

(a) What are the necessary conditions for $\overline{x}$ to be an optimum solution of the problem?

(b) Using the above conditions, prove that if D is not PSD, $\overline{x}$ could not be an optimum solution of the problem.

1.7 For the following quadratic program write down the corresponding LCP.

$$\begin{array}{lll} \text{Minimize} & -6x_1 - 4x_2 - 2x_3 + 3x_1^2 + 2x_2^2 + \frac{1}{3}x_3^2 & \\ \text{Subject to} & x_1 + 2x_2 + x_3 \leqq 4 & \\ & x_j \geqq 0 & \text{for all} j\,. \end{array}$$

If it is known that this LCP has a solution in which all the variables x_1, x_2, x_3 are positive, find it.

1.8 Write down the LCP corresponding to

$$\begin{array}{ll} \text{Minimize} & cx + \frac{1}{2}x^T Dx \\ \text{Subject to} & x \geqq 0\,. \end{array}$$

1.9 Let

$$M = \begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix}, \quad q = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Show that the LCP (q, M) has four distinct solutions. For $n = 3$, construct a square matrix M of order 3 and a $q \in \mathbf{R}^3$ such that (q, M) has eight distinct solutions.

Hint. Try $-M = \begin{pmatrix} 2 & -1 & -1 \\ -1 & 3 & -1 \\ -1 & -1 & 4 \end{pmatrix}$ $q = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$; or try $M = -I, q > 0\,.$

1.10 Let

$$M = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \quad q = \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix}.$$

Find out a solution of the LCP (q, M) by inspection. However, prove that there exists no complementary feasible basis for this problem.
(L. Watson)

1.11 Test whether the following matrices are PD, PSD, or not PSD by using the algorithms described in Section 1.3.1

$$\begin{pmatrix} 0 & 1 & -1 \\ 0 & 0 & -2 \\ 1 & 2 & 1 \end{pmatrix}, \begin{pmatrix} 4 & 3 & -7 \\ 0 & 0 & -2 \\ 0 & 0 & 6 \end{pmatrix}, \begin{pmatrix} 4 & 100 & 2 \\ 0 & 2 & 10 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 5 & -2 & -2 \\ 0 & 5 & -2 \\ 0 & 0 & 5 \end{pmatrix}.$$

1.12 Let $Q(x) = (1/2)x^T Dx - cx$. If D is PD, prove that $Q(x)$ is bounded below.

1.13 Let $\mathbf{K}$ be a nonempty closed convex polytope in $\mathbf{R}^n$. Let $f(x)$ be a real valued function defined on $\mathbf{R}^n$. If $f(x)$ is a concave function, prove that there exists an extreme point of $\mathbf{K}$ which minimizes $f(x)$ on $\mathbf{K}$.

1.14 Let D be an arbitrary square matrix of order n. Prove that, for every positive and sufficiently large λ, the function $Q_\lambda(x) = x^T(D - \lambda I)x + cx$ is a concave function on $\mathbf{R}^n$.

1.15 Consider the following quadratic assignment problem.

$$\begin{array}{ll} \text{minimize} & z(x) = \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{p=1}^{n}\sum_{q=1}^{n} c_{ijpq}x_{ij}x_{pq} \\ \text{subject to} & \sum_{j=1}^{n} x_{ij} = 1, \text{ for all } i = 1 \text{ to } n \\ & \sum_{i=1}^{n} x_{ij} = 1, \text{ for all } j = 1 \text{ to } n \\ & x_{ij} \geqq 0, \text{ for all } i, j = 1 \text{ to } n \end{array} \tag{1.45}$$

and

$$x_{ij} \text{ integral for } i, j = 1 \text{ to } n\,. \tag{1.46}$$

Show that this discrete problem (1.45), (1.46) can be posed as another problem of the same form as (1.45) without the integrality constraints (1.46).

1.16 Consider an optimization problem of the following form

$$\begin{array}{ll} \text{minimize} & \dfrac{(x^T Dx)^{\mathbf{1/2}}}{dx + \beta} \\ \text{subject to} & Ax \geqq b \\ & x \geqq 0 \end{array}$$

where D is a given PSD matrix and it is known that $dx + \beta > 0$ on the set of feasible solutions of this problem. Using the techniques of fractional programming (see Section 3.20 in [2.26]), show how this problem can be solved by solving a single convex quadratic programming problem. Using this, develop an approach for solving this problem efficiently by algorithms for solving LCPs
(J. S. Pang, [1.33]).

1.17 Let D be a given square matrix of order n. Develop an efficient algorithm which either confirms that D is PSD or produces a vector $y \in \mathbf{R}^n$ satisfying $y^T Dy < 0$.

1.18 Consider the following quadratic programming problem

$$\begin{array}{ll} \text{minimize} & Q(x) = cx + \dfrac{1}{2}x^T Dx \\ \text{subject to} & a \leqq Ax \leqq b \\ & l \leqq x \leqq u \end{array}$$

where A, D, c, a, b, l, u are given matrices of orders $m \times n$, $n \times n$, $1 \times n$, $m \times 1$, $m \times 1$, $n \times 1$, $n \times 1$ respectively, and D is symmetric. Express the necessary optimality conditions for this problem in the form of an LCP. (R. W. H. Sargent, [1.37])

1.19 Suppose D is a symmetric matrix of order n. Show that the KKT necessary optimality conditions for the quadratic program

$$\begin{array}{ll} \text{minimize} & cx + (1/2)x^T Dx \\ \text{subject to} & 0 \leqq x \leqq b \end{array}$$

where $b > 0$ is a given vector, are of the form: find, $x, y \geqq 0$ in $\mathbf{R}^n$ satisfying $c^T + Dx + y \geqq 0$, $b - x \geqq 0$, $x^T(c^T + Dx + y) = y^T(b - x) = 0$. Express these conditions in the form of an LCP. Also prove that this is equivalent to finding an $x \in \mathbf{R}^n$ satisfying $0 \leqq x \leqq b$ and $(u - x)^T(Dx + c^T) \geqq 0$ for all $0 \leqq u \leqq b$. Prove that this LCP always has a solution and that the solution is unique if D is a P-matrix.
(B. H. Ahn [9.4], S. Karamardian [1.15])

1.20 Weighted Min-Max Location Problem: Given m points $a^i = (a^i_1, \ldots, a^i_n)^T \in \mathbf{R}^n$, $i = 1$ to m, and positive weights δ_i, $i = 1$ to m associated with these points, define the function $\theta(x) =$ maximum $\{\delta_i \sqrt{(x - a^i)^T(x - a^i)} : i = 1 \text{ to } m\}$ over $x \in \mathbf{R}^n$. The weighted min-max location problem is to find an $x \in \mathbf{R}^n$ that minimizes $\theta(x)$. Show that this problem is equivalent to the problem

$$\begin{array}{ll} \text{minimize} & \lambda \\ \text{subject to} & \lambda - \delta_i^{\mathbf{2}}(\|a^i\|^{\mathbf{2}} + \sum_{j=1}^{n} x_j^{\mathbf{2}} - 2\sum_{j=1}^{n} a^i_j x_j) \geqq 0, \ i = 1 \text{ to } m \end{array} \tag{1.47}$$

where λ is treated as another variable in (1.47). Consider the following quadratic program

$$\begin{array}{ll} \text{minimize} & Q(X) = \sum_{j=1}^{n} x_j^{\mathbf{2}} - x_{n+1} \\ \text{subject to} & x_{n+1} - 2\sum_{j=1}^{n} a^i_j x_j \leqq \|a^i\|^{\mathbf{2}} + \dfrac{\lambda}{\delta_i^{\mathbf{2}}}, \quad i = 1 \text{ to } m \end{array} \tag{1.48}$$

where x_{n+1} is an additional variable in (1.48), $X = (x_1, \ldots, x_n, x_{n+1})$. Prove that if $(\overline{x}, \overline{\lambda})$ is feasible to (1.47), $(\overline{x}, \overline{\lambda}, \overline{x}_{n+1})$ where $\overline{x}_{n+1} = \sum_{j=1}^{n} \overline{x}_j^{\mathbf{2}}$, is feasible to (1.48) with $Q(\overline{X}) = 0$. Conversely if $(\hat{x}, \hat{\lambda})$ is feasible to (1.48) with $Q(\hat{X}) \leqq 0$, then show that $(\hat{x} = (\hat{x}_1, \ldots, \hat{x}_n), \hat{\lambda})$ is feasible to (1.47). Also, for each $\lambda > 0$, prove that the optimum solution of (1.48) is unique. Treating λ as a parameter, denote the optimum solution of (1.48) as a function of λ by $X(\lambda)$. Let $\tilde{\lambda}$ be the smallest value of λ for which $Q(X(\lambda)) \leqq 0$. Prove that $x(\tilde{\lambda})$ is the optimum solution of the min-max location problem. Use these results to develop an algorithm for the min-max location problem based on solving a parametric right hand side LCP.
(R. Chandrasekaran and M. J. A. P. Pacca, [1.2])

1.21 Let F be a square matrix of order n. In general there may be no relation between determinant $((F + F^T)/2)$ and determinant (F). Establish conditions under which determinant $((F + F^T)/2) \leqq$ determinant (F).

1.22 Let $\mathbf{K} \subset \mathbf{R}^n$ convex and $Q(x) = cx + \frac{1}{2}x^T Dx$. If $Q(x)$ is convex over $\mathbf{K}$ and $\mathbf{K}$ has nonempty interior, prove that $Q(x)$ is convex over the whole space $\mathbf{R}^n$.

1.23 Concave Regression Problem: Here, given a real valued function $\theta(t)$ defined on an interval, it is desired to find a convex (or concave, depending on the application) function that approximates it as closely as possible. Specifically, suppose we are given $\theta_i = \theta(\alpha_i)$, $i = 1$ to n, where $\alpha_1 < \alpha_2 < \ldots < \alpha_n$. So we are given the values of $\theta(t)$ at the points $t = \alpha_1, \ldots, \alpha_n$. It is required to find real values $f_1, \ldots, f_n$ so that $f_i = f(\alpha_i)$, $i = 1$ to n where f is a convex function defined on the real line, that minimizes the measure of deviation $\sum_{i=1}^{n} \gamma_i(\theta_i - f_i)^{\mathbf{2}}$ where γ_i, $i = 1$ to n are given positive weights. Formulate this problem as an LCP.

1.24 $\mathbf{K}_1$ and $\mathbf{K}_2$ are two convex polyhedra in $\mathbf{R}^n$, each of them provided as the set of feasible solutions of a given system of linear inequalities. Develop an algorithm for the problem

$$\begin{aligned} &\text{minimize } \|x - y\| \\ &x \in \mathbf{K}_1, y \in \mathbf{K}_2 \,. \end{aligned}$$

1.25 Sylvester's Problem: We are given a set of n points in $\mathbf{R}^m$, $\{A_{.1}, \ldots, A_{.n}\}$, where $A_{.j} = (a_{1j}, \ldots, a_{mj})^T$, $j = 1$ to n. It is required to find the smallest diameter sphere in $\mathbf{R}^m$ containing all the points in the set $\{A_{.1}, \ldots, A_{.n}\}$. Transform this into a quadratic program and discuss an algorithm for solving it. Apply your algorithm to find the smallest diameter circle containing all the points in $\{(1,1), (-3,2), (1,-5), (-2,4)\}$ in $\mathbf{R}^2$.
(References [1.5, 1.29])

1.26 Let $\mathbf{K}$ be any convex polyhedral subset of R^n (you can assume that $\mathbf{K}$ is the set of feasible solutions of $Ax \geqq b$ where A, b are given). Let x^0, x^1 be given points in $\mathbf{R}^n$. Let $\tilde{x}$, $\hat{x}$ be respectively the nearest points in $\mathbf{K}$ (in terms of the usual Euclidean distance) to x^0, x^1 respectively. Prove that $\|\tilde{x} - \hat{x}\| \leqq \|x^0 - x^1\|$.

1.27 Let $\alpha = (\alpha_1, \ldots, \alpha_n)$ be a given row vector of $\mathbf{R}^n$ and let $x^0 \in \mathbf{R}^n$ be another given column vector. It is required to find the nearest point in $\mathbf{K} = \{x : \alpha x \leqq 0, x \geqq 0\}$ to x^0, in terms of the usual Euclidean distance. For this, do the following. Let λ be a real valued parameter. Let λ_0 be the smallest nonnegative value of λ for which the piecewise linear, monotonically decreasing function $\alpha(x^0 - \lambda\alpha^T)^+$ assumes a nonpositive value. Let $\overline{x} = (x^0 - \lambda_0\alpha^T)^+$. (For any vector $y = (y_j) \in \mathbf{R}^n$, $y^+ = (y_j^+)$ where $y_j^+ =$ Maximum $\{0, y_j\}$ for each j.) Prove that $\overline{x}$ is the nearest point in $\mathbf{K}$ to x^0.

Extend this method into one for finding the nearest point in $\mathbf{\Gamma} = \{x : x \geqq 0, \alpha x \leqq \delta\}$ to x^0, where δ is a given number, assuming that $\mathbf{\Gamma} \neq \emptyset$.
(W. Oettli [1.30])

1.28 Let M be a square matrix of order n and $q \in \mathbf{R}^n$. Let $z \in \mathbf{R}^n$ be a vector of variables. Define: $f_i(z) =$ minimum $\{z_i, M_i.z + q_i\}$, that is

$$\begin{aligned} f_i(z) &= I_i.z \quad \text{if} \quad (M_i. - I_i.)z + q_i \geqq 0 \\ &= M_i.z + q_i \quad \text{if} \quad (M_i. - I_i.)z + q_i \leqq 0 \end{aligned}$$

for each $i = 1$ to n.

(a) Show that $f_i(z)$ is a piecewise linear concave function defined on $\mathbf{R}^n$
(b) Consider the system of equations

$$f_i(z) = 0 \quad i = 1 \text{ to } n\,.$$

Let $\overline{z}$ be a solution of this system. Let $\overline{w} = M\overline{z} + q$. Prove that $(\overline{w}, \overline{z})$ is a complementary feasible solution of the LCP (q, M).
(c) Using (b) show that every LCP is equivalent to solving a system of piecewise linear equations.
(R. Saigal)

1.29 For $j = 1$ to n define $x_j^+ =$ Maximum $\{0, x_j\}$, $x_j^- = -$ Minimum $\{0, x_j\}$. Let $x = (x_j) \in \mathbf{R}^n$, $x^+ = (x_j^+)$, $x^- = (x_j^-)$. Given the square matrix M of order n, define the piecewise linear function

$$T_M(x) = x^+ - Mx^-\,.$$

Show that $T_M(x)$ is linear in each orthant of $\mathbf{R}^n$. Prove that $(w = x^+, z = x^-)$ solves the LCP (q, M) iff $q = T_M(x)$.
(R. E. Stone [3.71])

1.30 Let D be a given square matrix of order n, and $f(x) = x^T Dx$. Prove that there exists a nonsingular linear transformation: $y = Ax$ (where A is a square nonsingular matrix of order n) such that

$$f(x) = y_1^2 + \ldots + y_p^2 - y_{p+1}^2 - \ldots - y_r^2$$

where $0 \leqq p \leqq r \leqq n$. Discuss an efficient method for finding such a matrix A, given D.
Find such a transformation for the quadratic form $f(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2 - 2x_1x_2 - 2x_1x_3 - 2x_2x_3$ (this dates back to Lagrange in 1759, see D. E. Knuth [10.20]).

1.31 Sylvester's Law of Inertia (dates from 1852): Let D be a given square matrix of order n, and $f(x) = x^T Dx$. If there exist nonsingular linear transformations: $y = Ax$, $z = Bx$ (A, B are both square nonsingular matrices of order n) such that

$$f(x) = y_1^2 + \ldots + y_p^2 - y_{p+1}^2 - \ldots - y_r^2 = z_1^2 + \ldots + z_q^2 - z_{q+1}^2 - \ldots - z_s^2$$

then prove that $p = q$ and $r = s$.

This shows that the numbers p and r associated with a quadratic form, defined in Exercise 1.30 are unique
(see D. E. Knuth [10.20]).

1.32 Using the notation of Exercise 1.30 prove that $r = n$ iff the matrix $(D + D^T)/2$ has no zero eigenvalues and that p is the number of positive eigenvalues of $(D + D^T)/2$.

Let D_0, D_1 be two given square matrices of order n, and let $D_\alpha = (1 - \alpha)D_0 + \alpha D_1$. Let $r(D_\alpha)$, $p(D_\alpha)$ be the numbers r, p, associated with the quadratic form $f_\alpha = x^T D_\alpha x$ as defined in Exercise 1.30. If $r(D_\alpha) = n$ for all $0 \leqq \alpha \leqq 1$, prove that $p(D_0) = p(D_1)$.
(See D. E. Knuth [10.20].)

1.33 To Determine Optimum Mix of Ingredients for Moulding Sand in a Foundry: In a heavy casting steel foundry, moulding sand is prepared by mixing sand, resin (Phenol formaledhyde) and catalyst (Para toluene sulfonic acid). In the mixture the resin undergoes a condensation polymerization reaction resulting in a phenol formaldehyde polymer that bonds and gives strength. The bench life of the mixed sand is defined to be the length of the time interval between mixing and the starting point of setting of the sand mix. In order to give the workers adequate time to use the sand and for proper mould strength, the bench life should be at least 10 minutes. Another important characteristic of the mixed sand is the dry compression strength which should be maximized. An important variable which influences these characteristics is the resin percentage in the mix, extensive studies have shown that the optimum level for this variable is 2 % of the weight of sand in the mix, so the company has fixed this variable at this optimal level. The other process variables which influence the output characteristics are:

$x_1 =$ temperature of sand at mixing time
$x_2 =$ % of catalyst, as a percent of resin added
$x_3 =$ dilution of catalyst added at mixing.

The variable x_3 can be varied by adding water to the catalyst before it is mixed. An experiment conducted yielded the following data.

Dry Compression Strength

	$x_3 = 0$				10			
x_1	$x_2 = 25$	30	35	40	25	30	35	40
20^c	31.4	32.4	33.7	37.3	32.7	33.7	36.3	34.0
30^c	33.4	34.1	34.9	32.6	30.1	31.1	35.0	35.2
40^c	33.8	31.4	38.0	32.4	31.6	32.3	34.7	34.8

Bench Life

	$x_3 = 0$				10			
x_1	$x_2 = 25$	30	35	40	25	30	35	40
20^c	13.3	11.5	10.8	10.3	15.8	14.0	12.8	11.8
30^c	10.3	9.0	8.0	6.8	12.3	11.0	10.3	9.3
40^c	7.0	6.3	5.0	4.3	11.8	10.5	7.3	5.8

Bench life can be approximated very closely by an affine function in the variables x_1, x_2, x_3; and dry compression strength can be approximated by a quadratic function in the same variables. Find the functional forms for these characteristics that provide the best approximation. Using them, formulate the problem of finding the optimal values of the variables in the region $0 \leqq x_3 \leqq 10$, $25 \leqq x_2 \leqq 40$, $20 \leqq x_1 \leqq 40$, so as to maximize the dry compression strength subject to the additional constraint that the bench life should be at least ten, as a mathematical programming problem. Find an optimum solution to this mathematical program. (Hint: For curve fitting use either the least squares method discussed in Section 1.3.5, or the minimum absolute deviation methods based on linear programming discussed in [2.26, Section 1.2.5].)
(Bharat Heavy Electricals Ltd., Hardwar, India).

1.34 Synchronous Motor Design Problem: There are 11 important design variables (these are variables like the gauge of the copper wring used, etc. etc.) denoted by x_1 to x_{11} and let $x = (x_1, \ldots, x_{11})^T$. These variables effect the raw material cost for this motor, denoted by $f_0(x)$; the efficiency of the motor (= (output energy)/(input energy) measured as a percentage) denoted by $f_1(x)$; and the power factor (this measures leakage, it is a loss measured as a percentage) denoted by $f_2(x)$. Subroutines are available for computing each of the functions $f_0(x)$, $f_1(x)$, $f_2(x)$ for given x. The problem is to find optimal values for the variables which minimizes $f_0(x)$ subject to $f_1(x) \geqq 86.8$ and $f_2(x) \leqq 90$ and $l \leqq x \leqq u$, where l, u are specified lower and upper bound vectors for the variables. Discuss a method for solving this problem.

1.35 Quadratic Programming Model to Determine State Taxes: It is required to determine optimum levels for various state government taxes that minimizes instability while meeting constraints on growth rates over time. Seven different taxes are considered, sales, motor fuel, alcoholic beverages, tobacco, motor vehicle, personal income, and corporate taxes. State government finance is based on the assumption of predictable and steady growth of each tax over time. Instability in tax revenue is measured by the degree to which the actual revenue differs from predicted revenue.

Using past data, a regression equation can be determined to measure the growth in tax revenue over time. Let s be the tax rate for a particular tax and S_t the expected tax revenue from this tax in year t. Then the regression equation used is

$$\log_e S_t = a + bt + cs$$

where a, b, c are parameters to be determined using past data to give the closest fit. Data for the past 10 years from a state is used for this parameter estimation. Clearly, the parameter c can only be estimated, if the tax rate s for that tax has changed during this period, this has happened only for the motor fuel and the tobacco taxes. The best fit parameter values for the various taxes are given below (for all but the motor fuel and tobacco taxes, the tax rate has remained the same over the 10 years period for which the tax data is available, and hence the parameter a given below for these taxes, is actually the value of $a + cs$, as it was not possible to estimate a and c individually from the data).

Table 1: Regression coefficient values

j	Tax j	a	b	c
1	Sales	12.61	.108	
2	Motor fuel	10.16	.020	.276
3	Alcoholic beverages	10.97	.044	
4	Tobacco	9.79	.027	.102
5	Motor vehicle	10.37	.036	
6	Personal income	11.89	.160	
7	Corporate	211.09	.112	

The annual growth rate is simply the regression coefficient b multiplied by 100 to convert it to percent.

For 1984, the tax revenue from each tax as a function of the tax rate can be determined by estimating the tax base. This data, available with the state, is given below.

j	Tax j	Tax base (millions of dollars)
1	Sales	34,329
2	Motor fuel	3,269
3	Alcoholic beverages	811
4	Tobacco	702
5	Motor vehicle	2,935
6	Personal income	30,809
7	Corporate	4,200

If s_j is the tax rate for tax j in 1984 as a fraction, x_j = tax revenue to be collected in 1984 in millions of dollars for the jth tax is expected to be: (tax base for tax j) s_j.

Choosing the decision variables to be x_j for $j = 1$ to 7, let $x = (x_1, \ldots, x_7)^T$. The total tax revenue is $\sum_{j=1}^{7} x_j$. Then the variability or instability in this revenue is measured by the quadratic function $Q(x) = x^T V x$ where V, the variance-covariance matrix estimated from past data is

$$\begin{pmatrix} .00070 & -.00007 & .00108 & -.00002 & .00050 & .00114 & .00105 \\ & .00115 & .00054 & -.00002 & .00058 & -.00055 & .00139 \\ & & .00279 & .00016 & .00142 & .00112 & .00183 \\ & & & .00010 & .00009 & -.00007 & -.00003 \\ & & & & .00156 & .00047 & .00177 \\ & & & & & .00274 & .00177 \\ & & & & & & .00652 \end{pmatrix}.$$

Since V is symmetric, only the upper half of V is recorded above.

The problem is to determine the vector x that minimizes $Q(x)$, subject to several constraints. One of the constraints is that the total expected tax revenue for 1984 should be $T = 3300$ in millions of dollars. The second constraint is that a specified growth rate of λ in the total tax revenue should be maintained. It can be assumed that this overall growth rate is the function $\sum_{i=1}^{7} \frac{x_j b_j}{T}$ which is a weighted average of the growth rates of the various taxes. We would like to solve the problem treating λ as a nonnegative parameter. Of particular interest are values $\lambda = 9$ % and 13 %.

The other constraints are lower and upper bounds on tax revenues x_j, these are of the form $0 \leqq x_j \leqq u_j$ for each j; where u_j is twice the 1983 revenue from tax j. The vector $u = (u_j)$ is (2216, 490, 195, 168, 95, 2074, 504) in millions of dollars.

Formulate this problem as an LCP and solve it using the complementary pivot algorithm discussed in Chapter 2. Using the tax base information given above, determine the optimal tax rates for 1984 for each tax.

(F. C. White [1.40], my thanks to H. Bunch for bringing this paper to my attention.)

1.36 Consider the equality constrained nonlinear program

$$\begin{array}{ll} \text{minimize} & \theta(x) \\ \text{subject to} & h_i(x) = 0, \quad i = 1 \text{ to } m\,. \end{array}$$

The quadratic merit function for this problem is $S(x) = \theta(x) + (\rho/2)\sum_{i=1}^{m}(h_i(x))^{\mathbf{2}}$ where ρ is a positive penalty parameter. Let $\overline{x} \in \mathbf{R}^n$ be an initial point and $\overline{\mu} = (\overline{\mu}_1, \ldots, \overline{\mu}_m) \in \mathbf{R}^m$ be a given Lagrange multiplier vector. Consider the equality constrained quadratic program in variables $d = (d_1, \ldots, d_n)^T$

$$\begin{array}{ll} \text{minimize} & \nabla\theta(\overline{x})d + \frac{1}{2}d^T B d \\ \text{subject to} & (h(\overline{x}))^T + (\nabla h(\overline{x}))d = \overline{\mu}/\rho \end{array}$$

where B is a symmetric PD matrix of order n. If $\overline{d} \neq 0$ is an optimum solution of this quadratic program, and $\overline{\pi} = (\overline{\pi}_1, \ldots, \overline{\pi}_m)$ the associated Lagrange multiplier vector, prove that $\overline{d}$ is a descent direction for $S(x)$ at $\overline{x}$.

1.37 Let $A = (a_{ij})$ be a given square matrix of order n. Consider the usual assignment problem

$$\begin{array}{lll} \text{minimize} & z(x) = \displaystyle\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_{ij} & \\ \text{subject to} & \displaystyle\sum_{i=1}^{n} x_{ij} = 1, & j = 1 \text{ to } n \\ \text{subject to} & \displaystyle\sum_{j=1}^{n} x_{ij} = 1, & i = 1 \text{ to } n \\ \text{subject to} & x_{ij} \geqq 0, & i, j = 1 \text{ to } n\,. \end{array}$$

i) Prove that if A is PD and symmetric, $\overline{x} = I_n$ = unit matrix of order n, is an optimum solution for this problem. Is the symmetry of A important for this result to be valid?

ii) Using the above, prove that if A is PD and symmetric, there exists a vector $u = (u_1, \ldots, u_n)$ satisfying

$$u_i - u_j \geqq a_{ij} - a_{jj}, \quad i, j = 1 \text{ to } n\,.$$

1.38 Consider the problem of an investor having one dollar to invest in assets $i = 1, \ldots, n$. If x_i is invested in asset i, then $\xi_i x_i$ is returned at the end of the investment period, where $(\xi_1, \ldots, \xi_n)$ are random variables independent of the choice of $x_i s$, with the row-vector of means $\mu = (\mu_1, \ldots, \mu_n)$ $(\mu > 0)$ and a positive definite symmetric variance-covariance matrix D. In portfolio theory, under certain assumptions, it is shown that optimal investment proportions, $x = (x_1, \ldots, x_n)^T$, may be obtained by maximizing the fractional objective function

$$g(x) = \frac{\mu x}{(x^T D x)^{\mathbf{1/2}}}\,.$$

i) A real valued function $f(x)$ defined on a convex set $\mathbf{K} \subset \mathbf{R}^n$ is said to be pseudo-concave on $\mathbf{K}$ if it is differentiable on $\mathbf{K}$ and for every $x^1, x^2 \in \mathbf{K}$, $\nabla f(x^2)(x^1 - x^2) \leqq 0$ implies $f(x^1) \leqq f(x^2)$.

Prove that $g(x)$ is pseudo-concave in $\{x : x > 0\}$, even though it is not in general concave on this set.

For the problem of maximizing a pseudo-concave function on a convex set, prove that every local maximum is a global maximum.
Consider the problem

$$\begin{array}{ll} \text{maximize} & g(x) \\ \text{subject to} & \sum_{j=1}^{n} x_j = 1 \\ & x_j \geqq 0, \text{ for all } j\,. \end{array}$$

Show that this problem has a unique optimum solution. Also, show that an optimum solution of this problem can be obtained from the solution of the LCP $(-\mu, D)$.
(W. T. Ziemba, C. Parkan and R. Brooks-Hill [3.80])

1.39 In Section 1.3.5, the computational problems associated with the Hilbert matrix were mentioned briefly. Consider the following linear program

$$\begin{array}{ll} \text{maximize} & cx \\ \text{subject to} & Ax \leqq b \end{array}$$

where

$$A = \begin{pmatrix} \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{n+1} \\ \frac{1}{3} & \frac{1}{4} & \cdots & \frac{1}{n+2} \\ \vdots & \vdots & & \vdots \\ \frac{1}{n+1} & \frac{1}{n+2} & \cdots & \frac{1}{2n} \end{pmatrix}$$

$$b = (b_i : i = 1 \text{ to } n)^T = \left(\sum_{j=1}^{n} \frac{1}{i+j}\right)$$

$$c = (c_j : j = 1 \text{ to } n) = \left(\frac{2}{j+1} + \sum_{i=2}^{n} \frac{1}{j+i}\right)$$

Clearly, this problem has the unique optimum solution $\overline{x} = (1, 1, \ldots, 1)^T$ and the dual problem has the unique optimum solution $\overline{\pi} = (2, 1, 1, \ldots, 1)$. The coefficient matrix A is related to the Hilbert matrix of order n. Verify that when this problem is solved by pivotal algorithms such as the simplex algorithm, or by the complementary pivot algorithm through an LCP formulation, using finite precision arithmetic, the results obtained are very bad, if n exceeds 10, say.
(E. Bernarczuk, "On the results of solving some linear programming problems using program packages of IBM and Robotron computers")

1.40 Consider the LCP (q, M). Define

$$f(z) = \sum_{i=1}^{n} [\text{minimum } \{0, M_i.z + q_i - z_i\} + z_i].$$

Show that the LCP (q, M) is equivalent to the following concave minimization problem

$$\begin{array}{ll} \text{minimize} & f(z) \\ \text{subject to} & Mz+q \geqq 0 \\ & z \quad \geqq 0. \end{array}$$

(O. L. Mangasarian [8.15])

1.41 Let n be a positive integer. Consider a square matrix $x = (x_{ij})$ of order n. Order the entries x_{ij} in the matrix in the form of a vector in $\mathbf{R}^{n^2}$, in some order. Let $\mathbf{K} \subset \mathbf{R}^{n^2}$ denote the set of all such vectors corresponding to PSD matrices x. Prove that $\mathbf{K}$ is a convex cone, but not polyhedral, and has a nonempty interior.

1.42 Consider the LCP (q, M) (1.6) to (1.8), of order n. Now consider the following mixed 0-1 integer programming problem (MIP)

$$\begin{array}{ll} \text{maximize} & y_{n+1} \\ \text{subject to} & 0 \leqq My + qy_{n+1} \leqq e - x \\ & 0 \leqq y \leqq x, 0 \leqq y_{n+1} \leqq 1 \\ & x_i = 0 \text{ or } 1 \text{ for all } i = 1 \text{ to } n \end{array} \tag{1.49}$$

where $y = (y_1, \ldots, y_n)^T$, $x = (x_1, \ldots, x_n)^T$ and e is the vector of all 1s in $\mathbf{R}^n$. Suppose the optimum objective value in the MIP (1.49) is y^*_{n+1}.

If $y^*_{n+1} = 0$, prove that the LCP (q, M) has no solution.
If $y^*_{n+1} > 0$ and (y^*, x^*, y^*_{n+1}) is any optimum solution of the MIP (1.49), prove that (w^*, z^*) is a solution of the LCP (q, M), where

$$\begin{aligned} z^* &= (1/y^*_{n+1})y^* \\ w^* &= Mz^* + q \end{aligned}$$

(J. Ben Rosen, "Solution of general LCP by 0-1 Mixed integer programming", Computer Science Tech. Report 86-23, University of Minnesota, Minneapolis, May, 1986).

1.8 References

1.1 R. Chandrasekaran, "The Weighted Euclidean 1-Center Problem", *O. R. Letters*, 1 (1982) 111–112.

1.2 R. Chandrasekaran and M. J. A. P. Pacca, "Weighted Min-Max, Max-Min Location Problems: Finite and Polynomially Bounded Algorithms", *Opsearch*, 17 (1980) 172–180.

1.3 R. W. Cottle and G. B. Dantzig, "Complementary Pivot Theory of Mathematical Programming", *Linear Algebra and its Applications*, 1 (1968) 103–125.

1.4 O. DeDonato and G. Maier, "Mathematical Programming Methods for the Inelastic Analysis of Reinforced Concrete Frames Allowing for Limited Rotation Capacity", *International Journal for Numerical Methods in Engineering*, 4 (1972) 307–329.

1.5 J. Elzinga and D. W. Hearn, "The Minimum Covering Sphere Problem", *Management Science*, 19 (1972) 96–104.

1.6 M. L. Fisher and F. J. Gould, "A Simplicial Algorithm for the Nonlinear Complementarity Problem", *Mathematical Programming*, 6 (1974) 281-300.

1.7 O. Fujiwara, S. P. Han and O. L. Mangasarian, "Local Duality of Nonlinear Programs", *SIAM Journal on Control and Optimization*, 22 (1984) 162—169.

1.8 R. F. Giese and P. C. Jones, "An Economic Model of Short-Rotation Forestry", *Mathematical Programming*, 28 (1984) 206–217.

1.9 S. P. Han and O. L. Mangasarian, "Characterization of Positive Definite and Semidefinite Matrices via Quadratic Programming Duality", *SIAM Journal of Algebraic and Discrete Methods*, 5 (1984) 26–32.

1.10 S. P. Han and O. L. Mangasarian, "Conjugate Cone Characterization of Positive and Semidefinite Matrices", *Linear Algebra and Its Applications*, (to appear).

1.11 M. R. Hestenes, *Optimization Theory, The Finite Dimensional Case*, Wiley-Interscience, 1975.

1.12 I. Kaneko, "A Mathematical Programming Method for the Inelastic Analysis of Reinforced Concrete Frames", *International Journal for Numerical Methods in Engineering*, 11 (1977) 1137–1154.

1.13 I. Kaneko, "Piecewise Linear Elastic-Plastic Analysis", *International Journal for Numerical Methods in Engineering*, 14 (1979) 757–767.

1.14 S. Karamardian, "The Nonlinear Complementarity Problem With Applications, Parts I and II", *Journal of Optimization Theory and Applications*, 4 (1969) 87–98, 167–181.

1.15 S. Karamardian, "Generalized Complementarity Problem", *Journal of Optimization Theory and Applications*, 8 (1971) 161–168.

1.16 S. Karamardian, "The Complementarity Problem", *Mathematical Programming*, 2 (1972) 107–129.

1.17 G. J. Kolhler, "A Complementarity Approach for Solving Leontief Substitution Systems and (Generalized) Markov Decision Processes", *RAIRO Rochereche Operationnelle/Operations Research*, 13 (1979) 75–80.

1.18 C. E. Lemke and J. T. Howson, "Equilibrium Points of Bimatrix Games", *SIAM Journal of Applied Mathematics*, 12 (1964) 413–423.

1.19 G. Maier, "A Quadratic Programming Approach for Certain Classes of Nonlinear Structural Problem", *Meccanica*, 3 (1968) 121–130.

1.20 A. Majthay, "Optimality Conditions for Quadratic Programming", *Mathematical Programming*, 1 (1971) 359–365.

1.21 H. M. Markowitz, "Portfolio Selection", *The Journal of Finance*, 12 (1952) 77–91.

1.22 C. MarMolinero, "Negative Chickens: An Exercise in Model Building", *Mathematical Spectrum*, 11 (1978-79) 13–18.

1.23 C. McLinden, "An Analogue of Moreau's Approximation Theorem with Application to the Nonlinear Complementarity Problem", *Pacific Journal of Mathematics*, 88 (1980) 101—161.

1.24 N. Megiddo and M. Kojima, "On One Existence and Uniqueness of Solutions in Nonlinear Complementarity Theory", *Mathematical Programming*, 12 (1977) 110–130.

1.25 J. J. Moré, "Classes of Functions and Feasibility Conditions in Nonlinear Complementarity Problems", *Mathematical Programming*, 6 (1974) 327–338.

1.26 K. G. Murty, "On the Number of Solutions of the Complementarity Problem and Spanning Properties of Complementary Cones", *Linear Algebra and Its Applications*, 5 (1972) 65–108.

1.27 K. G. Murty, Article on "Complementarity Problems", in: J. Belzer, A. G. Holzman and A. Kent (editors), *Encyclopedia of Computer Science and Technology*, Marcel Dekker, 1976.

1.28 K. G. Murty, *Linear and Combinatorial Programming,* Krieger Publishing Company, Malabar, Florida, USA, 1976.

1.29 K. P. K. Nair and R. Chandrasekaran, "Optimal Location of a Single Service Center of Certain Types", *Naval Research Logistics Quarterly*, 18 (1971) 503–510.

1.30 W. Oettli, "An Iterative Method Having Linear Rate of Convergence, for Solving a Pair of Dual Linear Programs", *Mathematical Programming*, 3 (1972) 302–311.

1.31 G. Owen, *Game Theory,* 2nd Edition, Academic Press, New York, 1982.

1.32 J. S. Pang, I. Kaneko and W. P. Hallman, "On the Solution of Some (Parametric) Linear Complementarity Problems With Applications to Portfolio Selection, Structural Engineering and Acturial Graduation", *Mathematical Programming*, 16 (1979) 325–347.

1.33 J. S. Pang, "A Parametric Linear Complementarity Technique for Optimal Portfolio Selection With a Risk-Free Asset", *Operations Research*, 28 (1980) 927–941.

1.34 J. S. Pang, "Linear Complementarity Problems Solvable by an Efficient Polynomially Bounded Pivoting Algorithm", Technical Report, MRC, University of Wisconsin, Madison, Wisconsin, 1983.

1.35 A. A. Raimondi and J. Boyd, "A Solution for the Finite Journal Bearing and Its Application to Analysis and Design", III, *Transactions of the American Society of Lubrication Engineers*, 1 (1958) 194–209.

1.36 J. R. Rice, *Matrix Computations and Mathematical Software,* McGraw-Hill, 1981.

1.37 R. W. H. Sargent, "An Efficient Implementation of the Lemke Algorithm and Its Extension to Deal with Upper and Lower Bounds", *Mathematical Programming Study*, 7 (1978) 36–54.

1.38 H. Tuy, "Minimization of a Concave Function Over an Unbounded Polyhedral Convex Set. Application to the LCP", Institute of Mathematics, Hanoi (1981).

1.39 J. W. Tolle and M. L. Fisher, "The Nonlinear Complementarity Problem: Existence and Determination of Solutions", *SIAM Journal on Control and Optimization*, 15 (1977) 612–624.

1.40 F. C. White, "Trade-Off in Growth and Stability in State Taxes", *National Tax Journal*, 36 (1983) 103–114.

References on Recursive (or Sequential) Quadratic Programming Methods for Nonlinear Programming

1.41 D. P. Bertsekas, *Constrained Optimization and Lagrange Multiplier Methods,* Academic Press, New York, 1982.

1.42 M. C. Biggs, "On the Convergence of Some Constrained Minimization Algorithms Based on Recursive Quadratic Programming", *Journal of the Institute of Mathematics and Its Applications*, 21 (1978) 67–82.

1.43 U. M. Garcia-Palomares and O. L. Mangasarian, "Superlinearly Convergent Quasi-Newton Algorithms for Nonlinearly Constrained Optimization Problems", Computer Science Technical Report No. 195, University of Wisconsin, Madison, Wisconsin, 1974.

1.44 S. P. Han, "Superlinearly Convergent Variable Metric Algorithms for General Nonlinear Programming Problems", *Mathematical Programming*, 11 (1976) 263–282.

1.45 S. P. Han, "A Globally Convergent Method for Nonlinear Programming", *Journal of Optimization Theory and Its Applications*, 22 (1977) 297–309.

1.46 M. J. D. Powell, "A Fast Algorithm for Nonlinearly Constrained Optimization Calculations", in: G. A. Watson (editor), *Numerical Analysis, Lecture Notes in Mathematics*, Vol. 630, Springer Verlag, New York, 1978.

1.47 M. J. D. Powell, "The Convergence of Variable Metric Methods for Nonlinearly Constrained Optimization Calculations", in: O. L. Mangasarian, R. R. Meyer and S. M. Robinson (editors), *Nonlinear Programming 3*, Academic Press, New York, 1978.

1.48 B. N. Pshenichny and Y. M. Danilin, *Numerical Methods in Extremal Problems,* (translated from Russian by V. Zhitomirsky), MIR Publishers, Moscow, 1978.

1.49 S. M. Robinson, "A Quadratically Convergent Algorithm for General Nonlinear Programming Problems", *Mathematical Programming*, 3 (1972) 145–156.

1.50 K. Schittkowski, "The Nonlinear Programming Method of Wilson, Han and Powell with an Augmented Lagrangian Type Line Search Function, Part I: Convergence Analysis and Part 2: An Efficient Implementation with Linear Least Squares Subproblems", *Numerische Mathematik*, 38 (1981) 83–114 and 115–127.

1.51 K. Schittkowski, "On the Convergence of a Sequential Quadratic Programming Method with an Augmented Lagrangian Line Search Function", Technical Report SOL 82-4, Department of Operations Research, Stanford University, Stanford, California, 1982.

1.52 R. A. Tapia, "Quasi-Newton Methods for Equality Constrained Optimization: Equivalents of Existing Methods and New Implementation", in: O. Mangasarian, R. Meyer and S. M. Robinson (editors), *Symposium on Nonlinear Programming III*, Academic Press, New York, 1978, pp. 125–164.

1.53 K. Tone, "Revisions of Constraint Approximations in the Successive QP method for Nonlinear Programming Problems", *Mathematical Programming*, 26 (1983) 144–152.

1.54 R. B. Wilson, "A Simplicial Algorithm for Concave Programming", Ph. D. Dissertation, Graduate School of Business Administration, Harvard University, Boston, Massachusetts, 1963.

Chapter 2

THE COMPLEMENTARY PIVOT ALGORITHM AND ITS EXTENSION TO FIXED POINT COMPUTING

LCPs of order 2 can be solved by drawing all the complementary cones in the q_1, q_2-plane as discussed in Chapter 1.

Example 2.1

Let $q = \begin{pmatrix} 4 \\ -1 \end{pmatrix}$, $M = \begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix}$ and consider the LCP (q, M). The class of complementary cones corresponding to this problem is shown in Figure 1.5.

w_1	w_2	z_1	z_2	q
1	0	2	-1	4
0	1	-1	2	-1

$w_1, w_2, z_1, z_2 \geqq 0, \quad w_1 z_1 = w_2 z_2 = 0$

q lies in two complementary cones Pos $(-M_{.1}, I_{.2})$ and Pos $(-M_{.1}, -M_{.2})$. This implies that the sets of usable variables (z_1, w_2) and (z_1, z_2) lead to solutions of the LCP.

Putting $w_1 = z_2 = 0$ and solving the remaining system for the values of the usable variables (z_1, w_2) lead to the solution $(z_1, w_2) = (2, 1)$. Here $(w_1, w_2, z_1, z_2) = (0, 1, 2, 0)$ is a solution of this LCP. Similarly putting $w_1 = w_2 = 0$ and solving it for the value of the usable variables (z_1, z_2) leads to the second solution $(w_1, w_2, z_1, z_2) = (0, 0, \frac{7}{3}, \frac{2}{3})$ of this LCP.

Example 2.2

Let $q = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$ and $M = \begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix}$ and consider the LCP (q, M). The class of complementary cones corresponding to this problem is in Figure 1.5. Verify that q is not contained in any complementary cone. Hence this LCP has no solution.

This graphic method can be conveniently used only for LCPs of order 2. In LCPs of higher order, in contrast to the graphic method where all the complementary cones were generated, we seek only one complementary cone in which q lies. In this chapter we discuss the **complementary pivot algorithm** (which is also called the **complementary pivot method**) for solving the LCP. In the LCP (1.1) if $q \geqq 0$, $(w, z) = (q, 0)$ is a solution and we are done. So we assume $q \ngeqq 0$. First we will briefly review some concepts from linear programming. See [2.26] for complete details.

2.1 BASES AND BASIC FEASIBLE SOLUTIONS

Consider the following system of linear equality constraints in nonnegative variables

$$\begin{aligned} Ax &= b \\ x &\geqq 0 \end{aligned} \tag{2.1}$$

where A is a given matrix of order $m \times n$. Without any loss of generality we assume that the rank of A is m (otherwise either (2.1) is inconsistent, or redundant equality constraints in (2.1) can be eliminated one by one until the remaining system satisfies this property. See [2.26]). In this system, the variable x_j is associated with the column $A_{.j}$, $j = 1$ to n. A **basis** B for (2.1) is a square matrix consisting of m columns of A which is nonsingular; and the column vector of variables x_B associated with the columns in B, arranged in the same order, is the **basic vector** corresponding to it. Let D be the matrix consisting of the $n - m$ columns of A not in B, and let x_D be the vector of variables associated with these columns. When considering the basis B for (2.1), columns in B, D, are called the **basic, nonbasic columns** respectively; and the variables in x_B, x_D are called the **basic, nonbasic variables** respectively. Rearranging the variables, (2.1) can be written in partitioned form as

$$\begin{aligned} Bx_B \; + \; Dx_D \; &= \; b \\ x_B \geqq 0, \qquad x_D &\geqq 0 \; . \end{aligned}$$

The **basic solution** of (2.1) corresponding to the basis B is obtained by setting $x_D = 0$ and then solving the remaining system for the values of the basic variables. Clearly it is $(x_B = B^{-1}b,\ x_D = 0)$. This solution is feasible to (2.1) iff $B^{-1}b \geqq 0$, and in this case B is said to be a **feasible basis** for (2.1) and the solution $(x_B = B^{-1}b,\ x_D = 0)$

is called the **basic feasible solution** (abbreviated as BFS) of (2.1) corresponding to it. A basis B which is not feasible (i. e., if at least one component of $B^{-1}b$ is strictly negative) is said to be an **infeasible basis** for (2.1). Thus each feasible basis B for (2.1) determines a unique BFS for it.

When referring to systems of type (2.1), the word **solution** refers to a vector x satisfying the equality constraints '$Ax = b$', that may or may not satisfy the nonnegativity restrictions '$x \geqq 0$'. A solution x of (2.1) is a **feasible solution** if it satisfies $x \geqq 0$.

Definition: Degeneracy, Nondegeneracy of Basic Solutions for (2.1); of (2.1) itself: and of the b-Vector in (2.1) The basic solution associated with a given basis B for (2.1), whether it is feasible or not, is said to be **degenerate** if at least one component in the vector $B^{-1}b$ is zero, **nondegenerate** otherwise.

A system of constraints of the form (2.1) is said to be nondegenerate if it has no degenerate basic solutions (i. e., iff in every solution of (2.1), at least m variables are nonzero, when the rank of A is m), degenerate otherwise. When A has full row rank, the system (2.1) is therefore degenerate iff the column vector b can be expressed as a linear combination of r columns of A, where $r < m$, nondegenerate otherwise. Thus whether the system of constraints (2.1) is degenerate or nondegenerate depends on the position of the right hand side constants vector b in $\mathbf{R}^m$ in relation to the columns of A; and if the system is degenerate, it can be made into a nondegenerate system by just perturbing the b-vector alone.

The right hand side constants vector b in the system of constraints (2.1) is said to be degenerate or nondegenerate in (2.1) depending on whether (2.1) is degenerate or nondegenerate. See Chapter 10 in [2.26].

The definitions given here are standard definitions of degeneracy, nondegeneracy that apply to either a system of constraints of the form (2.1) or the right hand constants vector b in such a system, or a particular basic solution of such a system. This should not be confused with the concepts of (principal) degeneracy or (principal) nondegeneracy of square matrices defined later on in Section 2.3, or the degeneracy of complementary cones defined in Chapter 1.

As an example, consider the system of constraints given in Example 2.4 in Section 2.2.2. The BFS of this system associated with the basic vector (x_1, x_2, x_3, x_4) is $\bar{x} = (3, 0, 6, 5, 0, 0, 0, 0)^T$ and it is degenerate since the basic variable x_2 is zero in this solution. The BFS of this system associated with the basic vector (x_8, x_2, x_3, x_4) can be verified to be $x = (0, 1, 2, 3, 0, 0, 0, 1)^T$ which is a nondegenerate BFS. Since the system has a degenerate basic solution, this system itself is degenerate, also the b-vector is degenerate in this system.

Definition: Lexico Positive A vector $a = (a_1, \ldots, a_r) \in \mathbf{R}^r$, is said to be **lexico positive**, denoted by $a \succ 0$, if $a \neq 0$ and the first nonzero component in a is strictly positive. A vector a is lexico negative, denoted by $a \prec 0$, if $-a \succ 0$. Given two vectors $x, y \in \mathbf{R}^r$, $x \succ y$ iff $x - y \succ 0$; $x \prec y$ iff $x - y \prec 0$. Given a set of vectors

$\{ a^1, \ldots, a^k \} \subset \mathbf{R}^r$, a lexico minimum in this set is a vector a^j satisfying the property that $a^i \succeq a^j$ for each $i = 1$ to k. To find the lexico minimum in a given set of vectors from $\mathbf{R}^r$, compare the first component in each vector and discard all vectors not corresponding to the minimum first component, from the set. Compare the second component in each remaining vector and again discard all vectors not corresponding to the minimum in this position. Repeat in the same manner with the third component, and so on. At any stage if there is a single vector left, it is the lexico minimum. This procedure terminates after at most r steps. At the end, if two or more vectors are left, they are all equal to each other, and each of them is a lexico minimum in the set.

Example 2.3

The vector $(0, 0, 0.001, -1000)$ is lexico positive. The vector $(0, -1, 20000, 5000)$ is lexico negative. In the set of vectors $\{ (-2, 0, -1, 0), (-2, 0, -1, 1), (-2, 1, -20, -30), (0, -10, -40, -50) \}$, the vector $(-2, 0, -1, 0)$ is the lexico minimum.

Perturbation of the Right Hand Side Constants Vector in (2.1) to make it Nondegenerate.

If (2.1) is degenerate, it is possible to perturb the right hand side constants vector b slightly, to make it nondegenerate. For example, let ε be a parameter, positive and sufficiently small. Let $b(\varepsilon) = b + (\varepsilon, \varepsilon^{\mathbf{2}}, \ldots, \varepsilon^{\mathbf{m}})^T$. It can be shown that if b in (2.1) is replaced by $b(\varepsilon)$, it becomes nondegenerate, for all ε positive and **sufficiently small** (this really means that there exists a positive number $\varepsilon_1 > 0$ such that whenever $0 < \varepsilon < \varepsilon_1$, the stated property holds). This leads to the perturbed problem

$$\begin{aligned} Ax &= b(\varepsilon) \\ x &\geqq 0 \end{aligned} \tag{2.2}$$

which is nondegenerate for all ε positive and sufficiently small. See Chapter 10 in [2.26] for a proof of this fact. A basis B and the associated basic vector x_B for (2.1) are said to be **lexico feasible** if they are feasible to (2.2) whenever ε is positive and sufficiently small, which can be verified to hold iff each row vector of the $m \times (m+1)$ matrix $(B^{-1}b \ \vdots \ B^{-1})$ is lexico positive. Thus lexico feasibility of a given basis for (2.1) can be determined by just checking the lexico positivity of each row of $(B^{-1}b \ \vdots \ B^{-1})$ without giving a specific value to ε. For example, if $b \geqq 0$, and A has the unit matrix of order m as a submatrix, that unit matrix forms a lexico feasible basis for (2.1).

Canonical Tableaus

Given a basis B, the canonical tableau of (2.1) with respect to it is obtained by multiplying the system of equality constraints in (2.1) on the left by B^{-1}. It is

Tableau 2.1 : Canonical Tableau of (2.1) with Respect to the Basis B

basic variables	x	
x_B	$B^{-1}A$	$B^{-1}b$

Let D be the matrix consisting of the $n-m$ columns of A not in B, and let x_D be the vector of variables associated with these columns. When the basic and nonbasic columns are rearranged in proper order, the canonical Tableau 2.1 becomes

Tableau 2.2

basic variables	x_B	x_D	
x_B	I	$B^{-1}D$	$B^{-1}b = \bar{b}$

$\bar{b}$ is known as the **updated right hand side constants vector** in the canonical tableau. The column of x_j in the canonical tableau, $B^{-1}A_{.j} = \bar{A}_{.j}$ is called the **update column** of x_j in the canonical tableau. The **inverse tableau** corresponding to the basis B is

Tableau 2.3 : Inverse Tableau

basic variables	Inverse	basic values
x_B	B^{-1}	$B^{-1}b$

It just provides the basis inverse and the updated right-hand-side constants column. From the information available in the inverse tableau, the update column corresponding to any nonbasic variable in the canonical tableau can be computed using the formulas given above.

2.2 THE COMPLEMENTARY PIVOT ALGORITHM

We will now discuss a pivotal algorithm for the LCP introduced by C. E. Lemke, known as the **Complementary Pivot Algorithm** (because it chooses the entering variable by a **complementary pivot rule**, the entering variable in a step is always the complement of the dropping variable in the previous step), and also referred to as **Lemke's Algorithm** in the literature.

2.2.1 The Original Tableau

An artificial variable z_0 associated with the column vector $-e_n$ (e_n is the column vector of all 1's in $\mathbf{R}^n$) is introduced into (1.6) to get a feasible basis for starting the algorithm. In detached coefficient tableau form, (1.6) then becomes

w	z	z_0	
I	$-M$	$-e_n$	q

$$w \geqq 0, \quad z \geqq 0, \quad z_0 \geqq 0 \tag{2.3}$$

2.2.2 Pivot Steps

The complementary pivot algorithm moves among feasible basic vectors for (2.3). The primary computational step used in this algorithm is the pivot step (or the **Gauss-Jordan pivot step**, or the Gauss-Jordan elimination pivot step), which is also the main step in the simplex algorithm for linear programs. In each stage of the algorithm, the basis is changed by bringing into the basic vector exactly one nonbasic variable known as the **entering variable**. Its updated column vector is the **pivot column** for this basis change. The **dropping variable** has to be determined according to the **minimum ratio test** to guarantee that the new basis obtained after the pivot step will also be a feasible basis.

For example, assume that the present feasible basic vector is $(y_1, \ldots, y_n)$ with y_r as the r^{th} basic variable, and let the entering variable be x_s. (The variables in (2.3) are $w_1, \ldots, w_n$; $z_1, \ldots, z_n$, z_0. Exactly n of these variables are present basic variables. For convenience in reference, we assume that these basic variables are called $y_1, \ldots, y_n$). After we rearrange the variables in (2.3), if necessary, the canonical form of (2.3), with respect to the present basis is of the form :

Basic variable	$y_1, \ldots, y_n$	x_s	Other variables	Right-hand constant vector
y_1	$1 \ \ldots \ 0$	$\bar{a}_{1s}$	$\cdots$	$\bar{q}_1$
$\vdots$	$\vdots \quad \vdots$	$\vdots$	$\vdots$	$\vdots$
y_n	$0 \ \ldots \ 1$	$\bar{a}_{ns}$	$\cdots$	$\bar{q}_n$

Keeping all the nonbasic variables other than x_s, equal to zero, and giving the value λ to the entering variable, x_s, leads to the new solutions :

$$\begin{aligned} &x_s = \lambda \\ &y_i = \bar{q}_i - \lambda \bar{a}_{is}, \quad i = 1, \ldots, n \\ &\text{All other variables} = 0 \end{aligned} \tag{2.4}$$

There are two possibilities here.

1. The pivot column may be nonpositive, that is, $\bar{a}_{is} \leqq 0$ for all $1 \leqq i \leqq n$. In this case, the solution in (2.4) remains nonnegative for all $\lambda \geqq 0$. As λ varies from 0 to ∞, this solution traces an **extreme half-line** (or an **unbounded edge**) of the set of feasible solutions of (2.3). In this case the minimum ratio, θ, in this pivot step is $+\infty$. See Example 2.4.
2. There is at least one positive entry in the pivot column. In this case, if the solution in (2.4) should remain nonnegative, the maximum value that λ can take is $\theta = \frac{\bar{q}_r}{\bar{a}_{rs}} = \text{minimum}\left\{ \frac{\bar{q}_i}{\bar{a}_{is}} : i \text{ such that } \bar{a}_{is} > 0 \right\}$. This θ is known as the **minimum ratio** in this pivot step. For any i that attains the minimum here, the present i^{th} basic variable y_i is eligible to be the dropping variable from the basic vector in this pivot step. The dropping basic variable can be chosen arbitrarily among those eligible, suppose it is y_r. y_r drops from the basic vector and x_s becomes the r^{th} basic variable in its place. The r^{th} row is the pivot row for this pivot step. The pivot step leads to the canonical tableau with respect to the new basis.

If the pivot column $(\bar{a}_{1s}, \ldots, \bar{a}_{ms})^T$ is placed by the side of the present inverse tableau and a pivot step performed with the element $\bar{a}_{rs}$ in it in the pivot row as the pivot element, the inverse tableau of the present basis gets transformed into the inverse tableau for the new basis.

The purpose of choosing the pivot row, or the dropping variable, by the minimum ratio test, is to guarantee that the basic vector obtained after this pivot step remains feasible.

In this case (when there is at least one positive entry in the pivot column) the pivot step is said to be a nondegenerate pivot step if the minimum ratio computed above is > 0, degenerate pivot step if it is 0. See Examples 2.5, 2.6.

Let B be the basis for (2.3) corresponding to the basic vector $(y_1, \ldots, y_n)$. As discussed above, the basic vector $(y_1, \ldots, y_n)$ is lexico feasible for (2.3) if each row vector of $(\bar{q} \ \vdots \ B^{-1})$ is lexico positive. If the initial basic vector $(y_1, \ldots, y_n)$ is lexico feasible, lexico feasibility can be maintained by choosing the pivot row according to the **lexico minimum ratio test**. Here the pivot row is chosen as the r^{th} row where r is the i that attains the lexico minimum in $\left\{ \frac{(\bar{q}_i, \beta_{i1}, \ldots, \beta_{in})}{\bar{a}_{is}} : i \text{ such that } \bar{a}_{is} > 0 \right\}$, where $\beta = (\beta_{ij}) = B^{-1}$. The lexico minimum ratio test identifies the pivot row (and hence the dropping basic variable) unambiguously, and guarantees that lexico feasibility is maintained after this pivot step. In the simplex algorithm for linear programming, the lexico minimum ratio test is used to guarantee that cycling will not occur under degeneracy (see Chapter 10 of [2.26]). The lexico minimum ratio test is one of the rules that can be used to resolve degeneracy in the simplex algorithm, and thus guarantee that it terminates in a finite number of pivot steps.

Example 2.4 Extreme Half-line

Consider the following canonical tableau with respect to the basic vector (x_1, x_2, x_3, x_4).

basic variables	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	b
x_1	1	0	0	0	1	−1	2	3	3
x_2	0	1	0	0	1	−2	1	−1	0
x_3	0	0	1	0	−1	0	5	4	6
x_4	0	0	0	1	−1	−3	8	2	5

$x_j \geqq 0$ for all j.

Suppose x_6 is the entering variable. The present BFS is $\bar{x} = (3, 0, 6, 5, 0, 0, 0, 0)^T$. The pivot column $(-1, -2, 0, -3)^T$ has no positive entry. Make the entering variable equal to λ, retain all other nonbasic variables equal to 0, this leads to the solution $x(\lambda) = (3+\lambda, 2\lambda, 6, 5+3\lambda, 0, \lambda, 0, 0)^T = \bar{x} + \lambda x^h$, where $x^h = (1, 2, 0, 3, 0, 1, 0, 0)^T$. x^h, the coefficient vector of λ in $x(\lambda)$, is obtained by making the entering variable equal to 1, all other nonbasic variables equal to zero, and each basic variable equal to the negative of the entry in the pivot column in its basic row. Since the pivot column is nonpositive here, $x^h \geqq 0$. it can be verified that x^h satisfies the homogeneous system obtained by replacing the right hand side constants vector by 0. Hence x^h is known as a **homogeneous solution** corresponding to the original system. Since $x^h \geqq 0$ here, $x(\lambda)$ remains $\geqq 0$ for all $\lambda \geqq 0$. The half-line $\{ \bar{x} + \lambda x^h : \lambda \geqq 0 \}$ is known as an **extreme half-line** of the set of feasible solutions of the original system.

A half-line is said to be a **feasible half-line** to a system of linear constraints, if every point on the half-line is feasible to the system.

Example 2.5 Nondegenerate Pivot Step

See Tableau 2.4 in Example 2.8 of Section 2.2.6 a few pages ahead. This is the canonical tableau with respect to the basic vector (w_1, w_2, z_0, w_4) and z_3 is the entering variable. The minimum ratio occurs uniquely in row 4, which is the pivot row in this step, and w_4 is the dropping variable. Performing the pivot step leads to the canonical tableau with respect to the new basic vector (w_1, w_2, z_0, z_3) in Tableau 2.5. This is a nondegenerate pivot step since the minimum ratio in it was $(\frac{4}{2}) > 0$. As a result of this pivot step the BFS has changed from $(w_1, w_2, w_3, w_4; z_1, z_2, z_3, z_4; z_0) = (12, 14, 0, 4; 0, 0, 0, 0; 9)$ to $(6, 8, 0, 0; 0, 0, 2, 0; 5)$.

Example 2.6 Degenerate Pivot Step

Consider the following canonical tableau :

Basic variable	x_1	x_2	x_3	x_4	x_5	x_6	$\bar{b}$	Ratio
x_1	1	0	0	1	2	-3	3	$\frac{3}{1}$
x_2	0	1	0	$\boxed{1}$	-2	1	0	$\frac{0}{1}$ *Min.*
x_3	0	0	1	-1	1	2	0	

$x_j \geqq 0$ for all j.

Here the BFS is $\bar{x} = (3, 0, 0, 0, 0, 0)^T$. It is degenerate. If x_4 is chosen as the entering variable, it can be verified that the minimum ratio of 0 occurs in row 2. Hence row 2 is the pivot row for this step, and x_2 is the dropping variable. Performing the pivot step leads to the canonical tableau with respect to the new basic vector (x_1, x_4, x_3).

basic variable	x_1	x_2	x_3	x_4	x_5	x_6	
x_1	1	-1	0	0	4	-4	3
x_4	0	1	0	1	-2	1	0
x_3	0	1	1	0	-1	3	0

Eventhough the basic vector has changed, the BFS has remained unchanged through this pivot step. A pivot step like this is called a **degenerate pivot step**.

A **pivot step** is **degenerate**, if the minimum ratio θ in it is 0, **nondegenerate** if the minimum ratio is positive and finite. In every pivot step the basic vector changes by one variable. In a degenerate pivot step there is no change in the correponding BFS (the entering variable replaces a zero valued basic variable in the solution). In a nondegenerate pivot step the BFS changes.

Example 2.7 Ties for Minimum Ratio lead to Degenerate Solution

Consider the following canonical tableau.

basic variable	x_1	x_2	x_3	x_4	x_5	x_6	$\bar{b}$	Ratio
x_1	1	0	0	1	-2	1	3	$\frac{3}{1}$
x_2	0	1	0	2	1	1	6	$\frac{6}{2}$
x_3	0	0	1	2	1	-2	16	$\frac{16}{2}$

The present BFS is $\bar{x} = (3, 6, 16, 0, 0, 0)^T$. Suppose x_4 is chosen as the entering variable. There is a tie for the minimum ratio. Both x_1, x_2 are eligible to be dropping variables. Irrespective of which of them is chosen as the dropping variable, it can be verified that the other remains a basic variable with a value of 0 in the next BFS. So the BFS obtained after this pivot step is degenerate.

In the same way it can be verified that the BFS obtained after a pivot step is always degenerate, if there is a tie for the minimum ratio in that step. Thus, if we know that the right hand side constants vector q is nondegenerate in (2.3), in every pivot step performed on (2.3), the minimum ratio test identifies the dropping variable uniquely and unambiguously.

2.2.3 Initialization

The artificial variable z_0 has been introduced into (2.3) for the sole purpose of obtaining a feasible basis to start the algorithm.

Identify row t such that $q_t = \text{minimum } \{ q_i : 1 \leqq i \leqq n \}$. Break ties for t in this equation arbitrarily. Since we assumed $q \ngeqq 0$, $q_t < 0$. When a pivot is made in (2.3) with the column vector of z_0 as the pivot column and the t^{th} row as the pivot row, the right-hand side constants vector becomes a nonnegative vector. The result is the canonical tableau with respect to the basic vector $(w_1, \ldots, w_{t-1}, z_0, w_{t+1}, \ldots, w_n)$. This is the initial basic vector for starting the algorithm.

2.2.4 Almost Complementary Feasible Basic Vectors

The initial basic vector satisfies the following properties :

(i) There is at most one basic variable from each complementary pair of variables (w_j, z_j).

(ii) It constains exactly one basic variable from each of $(n-1)$ complementary pairs of variables, and both the variables in the remaining complementary pair are nonbasic.

(iii) z_0 is a basic variable in it.

A feasible basic vector for (2.3) in which there is exactly one basic variable from each complementary pair (w_j, z_j) is known as a **complementary feasible basic vector**. A feasible basic vector for (2.3) satisfying properties (i), (ii), and (iii) above is known as an **almost complementary feasible basic vector**. Given an almost complementary feasible basic vector for (2.3), the complementary pair both of whose variables are nonbasic, is known as the **left-out complementary pair of variables** in it. All the basic vectors obtained in the algorithm with the possible exception of the

final basic vector are almost complementary feasible basic vectors. If at some stage of the algorithm, a complementary feasible basic vector is obtained, it is a final basic vector and the algorithm terminates.

Adjacent Almost Complementary Feasible Basic Vectors

Let $(y_1, \ldots, y_{j-1}, z_0, y_{j+1}, \ldots, y_n)$ be an almost complementary feasible basic vector for (2.3), where $y_i \in \{w_i, z_i\}$ for each $i \neq j$. Both the variables in the complementary pair (w_j, z_j) are not in this basic vector. Adjacent almost complementary feasible basic vectors can only be obtained by picking as the entering variable either w_j or z_j. Thus from each almost complementary feasible basic vector there are exactly two possible ways of generating adjcent almost complementary feasible basic vectors.

In the initial almost complementary feasible basic vector, both w_t and z_t are nonbasic variables. In the canonical tableau with respect to the initial basis, the updated column vector of w_t can be verified to be $-e_n$, which is negative. Hence, if w_t is picked as the entering variable into the initial basic vector, an extreme half-line is generated. Hence, the initial almost complementary BFS is at the end of an **almost complementary ray**.

So there is a unique way of obtaining an adjacent almost complementary feasible basic vector from the initial basic vector, and that is to pick z_t as the entering variable.

2.2.5 Complementary Pivot Rule

In the subsequent stages of the algorithm there is a unique way to continue the algorithm, which is to pick as the entering variable, the complement of the variable that just dropped from the basic vector. This is known as the **complementary pivot rule**.

The main property of the path generated by the algorithm is the following. Each BFS obtained in the algorithm has two almost complementary edges containing it. We arrive at this solution along one of these edges. And we leave it by the other edge. So the algorithm continues in a unique manner. It is also clear that a basic vector that was obtained in some stage of the algorithm can never reappear.

The path taken by the complementary pivot algorithm is illustrated in Figure 2.1. The initial BFS is that corresponding to the basic vector $(w_1, \ldots, w_{t-1}, z_0, w_{t+1}, \ldots, w_n)$ for (2.3). In Figure 2.1, each BFS obtained during the algorithm is indicated by a point, with the basic vector corresponding to it entered by its side; and consecutive BFSs are joined by an edge. If w_t is choosen as the entering variable into the initial basic vector we get an extreme half-line (discussed above) and the initial BFS is at end of this extreme half-line. When z_t is choosen as the entering variable into the initial basic vector, suppose w_i is the dropping variable. Then its complement z_i will be the entering variable into the next basic vector (this is the complementary pivot rule).

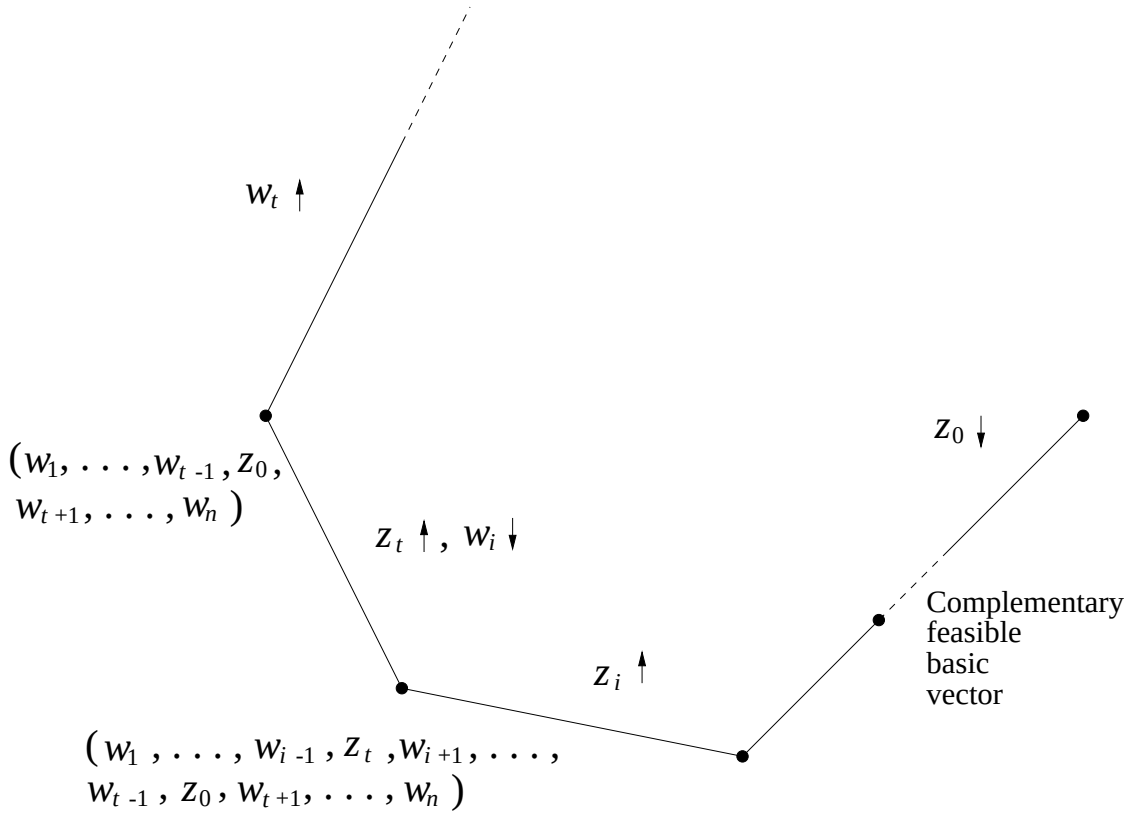


Figure 2.1 Path taken by the complementary pivot method. The ↑ indicates entering variable, ↓ indicates dropping variable. The basic vector corresponding to each point (BFS) is entered by its side. Finally if z_0 drops from the basic vector, we get a complementary feasible basic vector.

The path continues in this unique manner. It can never return to a basic vector visited earlier, since each BFS obtained in the algorithm has exactly two edges of the path incident at it, through one of which we arrive at that BFS and through the other we leave (if the path returns to a basic vector visited earlier, the BFS corresponding to it has three edges in the path incident at it, a contradiction). So the path must terminate after a finite number of steps either by going off along another extreme half-line at the end (**ray termination**, this happens when in some step, the pivot column, the updated column of the entering variable, has no positive entries in it), or by reaching a complementary feasible basic vector of the LCP (which happens when z_0 becomes the dropping variable). If ray termination occurs the extreme half-line obtained at the end, cannot be the same as the initial extreme half-line at the beginning of the path (this follows from the properties of the path discussed above, namely, that it never returns to a basic vector visited earlier).

2.2.6 Termination

There are exactly two possible ways in which the algorithm can terminate.

1. At some stage of the algorithm, z_0 may drop out of the basic vector, or become equal to zero in the BFS of (2.3). If $(\bar{w}, \bar{z}, \bar{z}_0 = 0)$ is the BFS of (2.3) at that stage, then $(\bar{w}, \bar{z})$ is a solution of the LCP (1.6) to (1.8).
2. At some stage of the algorithm, z_0 may be strictly positive in the BFS of (2.3), and the pivot column in that stage may turn out to be nonpositive, and in this case the algorithm terminates with another almost complementary extreme half-line, referred to in some publications as the **secondary ray** (distinct from the initial almost complementary extreme half-line or **initial ray** at the beginning of the algorithm). This is called **ray termination**.

When ray termination occurs, the algorithm is unable to solve the LCP. It is possible that the LCP (1.6) to (1.8) may not have a solution, but if it does have a solution, the algorithm is unable to find it. If ray termination occurs the algorithm is also unable to determine whether a solution to the LCP exists in the general case. However, when M satisfies some conditions, it can be proved that ray termination in the algorithm will only occur, when the LCP has no solution. See Section 2.3.

Problems Posed by Degeneracy of (2.3).

Definition: Nondegeneracy, or Degeneracy of q in the LCP (q, M) As defined earlier, the LCP (q, M) is the problem of finding w, z satisfying

w	z	
I	$-M$	q

$w, z \geqq 0$, $w^T z = 0$

This LCP is said to be nondegenerate (in this case q is said to be nondegenrate in the LCP (q, M)) if in every solution (w, z) of the system of linear equations "$w - Mz = q$", at least n variables are non-zero. This condition holds iff q cannot be expressed as a linear combination of $(n-1)$ or less column vectors of $(I \vdots -M)$.

The LCP (q, M) is said to be degenerate (in this case q is said to be degenerate in the LCP (q, M)) if q can be expressed as a linear combination of a set consisting of $(n-1)$ or less column vectors of $(I \vdots -M)$.

Definition: Nondegeneracy, Degeneracy of q in the Complementary Pivot Algorithm The system of contraints on which pivot operations are performed in the complementary pivot algorithm is (2.3). This system is said to be degenerate (and q is said to be degenerate in it) if q can be expressed as a linear combination of a set of $(n-1)$ or less column vectors of $(I \vdots -M \vdots -e)$; nondegenerate otherwise. If

(2.3) is nondegenerate, in every BFS of (2.3) obtained during the complementary pivot algorithm, all basic variables are strictly positive, and the minimum ratio test identifies the dropping basic variable in each pivot step uniquely and unambiguously.

The argument that each almost complementary feasible basis has at most two adjacent almost complementary feasible bases is used in developing the algorithm. This guarantees that the path taken by the algorithm continues unambiguously in a unique manner till termination occurs in one of the two possibilities. This property that each almost complementary feasible basis has at most two adjacent almost complementary feasible bases holds when (2.3) is nondegenerate. If (2.3) is degenerate, the dropping variable during some pivots may not be uniquely determined. In such a pivot step, by picking different dropping variables, different adjacent almost complementary feasible bases may be generated. If this happens, the almost complementary feasible basis in this step may have more than two adjacent almost complementary feasible bases. The algorithm can still be continued unambiguously according to the complementary pivot rule, but the path taken by the algorithm may depend on the dropping variables selected during the pivots in which these variables are not uniquely identified by the minimum ratio test. All the arguments mentioned in earlier sections are still valid, but in this case termination may not occur in a finite number of steps if the algorithm keeps cycling along a finite sequence of degenerate pivot steps. This can be avoided by using the concept of **lexico feasibility** of the solution. In this case the algorithm deals with **almost complementary lexico feasible bases** throughout. In each pivot step the lexico minimum ratio test determines the dropping variable unambiguously and, hence, each almost complementary lexico feasible basis can have at most two adjacent almost complementary lexico feasible bases. With this, the path taken by the algorithm is again unique and unambiguous, no cycling can occur and termination occurs after a finite number of pivot steps. See Section 2.2.8.

Interpretation of the Path Taken by the Complementary Pivot Algorithm

B. C. Eaves has given a simple haunted house interpretation of the path taken by the complementary pivot algorithm. A man who is afraid of ghosts has entered a haunted house from the outside through a door in one of its rooms. The house has the following properties :

(i) It has a finite number of rooms.

(ii) Each door is on a boundary wall between two rooms or on a boundary wall of a room on the outside.

(iii) Each room may have a ghost in it or may not. However, every room which has a ghost has exactly two doors.

All the doors in the house are open initially. The man's walk proceeds according to the following property.

(iv) When the man walks through a door, it is instantly sealed permanently and he can never walk back through it.

The man finds a ghost in the room he has entered initially, by properties (iii) and (iv) this room has exactly one open door when the man is inside it. In great fear he runs out of the room through that door. If the next room that he has entered has a ghost again, it also satisfies the property that it has exactly one open door when the man is inside it, and he runs out through that as fast as he can. In his walk, every room with a ghost satisfies the same property. He enters that room through one of its doors and leaves through the other. A **sanctuary** is defined to be either a room that has no ghost, or the outside of the house. The man keeps running until he finds a sanctuary. Property (i) guarantees that the man finds a sanctuary after running through at most a finite number of rooms. The sanctuary that he finds may be either a room without a ghost or the outside of the house.

We leave it to the reader to construct parallels between the ghost story and the complementary pivot algorithm and to find the walk of the man through the haunted house in Figure 2.2. The man walks into the house initially from the outside through the door marked with an arrow.

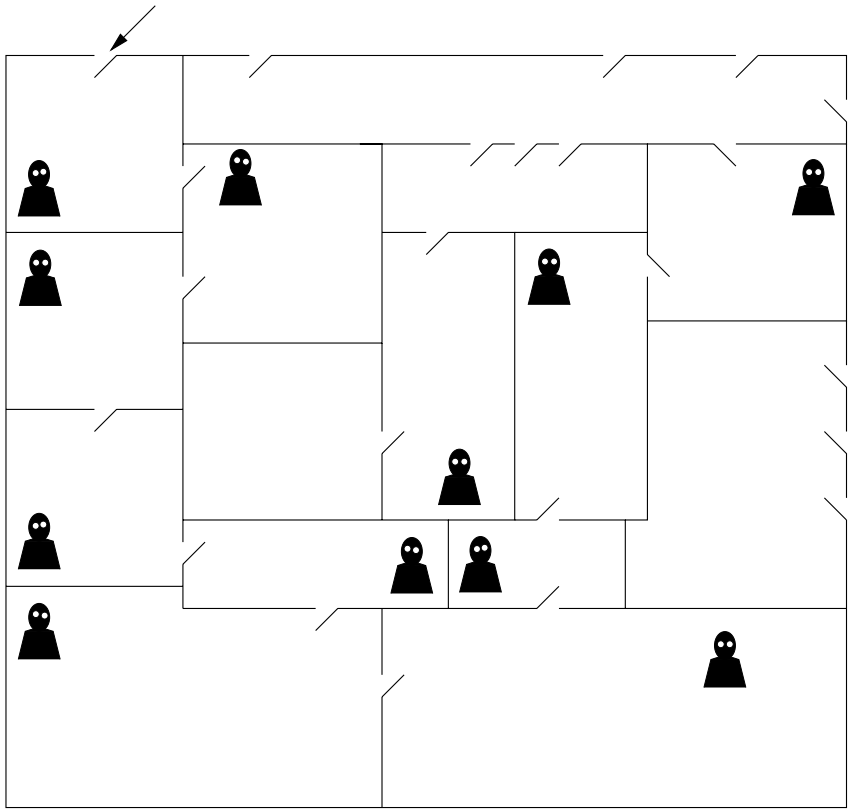


Figure 2.2 Haunted house

Geometric Interpretation of a Pivot Step in the Complementary Pivot Method

In a pivot step of the complementary pivot method, the current point moves between two facets of a complementary cone in the direction of $-e$. This geometric interpreta-

tion of a pivot step in the complementary pivot method as a walk between two facets of a complementary cone is given in Section 6.2.

Example 2.8

Consider the following LCP. (This is not an LCP corresponding to an LP.)

w_1	w_2	w_3	w_4	z_1	z_2	z_3	z_4	q
1	0	0	0	−1	1	1	1	3
0	1	0	0	1	−1	1	1	5
0	0	1	0	−1	−1	−2	0	−9
0	0	0	1	−1	−1	0	−2	−5

$w_i \geqq 0, \quad z_i \geqq 0, \quad w_i z_i = 0$ for all i

When we introduce the artificial variable z_0 the tableau becomes :

w_1	w_2	w_3	w_4	z_1	z_2	z_3	z_4	z_0	q
1	0	0	0	−1	1	1	1	−1	3
0	1	0	0	1	−1	1	1	−1	5
0	0	1	0	−1	−1	−2	0	[−1]	−9
0	0	0	1	−1	−1	0	−2	−1	−5

The most negative q_i is q_3. Therefore pivot in the column vector of z_0 with the third row as the pivot row. The pivot element is inside a box.

Tableau 2.4

Basic variables	w_1	w_2	w_3	w_4	z_1	z_2	z_3	z_4	z_0	q	Ratios
w_1	1	0	−1	0	0	2	3	1	0	12	$\frac{12}{3}$
w_2	0	1	−1	0	2	0	3	1	0	14	$\frac{14}{3}$
z_0	0	0	−1	0	1	1	2	0	1	9	$\frac{9}{2}$
w_4	0	0	−1	1	0	0	[2]	−2	0	4	$\frac{4}{2}$ *Min.*

By the complementary pivot rule we have to pick z_3 as the entering variable. The column vector of z_3 is the pivot column, w_4 drops from the basic vector.

Tableau 2.5

Basic variables	w_1	w_2	w_3	w_4	z_1	z_2	z_3	z_4	z_0	q	Ratios
w_1	1	0	$\frac{1}{2}$	$-\frac{3}{2}$	0	2	0	$\boxed{4}$	0	6	$\frac{6}{4}$ *Min.*
w_2	0	1	$\frac{1}{2}$	$-\frac{3}{2}$	2	0	0	4	0	8	$\frac{8}{4}$
z_0	0	0	0	-1	1	1	0	2	1	5	$\frac{5}{2}$
z_3	0	0	$-\frac{1}{2}$	$\frac{1}{2}$	0	0	1	-1	0	2	

Since w_4 has dropped from the basic vector, its complement, z_4 is the entering variable for the next step. w_1 drops from the basic vector.

Basic variables	w_1	w_2	w_3	w_4	z_1	z_2	z_3	z_4	z_0	q	Ratios
z_4	$\frac{1}{4}$	0	$\frac{1}{8}$	$-\frac{3}{8}$	0	$\frac{1}{2}$	0	1	0	$\frac{6}{4}$	
w_2	-1	1	0	0	$\boxed{2}$	-2	0	0	0	2	$\frac{2}{2}$ *Min.*
z_0	$-\frac{1}{2}$	0	$-\frac{1}{4}$	$-\frac{1}{4}$	1	0	0	0	1	2	$\frac{2}{1}$
z_3	$\frac{1}{4}$	0	$-\frac{3}{8}$	$\frac{1}{8}$	0	$\frac{1}{2}$	1	0	0	$\frac{14}{4}$	

Since w_1 has dropped from the basic vector, its complement, z_1 is the new entering variable. Now w_2 drops from the basic vector.

Basic variables	w_1	w_2	w_3	w_4	z_1	z_2	z_3	z_4	z_0	q	Ratios
z_4	$\frac{1}{4}$	0	$\frac{1}{8}$	$-\frac{3}{8}$	0	$\frac{1}{2}$	0	1	0	$\frac{6}{4}$	3
z_1	$-\frac{1}{2}$	$\frac{1}{2}$	0	0	1	-1	0	0	0	1	
z_0	0	$-\frac{1}{2}$	$-\frac{1}{4}$	$-\frac{1}{4}$	0	$\boxed{1}$	0	0	1	1	1 *Min.*
z_3	$\frac{1}{4}$	0	$-\frac{3}{8}$	$\frac{1}{8}$	0	$\frac{1}{2}$	1	0	0	$\frac{14}{4}$	7

Since w_2 has dropped from the basic vector, its complement, z_2 is the entering variable. Now z_0 drops from the basic vector.

Basic variables	w_1	w_2	w_3	w_4	z_1	z_2	z_3	z_4	z_0	q
z_4	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$-\frac{1}{4}$	0	0	0	1	$-\frac{1}{2}$	1
z_1	$-\frac{1}{2}$	0	$-\frac{1}{4}$	$-\frac{1}{4}$	1	0	0	0	1	2
z_2	0	$-\frac{1}{2}$	$-\frac{1}{4}$	$-\frac{1}{4}$	0	1	0	0	1	1
z_3	$\frac{1}{4}$	$\frac{1}{4}$	$-\frac{1}{4}$	$\frac{1}{4}$	0	0	1	0	$-\frac{1}{2}$	3

Since the present basis is a complementary feasible basis, the algorithm terminates. The correponding solution of the LCP is $w = 0$, $(z_1, z_2, z_3, z_4) = (2, 1, 3, 1)$.

Example 2.9

---

w_1	w_2	w_3	z_1	z_2	z_3	q
1	0	0	1	0	3	-3
0	1	0	-1	2	5	-2
0	0	1	2	1	2	-1

$w_i \geqq 0, \quad z_i \geqq 0, \quad w_i z_i = 0 \qquad$ for all i

The tableau with the artificial variable z_0 is :

w_1	w_2	w_3	z_1	z_2	z_3	z_0	q
1	0	0	1	0	3	$\boxed{-1}$	-3
0	1	0	-1	2	5	-1	-2
0	0	1	2	1	2	-1	-1

The initial canonical tableau is :

Basic variables	w_1	w_2	w_3	z_1	z_2	z_3	z_0	q	Ratios
z_0	-1	0	0	-1	0	-3	1	3	
w_2	-1	1	0	-2	2	2	0	1	
w_3	-1	0	1	$\boxed{1}$	1	-1	0	2	$\frac{2}{1}$

The next tableau is :

Basic variables	w_1	w_2	w_3	z_1	z_2	z_3	z_0	q
z_0	-2	0	1	0	1	-4	1	5
w_2	-3	1	2	0	4	0	0	5
z_1	-1	0	1	1	1	-1	0	2

The entering variable here is z_3. The pivot column is nonpositive. Hence, the algorithm stops here with ray termination. The algorithm has been unable to solve this LCP.

2.2.7 IMPLEMENTATION OF THE COMPLEMENTARY PIVOT METHOD USING THE INVERSE OF THE BASIS

Let (2.3) be the original tableau for the LCP being solved by the complementary pivot method. Let t be determined as in Section 2.2.3. After performing the pivot with row t as the pivot row and the column vector of z_0 as the pivot column, we get the **initial tableau** for this algorithm. Let P_0 be the pivot matrix of order n obtained by replacing the t^{th} column in I (the unit matrix of order n) by $-e_n$ (the column vector in $\mathbf{R}^n$ all of whose entires are -1). Let $M' = P_0M$, $q' = P_0q$. Then the initial tableau in this algorithm is

Tableau 2.6 : Initial Tableau

w	z	z_0	
P_0	$-M'$	$I_{.t}$	q'

The initial basic vector is $(w_1, \ldots, w_{t-1}, z_0, w_{t+1}, \ldots, w_n)$ and the basis corresponding to it in Tableau 2.6 is I. By choice of t, $q' \geq 0$. So each row of $(q' \ \vdots \ I)$ is lexico-positive, and hence the initial basic vector in this algorithm is lexico-feasible for the problem in Tableau 2.6.

At some stage of the algorithm, let B be the basis from Tableau 2.6, corresponding to the present basic vector. Let $\beta = (\beta_{ij}) = B^{-1}$ and $\bar{q} = B^{-1}q'$. Then the inverse tableau at this stage is

Basic vector	Inverse	
	$\beta = B^{-1}$	$\bar{q}$

If the entering variable in this step, determined by the complementary pivot rule, is $y_s \in \{ w_s, z_s \}$, then the pivot column, the updated column of y_s, is $\beta P_0 I_{.s}$ if $y_s = w_s$, or $\beta P_0(-M_{.s})$ if $y_s = z_s$. Suppose this pivot column is $(\bar{a}_{1s}, \ldots, \bar{a}_{ns})^T$. If $(\bar{a}_{1s}, \ldots, \bar{a}_{ns})^T \leq 0$, we have ray termination and the method has been unable to solve this LCP. If $(\bar{a}_{1s}, \ldots, \bar{a}_{ns})^T \not\leq 0$, the minimum ratio in this step is $\theta =$ minimum $\left\{ \frac{\bar{q}_i}{\bar{a}_{is}} : i \text{ such that } \bar{a}_{is} > 0 \right\}$. If the i that attains this minimum is unique, it determines the pivot row uniquely. The present basic variable in the pivot row is the dropping variable. If the minimum ratio does not identify the dropping variable uniquely, check whether z_0 is eligible to drop, and if so choose it as the dropping variable. If z_0 is not eligible to drop, one of those eligible to drop can be choosen as the dropping variable arbitrarily, but this can lead to cycling under degeneracy. To avoid cycling, we can use the lexico-minimum ratio rule, which chooses the dropping basic variable so that the pivot row is the row corresponding to the lexico-minimum among $\left\{ \frac{(\bar{q}_i; \beta_{i1}, \ldots, \beta_{in})}{\bar{a}_{is}} : i \text{ such that } \bar{a}_{is} > 0 \right\}$. This lexico minimum ratio rule determines the dropping variable uniquely and unambiguously. If the lexico-minimum ratio rule is used in all steps beginning with the initial step, the dropping variable is identified uniquely in every step, each of the updated vectors $(\bar{q}_i; \beta_{i1}, \ldots, \beta_{in})$, $i = 1$ to n, remain lexico-positive throught, and cycling cannot occur by the properties of the almost complementary path generated by this method, discussed above (see Section 2.2.8). Once the dropping variable is identified, performing the pivot leads to the next basis inverse, and the entering variable in the next step is the complement of the dropping variable, and the method is continued in the same way.

Clearly it is not necessary to maintain the basis inverse explicitly. The complementary pivot algorithm can also be implemented with the basis inverse maintained in product form (PFI) or in elimination form (EFI) just as the simplex algorithm for linear programming (see Chapters 5, 7 of [2.26]).

Example 2.10

Consider the LCP (q, M) where

$$M = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 2 & 2 & 1 \end{pmatrix} \qquad q = \begin{pmatrix} -8 \\ -12 \\ -14 \end{pmatrix}$$

To solve this LCP by the complementary pivot algorithm, we introduce the artificial variable z_0 and construct the original tableau as in (2.3). When z_0 replaces w_3 in the basic vector (w_1, w_2, w_3), we get a feasible basic vector for the original tableau. So the initial tableau for this problem is :

Initial Basic Vector	w_1	w_2	w_3	z_1	z_2	z_3	z_0	q
w_1	1	0	-1	1	2	1	0	6
w_2	0	1	-1	0	1	1	0	2
z_0	0	0	-1	2	2	1	1	14

The various basis inverses obtained when this LCP is solved by the complementary pivot algorithm are given below.

Basic Vector	Inverse			$\bar{q}$	Pivot Column	Ratios
					z_3	
w_1	1	0	0	6	1	6
w_2	0	1	0	2	$\boxed{1}$	2 *Min.*
z_0	0	0	1	14	1	14
					z_2	
w_1	1	-1	0	4	1	4
z_3	0	1	0	2	$\boxed{1}$	2 *Min.*
z_0	0	-1	1	12	1	12
					w_3	
w_1	1	-2	0	2	$\boxed{1}$	2 *Min.*
z_2	0	1	0	2	-1	
z_0	0	-2	1	10	1	10
					z_1	
w_3	1	-2	0	2	$\boxed{1}$	2 *Min.*
z_2	1	-1	0	4	1	4
z_0	-1	0	1	8	1	8

Basic Vector	Inverse			$\bar{q}$	Pivot Column	Ratios
					z_3	
z_1	1	−2	0	2	−1	
z_2	0	1	0	2	[1]	2 *Min.*
z_0	−2	2	1	6	1	6
					w_2	
z_1	1	−1	0	4	−1	
z_3	0	1	0	2	[1]	2 *Min.*
z_0	−2	1	1	4	1	4
					w_3	
z_1	1	0	0	6	−1	
w_2	0	1	0	2	−1	
z_0	−2	0	1	2	[1]	2 *Min.*
z_1	−1	0	1	8		
w_2	−2	1	1	4		
w_3	−2	0	1	2		

So the solution of this LCP is $(w_1, w_2, w_3; z_1, z_2, z_3) = (0, 4, 2; 8, 0, 0)$.

2.2.8 Cycling Under Degeneracy in the Complementary Pivot Method

Whenever there is a tie for the pivot row in any step of the complementary pivot method, suppose we adopt the rule that the pivot row will be chosen to be the topmost among those eligible for it in that step. Under this rule it is possible that cycling occurs under degeneracy. Here we provide an example of cycling under this rule, constructed by M. M. Kostreva [2.20]. Let

$$M = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 2 \\ 2 & 0 & 1 \end{pmatrix} \qquad q = \begin{pmatrix} -1 \\ -1 \\ -1 \end{pmatrix}$$

and solve the LCP (q, M) by the complementary pivot method using the above pivot row choice rule in each pivot step. It can be verified that we get the following almost complementary feasible basic vectors: initial basic vector (z_0, w_2, w_3) followed by (z_0, z_1, w_3), (z_0, z_2, w_3), (z_0, z_2, w_1), (z_0, z_3, w_1), (z_0, z_3, w_2), (z_0, z_1, w_2), (z_0, z_1, w_3), in this order. After the initial basic vector (z_0, w_2, w_3) is obtained, all pivots made are degenerate pivot steps, and at the end the method has returned to the basic vector (z_0, z_1, w_3) and so the method has cycled on this problem. The matrix M is a P-matrix, it will be proved later on the LCP (q, M) has a unique solution, and that the complementary pivot method always terminates in a finite number of pivot steps with that solution, if it is carried out in such a way that cycling does not occur under degeneracy. Actually, for the LCP (q, M) considered here, it can be verified that (z_1, z_2, z_3) is the complementary feasible basic vector.

As discussed above, after obtaining the initial basic vector, if the complementary pivot method is carried out using the lexico-minimum ratio rule for choosing the pivot row in each pivot step, cycling cannot occur, and the method must terminate either by obtaining a complementary feasible vector, or in ray termination, after a finite number of pivot steps, because of the following arguments. If q is nondegenerate in (2.3), the dropping basic variable is identified uniquely by the usual minimum ratio test, in every step of the complementary pivot algorithm applied on it. Using the properties of the path traced by this algorithm we verify that in this case, the algorithm must terminate after a finite number of pivot steps either with a complementary feasible basic vector or in ray termination. Suppose q is degenerate in (2.3). Perturb (2.3) by replacing q by $q(\varepsilon) = q + (\varepsilon, \varepsilon^{\mathbf{2}}, \ldots, \varepsilon^{\mathbf{n}})^T$, as in (2.2). When ε is positive but sufficiently small, the perturbed problem is nondegenerate. So when the perturbed problem is solved by the complementary pivot algorithm treating $\varepsilon > 0$ to be sufficiently small, it must terminate in a finite number of pivot steps. If a complementary feasible basic vector is obtained at the end for the perturbed problem, that basic vector is also a complementary basic vector for the original LCP (unperturbed original problem, with $\varepsilon = 0$). If ray termination occurs at the end on the perturbed problem, the final almost complementary feasible basic vector is also feasible to the original LCP and satisfies the condition for ray termination in it. The sequence of basic vectors obtained when the complementary pivot algorithm is applied on the original problem (2.3) using the lexico-minimum ratio rule for chosing the dropping variable in every pivot step, is exactly the same as the sequence of basic vectors obtained when the complementary pivot algorithm is applied on the perturbed problem got by replacing q in (2.3) by $q(\varepsilon)$ with $\varepsilon > 0$ and sufficiently small. These facts show that the complementary pivot algorithm must terminate in a finite number of pivot steps (i. e., can not cycle) when operated with the lexico minimum ratio test for chosing the dropping variable in every pivot step.

2.3 CONDITIONS UNDER WHICH THE COMPLEMENTARY PIVOT ALGORITHM WORKS

We define several classes of matrices that are useful in the study of the LCP. Let $M = (m_{ij})$ be a square matrix of order n. It is said to be a

Copositive matrix if $y^T My \geqq 0$ for all $y \geqq 0$.

Strict copositive matrix if $y^T My > 0$ for all $y \geq 0$.

Copositive plus matrix if it is a copositive matrix and whenever $y \geqq 0$, and satisfies $y^T My = 0$, we have $y^T(M + M^T) = 0$.

P-matrix if all its principal subdeterminantes are positive.

Q-matrix if the LCP (q, M) has a solution for every $q \in \mathbf{R}^n$.

Negative definite matrix if $y^T My < 0$ for all $y \neq 0$.

Negative semidefinite matrix if $y^T My \leqq 0$ for all $y \in \mathbf{R}^n$.

Z-matrix if $m_{ij} \leqq 0$ for all $i \neq j$

Principally nondegenerate matrix if all its principal subdeterminants are non-zero.

Principally degenerate matrix if at least one of its principal subdeterminants is zero.

L_1-matrix if for every $y \geq 0$, $y \in \mathbf{R}^n$, there is an i such that $y_i > 0$ and $M_{i.}y \geqq 0$. If M is an L_1-matrix, an i like it is called a **defining index** for M and y. These matrices are also called **semimonotone matrices**.

L_2-matrix if for every $y \geq 0$, $y \in \mathbf{R}^n$, such that $My \geqq 0$ and $y^T My = 0$, there are diagonal matrices, $\Lambda \geqq 0$, $\Omega \geqq 0$ such that $\Omega y \neq 0$ and $(\Lambda M + M^T\Omega)y = 0$. An equivalent definition is that for each $z \geq 0$, satisfying $w = Mz \geqq 0$ and $w^T z = 0$; there exists a $\hat{z} \geq 0$ satisfying $\hat{w} = -(\hat{z}^T M)^T$, $w \geqq \hat{w} \geqq 0$, $z \geqq \hat{z} \geqq 0$.

L-matrix if it is both an L_1-matrix and an L_2-matrix.

$L_\star$-matrix if for every $y \geq 0$, $y \in \mathbf{R}^n$, there is an i such that $y_i > 0$ and $M_{i.}y > 0$. If M is an $L_\star$-matrix, an i like it is called a **defining index** for M and y.

P_0-matrix if all its principal subdeterminants are $\geqq 0$.

Row adequate matrix if it is a P_0-matrix and whenever the principal subdeterminant corresponding to some subset $\mathbf{J} \subset \{1, \ldots, n\}$ is zero, then the set of row vectors of M corresponding to $\mathbf{J}$, $\{M_{i.} : i \in \mathbf{J}\}$ is linearly dependent.

Column adequate matrix if it is a P_0-matrix and whenever the principal subdeterminant corresponding to some subset $\mathbf{J} \subset \{1, \ldots, n\}$ is zero, then

the set of column vectors of M corresponding to $\mathbf{J}$, $\{ M_{.j} : j \in \mathbf{J} \}$ is linearly dependent.

Adequate matrix if it is both row and column adequate.

In this book the only type of degeneracy, nondegeneracy of square matrices that we discuss is principal degeneracy or principal nondegeneracy defined above. So, for notational convenience we omit the term "principally" and refer to these matrices as being degenerate or nondegenerate matrices. Examples of degenerate matrices are $\begin{pmatrix} 0 & 4 \\ 3 & -10 \end{pmatrix}$, $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$. Examples of nondegenerate matrices are $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$, $\begin{pmatrix} 3 & 1 \\ 0 & -2 \end{pmatrix}$. The notation C_0-**matrix** is used to denote copositive matrices, and the notation C_+-**matrix** is used to denote copositive plus matrices.

Theorem 1.11 implies that every PSD matrix is also a copositive plus matrix. Also, the square matrix M is negative definite or negative semi-definite, iff $-M$ is PD or PSD respectively.

2.3.1 Results on LCPs Associated with Copositive Plus Matrices

Theorem 2.1 *If M is a copositive plus matrix and the system of constraints (1.6) and (1.7) of Section 1.1.3 has a feasible solution, then the LCP (1.6) — (1.8) has a solution and the complementary pivot algorithm will terminate with the complementary feasible basis. Conversely, when M is a copositive plus matrix, if the complementary pivot algorithm applied on (1.6) — (1.8) terminates in ray termination, the system of constraints (1.6), (1.7) must be infeasible.*

Proof. Assume that either (2.3) is nondegenerate, or that the lexico-minimum ratio rule is used throughout the algorithm to determine the dropping basic variable in each step of the algorithm. This implies that each almost complementary feasible (or lexico feasible) basis obtained during the algorithm has exactly two adjacent almost complementary feasible (or lexico feasible) bases, excepting the initial and terminal bases, which have exactly one such adjacent basis only. The complementary pivot algorithm operates on the system (2.3).

The initial basic vector is $(w_1, \ldots, w_{t-1}, z_0, w_{t+1}, \ldots, w_n)$ (as in Section 2.2.3). The corresponding BFS is $z = 0$, $w_t = 0$, $z_0 = -q_t$, and $w_i = q_i - q_t$ for all $i \neq t$. If w_t is taken as the entering variable into this basic vector, it generates the half-line (called the **initial extreme half-line**)

$$\begin{aligned} w_i &= q_i - q_t + \lambda \qquad \text{for all } i \neq t \\ w_t &= \lambda \\ z &= 0 \\ z_0 &= -q_t + \lambda \end{aligned}$$

where $\lambda \geqq 0$. (This can be seen by obtaining the canonical tableau corresponding to the initial basic vector.) This initial extreme half-line contains the initial BFS of (2.3) as its end point. Among the basic vectors obtained during the algorithm, the only one that can be adjacent to the initial basic vector is the one obtained by introducing z_t into it. Once the algorithm moves to this adjacent basic vector, the initial basic vector will never again appear during the algorithm. Hence, if the algorithm terminates with ray termination, the extreme half-line obtained at termination cannot be the initial extreme half-line.

At every point on the initial extreme half-line all the variables w, z_0 are strictly positive. It is clear that the only edge of (2.3) that contains a point in which all the variables w, z_0 are strictly positive is the initial extreme half-line.

Suppose the algorithm terminates in ray termination without producing a solution of the LCP. Let B_k be the terminal basis. When the complementary pivot algorithm is continued from this basis B_k, the updated column vector of the entering variable must be nonpositive resulting in the generation of an extreme half-line. Let the terminal extreme half-line be

$$\left\{ (w, z, z_0) = (w^k + \lambda w^h, z^k + \lambda z^h, z_0^k + \lambda z_0^h) : \quad \lambda \geqq 0 \right\} \tag{2.5}$$

where (w^k, z^k, z_0^k) is the BFS of (2.3) with respect to the terminal basis B_k, and (w^h, z^h, z_0^h) is a homogeneous solution corresponding to (2.3) that is,

$$\begin{aligned} & w^h - Mz^h - e_n z_0^h = 0 \\ & w^h \geqq 0, \quad z^h \geqq 0, \quad z_0^h \geqq 0 \end{aligned} \tag{2.6}$$

$(w^h, z^h, z_0^h) \neq 0$. If $z^h = 0$, (2.6) and the fact that $(w^h, z^h, z_0^h) \neq 0$ together imply that $w^h \neq 0$ and hence $z_0^h > 0$, and consequently $w^h > 0$. Hence, if $z^h = 0$, points on this terminal extreme half-line have all the variables w, z_0 strictly positive, which by earlier arguments would imply that the terminal extreme half-line is the initial extreme half-line, a contradiction. So $z^h \neq 0$.

Since every solution obtained under the algorithm satisfies the complementarity constraint, $w^T z = 0$, we must have $(w^k + \lambda w^h)^T (z^k + \lambda z^h) = 0$ for all $\lambda \geqq 0$. This implies that $(w^k)^T z^k = (w^k)^T z^h = (w^h)^T z^k = (w^h)^T z^h = 0$. From (2.6) $(w^h)^T = (Mz^h + e_n z_0^h)^T$. Hence from $(w^h)^T z^h = 0$, we can conclude that $(z^h)^T M^T z^h = (z^h)^T M z^h = -e_n^T z^h z_0^h \leqq 0$. Since $z^h \geqq 0$, and M is copositive plus by the hypothesis, $(z^h)^T M z^h$ cannot be < 0, and hence, by the above, we conclude that $(z^h)^T M z^h = 0$. This implies that $(z^h)^T (M + M^T) = 0$, by the copositive plus property of M. So $(z^h)^T M = -(z^h)^T M^T$. Also since $-e_n^T z^h z_0^h = (z^h)^T M z^h = 0$, z_0^h must be zero (since $z^h \geq 0$). Since (w^k, z^k, z_0^k) is the BFS of (2.3) with respect to the feasible basis B_k, $w^k = Mz^k + q + e_n z_0^k$. Now

$$\begin{aligned}
0 = (w^k)^T z^h &= (Mz^k + q + e_n z_0^k)^T z^h \\
&= (z^k)^T M^T z^h + q^T z^h + z_0^k e_n^T z^h \\
&= (z^h)^T M z^k + (z^h)^T q + z_0^k e_n^T z^h \\
&= -(z^h)^T M^T z^k + (z^h)^T q + z_0^k e_n^T z^h \\
&= -(z^k)^T M z^h + (z^h)^T q + z_0^k e_n^T z^h \\
&= -(z^k)^T w^h + (z^h)^T q + z_0^k e_n^T z^h \\
&= (z^h)^T q + z_0^k e_n^T z^h
\end{aligned}$$

So $(z^h)^T q = -z_0^k e_n^T z^h$. Since $z^h \geq 0$ and $z_0^k > 0$ [otherwise (w^k, z^k) would be a solution of the LCP], $z_0^k e_n^T z^h > 0$. Hence, $(z^h)^T q < 0$. Hence, if $\pi = (z^h)^T$ we have, $\pi q < 0$, $\pi \geqq 0$, $\pi(-M) = -(z^h)^T M = (z^h)^T M^T = (w^h)^T \geqq 0$, that is,

$$\pi q < 0$$

$$\pi(I \ \vdots \ -M) \geqq 0$$

By Farakas lemma (Theorem 3 of Appendix 1), this implies that the system :

$$(I \ \vdots \ -M) \begin{pmatrix} w \\ \dots \\ z \end{pmatrix} = q \ , \quad \begin{pmatrix} w \\ \dots \\ z \end{pmatrix} \geqq 0$$

has no feasible solution. Hence, if the complementary pivot algorithm terminates in ray termination, the system (1.6) and (1.7) has no feasible solutions in this case and thus there cannot be any solution to the LCP.

This also implies that whenever (1.6) and (1.7) have a feasible solution, the LCP (1.6) to (1.8) has a solution in this case and the complementary pivot algorithm finds it.

□

The following results can be derived as corollaries.

Result 2.1 *In the LCPs corresponding to LPs and convex quadratic programs, the matrix M is* PSD *and hence copositive plus. Hence, if the complementary pivot algorithm applied to the LCP corresponding to an LP or a convex quadratic program terminates in ray termination, that LP or convex quadratic program must either be infeasible, or if it is feasible, the objective function must be unbounded below on the set of feasible solutions of that problem.*

Hence the complementary pivot algorithm works when used to solve LPs or convex quadratic programs.

Result 2.2 *If M is strict copositive the complementary pivot algorithm applied on (1.6) to (1.8) terminates with a solution of the LCP.*

Proof. If the complementary pivot algorithm terminates in ray termination, as seen in the proof of the above theorem there exists a $z^h \geq 0$ such that $(z^h)^T M z^h = 0$, contradicting the hypothesis that M is strict copositive.

□

Thus all strict copositive matrices are Q-matrices. Also, if $M = (m_{ij}) \geqq 0$ and $m_{ii} > 0$ for all i, M is strict copositive and hence a Q-matrix.

Exercise

2.1 Suppose $M \geqq 0$ and $m_{11} = 0$. Prove that if $q = (-1, 1, \ldots, 1)^T$, the LCP (1.6) to (1.8) cannot have a solution. Thus prove that a square nonnegative matrix is a Q-matrix iff all its diagonal entries are strictly positive.

Later on we prove that if M is a P-matrix, the complementary pivot algorithm terminates with a complementary feasible solution when applied on the LCP (q, M). When the complementary pivot algorithm is applied on a LCP in which the matrix M is not a copositive plus matrix or a P-matrix, it is still possible that the algorithm terminates with a complementary feasible basis for the problem. However, in this general case it is also possible that the algorithm stops with ray termination even if a solution to the LCP exists.

To Process an LCP (q, M)

An algorithm for solving LCPs is said to **process** a particular LCP (q, M) for given q and M, if the algorithm is guaranteed to either determine that the LCP (q, M) has no solution, or find a solution for it, after a finite amount of computational effort.

Suppose M is a copositive plus matrix, and consider the LCP (q, M), for given q. When the complementary pivot algorithm is applied on this LCP (q, M), either it finds a solution; or ends up in ray termination which implies that this LCP has no solution by the above theorem. Hence, the complementary pivot algorithm processes the LCP (q, M) whenever M is a copositive plus matrix.

2.3.2 Results on LCPs Associated with L- and $L_\star$-Matrices

Here we show that the complementary pivot algorithm will process the LCP (q, M) whenever M is an L- or $L_\star$-matrix. The results in this section are from B. C. Eaves [2.8, 2.9], they extend the results proved in Section 2.3.1 considerably. Later on, in Section 2.9.2 we derive some results on the general nonconvex programming problem using those proved in this section.

Lemma 2.1 *If M is an L_1-matrix, the LCP (q, M) has a unique solution for all $q > 0$, and conversely.*

Proof. When $q > 0$, one solution of the LCP (q, M) is $(w = q, z = 0)$. So if $(\bar{w}, \bar{z})$ is an alternate solution, we must have $\bar{z} \geq 0$. But $\bar{w} - M\bar{z} = q$. Let M be an L_1-matrix and let i be the defining index for M and $\bar{z}$. We have

$$\bar{w}_i = (M\bar{z})_i + q_i > 0$$

So $\bar{w}_i \bar{z}_i > 0$, contradiction to complementarity.

Now suppose M is not an L_1-matrix. So, there must exist a $\bar{y} = (\bar{y}_i) \geq 0$ such that for all i such that $\bar{y}_i > 0$, $M_{i.}\bar{y} < 0$. Let $\mathbf{J} = \{ i : \bar{y}_i > 0 \}$. Select a positive number α such that $\alpha > \{ |M_{i.}\bar{y}| : i \notin \mathbf{J} \}$. Define the vector $q = (q_j) \in \mathbf{R}^n$ by

$$q_j = \begin{cases} -M_{j.}\bar{y}, & \text{for all } j \in \mathbf{J} \\ \alpha, & \text{for all } j \notin \mathbf{J}. \end{cases}$$

Then $q > 0$ and the LCP (q, M) has two distinct solutions namely $(w, z) = (q, 0)$ and $\big(\bar{w} = (\bar{w}_j), \bar{z} = \bar{y}\big)$, where

$$\bar{w}_j = \begin{cases} 0, & \text{for all } j \in \mathbf{J} \\ \alpha + M_{j.}\bar{y}, & \text{for all } j \notin \mathbf{J}. \end{cases}$$

This establishes the converse.

□

Lemma 2.2 *If M is an $L_\star$-matrix, the LCP (q, M) has a unique solution for every $q \geqq 0$, and conversely.*

Proof. Similar to Lemma 2.1.

□

Lemma 2.3 *If M is an L_2-matrix and the complementary pivot method applied on the LCP (q, M) terminates with the secondary ray $\{ (\bar{w}^k, z^k, z_0^k) + \lambda(w^h, z^h, z_0^h) : \lambda \geqq 0 \}$ as in (2.5), where (w^k, z^k, z_0^k) is the terminal BFS of (2.3) and (w^h, z^h, z_0^h) is a homogeneous solution corresponding to (2.3) satisfying (2.6); and $z_0^k > 0$ and $z_0^h = 0$; then the LCP (q, M) is infeasible, that is, the system "$w - Mz = q$, $w \geqq 0$, $z \geqq 0$" has no feasible solution.*

Proof. As in the proof of Theorem 2.1 we assume that either (2.3) is nondegenerate or that the lexico minimum ratio rule is used throughout the algorithm to determine the dropping basic variable in each step of the algorithm. Using the hypothesis that $z_0^h = 0$ in (2.6), we have

$$w^h - Mz^h = 0$$
$$(z^h)^T w^h = 0$$

Since $(w^h, z^h, z_0^h) \neq 0$, this implies that $z^h \geq 0$. Therefore $0 = (z^h)^T w^h = (z^h)^T M z^h = 0$, and $z^h \geq 0$. So, using the hypothesis that M is an L_2-matrix, we have diagonal

matrices $\Omega \geqq 0$, $\Lambda \geqq 0$ such that $\Omega z^h \neq 0$ and $(\Lambda M + M^T\Omega)z^h = 0$. Since $\Lambda M z^h \geqq 0$ (since $Mz^h = w^h \geqq 0$ and $\Lambda \geqq 0$ is a diagonal matrix) this implies that $M^T\Omega z^h = (z^h)^T\Omega M \leqq 0$. Now $0 = (z^k)^T w^h = (z^k)^T \Lambda w^h$ (since Λ is a diagonal matrix with nonnegative entries and $w^h \geqq 0$, $z^k \geqq 0$) $= (z^k)^T \Lambda M z^h = (z^k)^T(-M^T\Omega z^h)$. So $(z^h)^T\Omega M z^k = 0$. Now

$$(z^h)^T\Omega(w^k - Mz^k - ez_0^k) = (z^h)^T\Omega q$$

Since Ω is a nonnegative diagonal matrix and $(z^h)^T w^k = 0$ and $(z^h)^T \geqq 0$, $w^k \geqq 0$, we have $(z^h)^T\Omega w^k = 0$. Also $(z^h)^T\Omega M z^k = (z^k)^T M^T\Omega z^h = -(z^k)^T\Lambda M z^h = -(z^k)^T\Lambda w^h = 0$ (since $z^k \geqq 0$, $w^h \geqq 0$, Λ is a diagonal matrix which is $\geqq 0$, $(z^k)^T w^h = 0$ implies $(z^k)^T\Lambda w^h = 0$). Using these in the above equation, we get

$$-(z^h)^T\Omega e z_0^k = (z^h)^T\Omega q$$

since $(z^h)^T \geq 0$, $\Omega \geqq 0$, $\Omega z^h \neq 0$, we have $\Omega z^h = (z^h)^T\Omega \geq 0$, this implies that $(z^h)^T\Omega e > 0$. Also, by hypothesis $z_0^k > 0$. So from the above equation $(z^h)^T\Omega q < 0$. So if $\pi = (z^h)^T\Omega$, we have

$$\begin{aligned} \pi &\geq 0 \\ -\pi M = -(z^h)^T\Omega M = -M^T\Omega z^h = \Lambda M z^h = \Lambda w^h &\geqq 0 \\ \pi q &< 0 \end{aligned}$$

which implies that $q \notin \text{Pos}(I, -M)$ by Farakas' theorem (Theorem 3 of Appendix 1). So the system

$$\begin{aligned} w - Mz &= q \\ w, z &\geqq 0 \end{aligned}$$

is itself infeasible.

□

Theorem 2.2 *The complementary pivot algorithm processes the LCP (q, M) if M is an L-matrix.*

Proof. When we apply the complementary pivot algorithm on the LCP (q, M), suppose the secondary ray $\{ (w^k + \lambda w^h, z^k + \lambda z^h, z_0^k + \lambda z_0^h) : \lambda \geqq 0 \}$ is generated. So we have

$$(w^k + \lambda w^h) - M(z^k + \lambda z^h) = q + e(z_0^k + \lambda z_0^h) \ .$$

If $z_0^h > 0$, and in the above equation if $\bar{\lambda}$ is a large positive value such that $q + e(z_0^k + \lambda z_0^h) > 0$, then $(w^k + \bar{\lambda} w^h, z^k + \bar{\lambda} z^h)$ is a complementary solution for the LCP $\left(q + e(z_0^k + \bar{\lambda} z^h), M\right)$ which by Lemma 2.1 implies that $z^k + \bar{\lambda} z^h = 0$, which means that $z^k = z^h = 0$, a contradiction to the fact that this is the secondary ray. So z_0^h cannot be > 0, that is $z_0^h = 0$, and in this case (q, M) has no solution by Lemma 2.3. So the complementary pivot algorithm processes the LCP (q, M).

□

Theorem 2.3 *If M is an $L_\star$-matrix, when the complementary pivot algorithm is applied on the LCP (q, M), it terminates with a complementary feasible solution.*

Proof. In this case we show that there can be no secondary ray. Suppose $\{ (w^k + \lambda w^h, z^k + \lambda z^h, z_0^k + \lambda z_0^h) : \lambda \geqq 0 \}$ is a secondary ray. As in the proof of Theorem 2.1, $z^h \geq 0$ (otherwise this ray will be the same as the initial ray, a contradiction). Let i be the defining index of M, z^h. So we have $z_i^h > 0$ which implies $w_i^h = 0$ by complementarity and

$$0 < (Mz^h)_i = -(ez_0^h)_i \leqq 0$$

a contradiction. So a secondary ray cannot exist in this case, and the complementary pivot method must terminate with a complementary feasible solution.

□

Theorem 2.2 and 2.3 make it possible for us to conclude that the complementary pivot algorithm processes that LCP (q, M) for a much larger class of matrices M than the copositive plus class proved in Theorem 2.1. We will now prove several results establishing that a variety of matrices are in fact L- or $L_\star$-matrices. By virtue of Theorem 2.2 and 2.3, this establishes that the complementary pivot method processes the LCP (q, M) whenever M is a matrix of one of these types.

All copositive plus matrices are L-matrices. This follows because when M is copositive plus, $y \geqq 0$ implies $y^T My \geqq 0$, and if y is such that $y \geqq 0$, $y^T My = 0$ then $(M + M^T)y = 0$, hence M satisfies the definition of being an L-matrix by taking the diagonal matrices Λ and Ω to be both I. A strictly copositive matrix is clearly an $L_\star$-matrix. From the definitions, it can be verified that PMP^T (obtained by principal rearrangement of M), $\Lambda M \Omega$ (obtained by positive row and column scaling of M) are L-matrices if M is, whenever P is a permutation matrix and Λ, Ω are diagonal matrices with positive diagonal elements. Copositive plus matrices M satisfy the property that PMP^T is also copositive plus whenever P is a permutation matrix, but if M is copositive plus, $\Lambda M \Omega$ may not be copositive when Λ, Ω are diagonal matrices with positive diagonal entries. Also if M, N are L-matrices, so is $\begin{pmatrix} M & 0 \\ 0 & N \end{pmatrix}$. Again, from Theorem 3.11 of Section 3.3, it follows that all P-matrices are $L_\star$ matrices.

Lemma 2.4 *M is row adequate iff for any y, $(y^T M_{.i})y_i \leqq 0$ for $i = 1$ to n implies that $y^T M = 0$.*

Proof. Suppose M is row adequate, and there exists a $y \geqq 0$ such that $(y^T M_{.i})y_i \leqq 0$ for $i = 1$ to n. By a standard reduction technique used in linear programming (see Section 3.4.2 in [2.26]) we can get a solution x of

$$x^T M = y^T M$$
$$x \geqq 0$$

such that $\{ M_{i.} : x_i > 0 \} \subset \{ M_{i.} : y_i > 0 \}$ and $\{ M_{i.} : x_i > 0 \}$ is linearly independent. So we also have $(x^T M_{.i})x_i \leqq 0$ for all $i = 1$ to n. Let $\mathbf{J} = \{ i : x_i > 0 \}$. Since M is a P_0-matrix, so is its principal submatrix $M_{\mathbf{JJ}} = (m_{ij} : i \in \mathbf{J}, j \in \mathbf{J})$. By linear

independence of the set of row vectors $\{ M_{i.} : i \in \mathbf{J} \}$, since M is row adequate, we know that the determinant of $M_{\mathbf{TT}} \neq 0$ for all $\mathbf{T} \subset \mathbf{J}$, and therefore that $M_{\mathbf{JJ}}$ is a P-matrix. The facts $\mathbf{J} = \{ i : x_i > 0 \}$, $x_i = 0$ if $i \notin \mathbf{J}$, and $(x^T M_{.i})x_i \leqq 0$ for all $i = 1$ to n, together imply that $M_{\mathbf{JJ}} x_{\mathbf{J}} \leqq 0$ where $x_{\mathbf{J}} = (x_j : j \in \mathbf{J})$, which implies by Theorem 3.11 of Section 3.3 that $x_{\mathbf{J}} = 0$ since $M_{\mathbf{JJ}}$ is a P-matrix, a contradiction. So $\mathbf{J}$ must be empty and $x = 0$, and hence $y^T M = 0$. Now if $y \in \mathbf{R}^n$, y not necessarily $\geqq 0$, satisfies $(y^T M_{.i})y_i \leqq 0$ for all $i = 1$ to n, let $\lambda_i = 1$ if $y_i \geqq 0$, or -1 if $y_i < 0$; and let Δ be the diagonal matrix with diagonal entries $\lambda_1, \ldots, \lambda_n$. Then $y^T \Delta \geqq 0$ and $\big(y^T \Delta(\Delta M)_{.i}\big)\lambda_i^2 y_i = \big((y^T \Delta)(\Delta M \Delta)_{.i}\big)(\lambda_i y_i) \leqq 0$ for all i. But $\Delta M \Delta$ is row adequate since M is, and by the above we therefore have $y^T \Delta(\Delta M \Delta) = 0$ or $y^T M = 0$.

Conversely, if M is a square matrix such that for any y, $(y^T M_{.i})y_i \leqq 0$ for all $i = 1$ to n implies that $y^T M = 0$, it follows that M is a P_0-matrix by the result in Exercise 3.5 and that M is row adequate.

□

Lemma 2.5 *Let M be a P_0-matrix. If*

$$\begin{aligned} My &= 0 \\ y &> 0 \end{aligned}$$

has a solution y, then the system

$$\begin{aligned} x^T M &= 0 \\ x &\geq 0 \end{aligned}$$

has a solution.

Proof. Let y satisfy $My = 0$, $y > 0$. By the result in Exercise 3.6 we know that since M is a P_0-matrix, there is a x satisfying $x^T M \geqq 0$, $x \geq 0$. If $x^T M \neq 0$, then $(x^T M)y > 0$ but $x^T(My) = 0$, a contradiction. So this x must satisfy $x^T M = 0$.

□

Theorem 2.4 *If M is row adequate, then M is an L-matrix.*

Proof. By the result in Exercise 3.5 M is a P_0-matrix iff for all $y \neq 0$, there exists an i such that $y_i \neq 0$ and $y_i(M_{i.}y) \geqq 0$. This implies that all P_0-matrices are L_1-matrices.

Suppose y satisfies $y \geq 0$, $My \geqq 0$, $y^T My = 0$. Let $\mathbf{J} = \{ i : y_i > 0 \}$. These facts imply $M_{\mathbf{JJ}} y_{\mathbf{J}} = 0$ where $M_{\mathbf{JJ}}$ is the principal submatrix $(m_{ij} : i \in \mathbf{J}, j \in \mathbf{J})$, and $y_{\mathbf{J}} = (y_j : j \in \mathbf{J}) > 0$. By Lemma 2.5, there exists an $x_{\mathbf{J}} = (x_j : j \in \mathbf{J})$ satisfying $x_{\mathbf{J}} \geq 0$, $x_{\mathbf{J}}^T M_{\mathbf{JJ}} = 0$. From Lemma 2.4, these facts imply that $x_{\mathbf{J}}^T M_{\mathbf{J}.} = 0$, where $M_{\mathbf{J}.}$ is the matrix with rows $M_{i.}$ for $i \in \mathbf{J}$. Select the diagonal matrix Ω so that $x_{\mathbf{J}} = (\Omega y)_{\mathbf{J}}$ (possible because $y_{\mathbf{J}} > 0$) and $0 = (\Omega y)_{\bar{\mathbf{J}}}$ where $\bar{\mathbf{J}} = \{ 1, \ldots, n \} \setminus \mathbf{J}$. Then $y^T \Omega M = 0$ and $(\Lambda M + M^T \Omega)y = 0$ with $\Lambda = 0$. So M is an L_2-matrix too. Thus M is an L-matrix.

□

Lemma 2.6 *If R, S are L-matrices and $P > 0$, $N < 0$ are matrices of appropriate orders, then*

$$A = \begin{pmatrix} R & P \\ N & S \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} S & N \\ P & R \end{pmatrix}$$

are also L-matrices.

Proof. Consider the product $A\xi$ where

$$\xi = \begin{pmatrix} x \\ y \end{pmatrix} \geq 0$$

Case 1 : Let $x \geq 0$, $y \geq 0$. Select a defining index for R and x, suppose it is i. Then

$$x_i(Rx + Py)_i > 0, \text{ since } P > 0 \text{ and } y \geq 0 \ .$$

This verifies that in this case the same i will serve as a defining index for A to satisfy the condition for being an L_1-matrix with this vector ξ. Also verify that in this case, A satisfies the condition for being an L_2-matrix, with this vector ξ, trivially.
Case 2 : Let $x \geq 0$, $y = 0$. The select i as in case 1 and it will serve as a defining index for A to satisfy the conditions for being an L_1-matrix, with this vector ξ. Also verify that in this case A satisfies the condition for being an L_2-matrix, with this vector ξ, trivially, since $A\xi \geqq 0$ would imply in this case $x = 0$, a contradiction.
Case 3 : Let $x = 0$, $y \geq 0$. Select a defining index for S and y, suppose it is i. Verify that the same i will serve as a defining index for A to satisfy the condition for being an L_1-matrix. If y is such that $A\xi \geqq 0$ and $\xi^T A\xi = 0$, then $Sy \geqq 0$, $y^T Sy = 0$. Since S is an L_2-matrix, there must exist diagonal matrices Λ_2, $\Omega_2 \geqq 0$ such that $\Omega_2 y \neq 0$ and $(\Lambda_2 S + S^T \Omega_2)y = 0$. Now, it can be verified easily that there is an appropriate choice of diagonal matrices Λ_1, Ω_1 such that (since $x = 0$ in this case)

$$\begin{aligned} &\begin{pmatrix} \Lambda_1 P + N^T \Omega_2 \\ \Lambda_2 S + S^T \Omega_2 \end{pmatrix} y = \\ &= \left(\begin{pmatrix} \Lambda_1 & \\ & \Lambda_2 \end{pmatrix} \begin{pmatrix} R & P \\ N & S \end{pmatrix} + \begin{pmatrix} R^T & N^T \\ P^T & S^T \end{pmatrix} \begin{pmatrix} \Omega_1 & \\ & \Omega_2 \end{pmatrix} \right) \begin{pmatrix} x \\ y \end{pmatrix} \\ &= 0 \end{aligned}$$

So A satisfies the condition for being an L_2-matrix, with this vector ξ.

These facts establish that A is an L-matrix. The proof that B is an L-matrix is similar.

□

Lemma 2.7 *If R, S are $L_\star$-matrices and $P \geqq 0$, Q arbitrary, are matrices of appropriate orders, then*

$$A = \begin{pmatrix} R & P \\ Q & S \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} S & Q \\ P & R \end{pmatrix}$$

are also $L_\star$-matrices.

Proof. Let

$$\xi = \begin{pmatrix} x \\ y \end{pmatrix}$$

Consider the product $A\xi$. If $x \geq 0$, select i to be a defining index for R and x. Since $P \geqq 0$, the same i serves as a defining index for A and ξ in the condition for A to be an $L_\star$-matrix with this ξ. If $x = 0$, then select i to be a defining index for S and y, the same i serves as a defining index for A and ξ in the condition for A to be an $L_\star$-matrix, with this ξ. So A is an $L_\star$-matrix. The proof that B is an $L_\star$-matrix is similar.

□

Lemma 2.8 *If $P > 0$ is of order $n \times m$ and $N < 0$ is of order $m \times n$, then* $\begin{pmatrix} 0 & P \\ N & 0 \end{pmatrix}$ *is an L-matrix.*

Proof. Since 0 is an L-matrix, this results follows from Lemma 2.6.

□

In Exercise 2.24 we ask the reader to prove that one formulation of the bimatrix game problem as an LCP can be solved directly by the complementary pivot algorithm, to yield a solution, using this lemma.

Lemma 2.9 *Let T $(n \times n)$, R $(n \times m)$, ρ $(n \times 1)$, S $(m \times n)$, σ $(1 \times n)$ be given matrices with $\rho > 0$, $\sigma < 0$; where $n \geqq 0$, $m \geqq 0$. If for each $x = (x_1, \ldots, x_m)^T$, δ real satisfying $(x_1, \ldots, x_m, \delta) \geq 0$, $Rx + \rho\delta \geqq 0$; there exist diagonal matrices $\Lambda \geqq 0$, $\Gamma \geqq 0$ of orders $n \times n$ and $(m+1) \times (m+1)$ respectively such that*

$$\Gamma \begin{pmatrix} x \\ \delta \end{pmatrix} \neq 0 \quad \text{and} \quad \left(\Lambda(R, \rho) + (S^T, \sigma^T)\Gamma\right) \begin{pmatrix} x \\ \delta \end{pmatrix} = 0$$

then the following matrix M is an L_2-matrix

$$M = \begin{pmatrix} T & R & \rho \\ S & 0 & 0 \\ \sigma & 0 & 0 \end{pmatrix}$$

Proof. Follows from the definition of L_2-matrices.

□

Notice that in Lemma 2.9, m could be zero, this will correspond to R, S being vacuous.

Theorem 2.5 *Let T $(n \times n)$, R $(n \times m)$, ρ $(n \times 1)$, S $(m \times n)$, σ $(1 \times n)$ be given matrices satisfying $\rho > 0$, $\sigma < 0$. Let $N = n + m + 1$. Let $\mathbf{J}_1 = \{1, \ldots, n\}$, $\mathbf{J}_2 = \{n+1, \ldots, n+m\}$, $\mathbf{J}_3 = \{n+m+1\}$. For vectors $w, z, q \in \mathbf{R}^N$, let $w_{\mathbf{J}_t}$ etc. be defined to be the vectors $w_{\mathbf{J}_t} = (w_j : j \in \mathbf{J}_t)$, etc. Assume that $q \in \mathbf{R}^N$ is a given column vector satisfying $q_{\mathbf{J}_3} = (q_{n+m+1}) > 0$. Let*

$$M = \begin{pmatrix} T & R & \rho \\ S & 0 & 0 \\ \sigma & 0 & 0 \end{pmatrix}$$

If M is an L_2-matrix, when the complementary pivot method is applied on the LCP (q, M) with the original column vector of the artificial variable z_0 taken to be $(-1, \ldots, -1, 0)^T \in \mathbf{R}^N$, either we get a complementary feasible solution of the problem, or the system

$$-\begin{pmatrix} S \\ \sigma \end{pmatrix} x \leqq \begin{pmatrix} q_{\mathbf{J}_2} \\ q_{\mathbf{J}_3} \end{pmatrix}$$

$$x \geqq 0$$

must be infeasible.

Proof. Suppose the complementary pivot algorithm is applied on the LCP (q, M) with the original column vector of the artificial variable z_0 taken to be $(-1, \ldots, -1, 0)^T \in \mathbf{R}^N$, and it terminates with the secondary ray $\left\{ (w^k + \lambda w^h, z^k + \lambda z^h, z_0^k + \lambda z_0^h) : \quad \lambda \geqq 0 \right\}$. Then

$$\begin{pmatrix} w^h_{\mathbf{J}_1} \\ w^h_{\mathbf{J}_2} \\ w^h_{\mathbf{J}_3} \end{pmatrix} - \begin{pmatrix} T & R & \rho \\ S & 0 & 0 \\ \sigma & 0 & 0 \end{pmatrix} \begin{pmatrix} z^h_{\mathbf{J}_1} \\ z^h_{\mathbf{J}_2} \\ z^h_{\mathbf{J}_3} \end{pmatrix} - \begin{pmatrix} e_n \\ e_m \\ 0 \end{pmatrix} z_0^h = 0$$

So $w^h_{\mathbf{J}_3} = \sigma z^h_{\mathbf{J}_1}$ and since $z^h_{\mathbf{J}_1} \geqq 0$, $w^h_{\mathbf{J}_3} \geqq 0$ and $\sigma < 0$, we have $z^h_{\mathbf{J}_1} = 0$, $w^h_{\mathbf{J}_3} = 0$.

If $z_0^h > 0$, then $w^h_{\mathbf{J}_2} = S z^h_{\mathbf{J}_1} + e_m z_0^h = e_m z_0^h > 0$, which by complementarity implies that $z^h_{\mathbf{J}_2} = z^k_{\mathbf{J}_2} = 0$. So $w^h_{\mathbf{J}_1} = R z^h_{\mathbf{J}_2} + \rho z^h_{\mathbf{J}_3} + e_n z_0^h = \rho z^h_{\mathbf{J}_3} + e_n z_0^h > 0$ (since $\rho > 0$). By complementarity $z^k_{\mathbf{J}_1} = 0$, and so $w^k_{\mathbf{J}_3} = \sigma z^k_{\mathbf{J}_1} + q_{\mathbf{J}_3} = q_{\mathbf{J}_3} > 0$. So by complementarity, $z^k_{\mathbf{J}_3} = z^h_{\mathbf{J}_3} = 0$. Thus $z^h = z^k = 0$, contradiction to the fact that this is a secondary ray. Therefore z_0^h must be zero. Since M is an L_2-matrix, by Lemma 2.3, the existence of this secondary ray with $z_0^h = 0$ implies that

$$w - Mz = q$$
$$w, z \geqq 0$$

has no feasible solution, which, by Faraka′s theorem (Theorem 3 of Appendix 1) implies that there exists a row vector $\alpha \in \mathbf{R}^N$ such that

$$\alpha M \leqq 0$$
$$\alpha q < 0$$
$$\alpha \geq 0$$

$\alpha M \leqq 0$ includes the constraints $\alpha_{\mathbf{J}_1} \leqq 0$ and since $\alpha_{\mathbf{J}_1} \geqq 0$, $\rho > 0$, this implies that $\alpha_{\mathbf{J}_1} = 0$. So the above system of constraints becomes

$$(\alpha_{\mathbf{J}_2}, \alpha_{\mathbf{J}_3}) \begin{pmatrix} S \\ \sigma \end{pmatrix} \leqq 0$$
$$(\alpha_{\mathbf{J}_2}, \alpha_{\mathbf{J}_3}) \begin{pmatrix} q_{\mathbf{J}_2} \\ q_{\mathbf{J}_3} \end{pmatrix} < 0$$
$$(\alpha_{\mathbf{J}_2}, \alpha_{\mathbf{J}_3}) \geqq 0$$

By Faraka′s theorem (Theorem 3 of Appendix 1) this implies that the system

$$-\begin{pmatrix} S \\ \sigma \end{pmatrix} x \leq \begin{pmatrix} q_{\mathbf{J}_2} \\ q_{\mathbf{J}_3} \end{pmatrix}$$

$$x \geqq 0$$

is infeasible.

□

In Section 2.9.2, Lemma 2.9 and Theorem 2.5 are applied to show that KKT points for general quadratic programs can be computed, when they exist, using the complementary pivot algorithm.

2.3.3 A Variant of the Complementary Pivot Algorithm

In the version of complementary pivot algorithm discussed so far, we have choosen the original column vector associated with the artificial variable z_0 to be $-e_n$. Given a column vector $d \in \mathbf{R}^n$ satisfying $d > 0$, clearly we can choose the original column vector associated with z_0 to be $-d$ instead of $-e_n$ in the complementary pivot algorithm. If this is done, the original tableau turns out to be :

w	z	z_0	
I	$-M$	$-d$	q

$$w \geqq 0, \quad z \geqq 0, \quad z_0 \geqq 0 \tag{2.7}$$

If $q \geqq 0$, $(w = q, z = 0)$ is a solution of the LCP (q, M) and we are done. So assume $q \ngeqq 0$. Determine t to satisfy $\left(\frac{q_t}{d_t}\right) =$ minimum $\left\{ \left(\frac{q_i}{d_i}\right) : i = 1 \text{ to } n \right\}$. Ties for t can be broken arbitrarily. It can be verified that if a pivot step is performed in (2.7), with the column vector of z_0 as the pivot column, and the t^{th} row as the pivot row; the right hand side constants vector becomes nonnegative after this pivot step. So $(w_1, \ldots, w_{t-1}, z_0, w_{t+1}, \ldots, w_n)$ is a feasible basic vector for (2.7). It is an almost complementary feasible basic vector as defined earlier. Choose z_t as the entering variable into this initial almost complementary feasible basic vector $(w_1, \ldots, w_{t-1}, z_0, w_{t+1}, \ldots, w_n)$, and continue by choosing entering variables using the complementary pivot rule as before.

We will now illustrate this variant of the complementary pivot algorithm using a numerical example by M. M. Kostreva [4.11].

Example 2.11

Consider the LCP (q, M), where

$$M = \begin{pmatrix} -1.5 & 2 \\ -4 & 4 \end{pmatrix} \qquad q = \begin{pmatrix} -5 \\ 17 \end{pmatrix}$$

Let $d = (5, 16)^T$. We will apply the complementary pivot algorithm on this LCP, using $-d$ as the original column of the artificial variable z_0.

Basic variables	w_1	w_2	z_1	z_2	z_0	q	
	1	0	1.5	−2	$\boxed{-5}$	−5	$t = 1$
	0	1	4	−4	−16	17	
z_0	$-\frac{1}{5}$	0	$-\frac{3}{10}$	$\frac{2}{5}$	1	1	
w_2	$-\frac{16}{5}$	1	$-\frac{8}{10}$	$\frac{12}{5}$	0	33	

The entering variable is z_1. The updated column vector of z_1 in the canonical tableau with respect to the basic vector (z_0, w_2) is nonpositive. So the algorithm ends up in ray termination.

Example 2.12

Consider the LCP (q, M) discussed in Example 2.11. Let $d = e_2 = (1, 1)^T$. We will apply the complementary pivot algorithm on this LCP with $-e_2$ as the original column of the artificial variable z_0.

Basic variables	w_1	w_2	z_1	z_2	z_0		
	1	0	1.5	−2	$\boxed{-1}$	−5	$t = 1$
	0	1	4	−4	−1	17	
z_0	−1	0	$-\frac{3}{2}$	2	1	5	
w_2	−1	1	$\boxed{\frac{5}{2}}$	−2	0	22	
z_0	$-\frac{8}{5}$	$\frac{3}{5}$	0	$\boxed{\frac{4}{5}}$	1	$\frac{91}{5}$	
z_1	$-\frac{2}{5}$	$\frac{2}{5}$	1	$-\frac{4}{5}$	0	$\frac{44}{5}$	
z_2	−2	$\frac{3}{4}$	0	1	$\frac{5}{4}$	$\frac{91}{4}$	
z_1	−2	1	1	0	1	27	

Now we have terminated with a complementary feasible basic vector, and the corresponding solution of the LCP is $w = 0$, $z = (z_1, z_2) = (27, \frac{91}{4})$.

These examples taken from M. M. Kostreva [4.11] illustrate the fact that, given a general LCP (q, M), the complementary pivot algorithm applied on it with a given

positive vector d may end up in ray termination; and yet when it is run with a different positive d vector it may terminate with a solution of the LCP. The question of how to find a good d vector seems to be a hard problem, for which no answer is known. There are LCPs which are known to have solutions, and yet when the complementary pivot algorithm is applied on them with any positive d vector, it always ends up in ray termination. See Exercise 2.11.

If M is a copositive plus matrix, and if the complementary pivot algorithm with any positive d vector ends up in ray termination when applied on the LCP (q, M), then it can be proved that the LCP (q, M) has no solution (in fact it can be proved that "$w - Mz = q$" does not even have a nonnegative solution), using arguments exactly similar to those in the proof of Theorem 2.1. Thus any LCP (q, M) where M is a copositive plus matrix, will be processed by the complementary pivot algorithm with any positive d vector.

Exercise

2.2 Prove that when M is an L-matrix or an $L_\star$-matrix, the variant of the complementary pivot algorithm discussed in this section, with any vector $d > 0$ of appropriate dimension, will process the LCP (q, M). (Proofs are similar to those in Section 2.3.2.)

2.3.4 Lexicographic Lemke Algorithm

This variant of the complementary pivot algorithm is known as the **Lexicographic Lemke Algorithm** if the original column vector of the artificial variable z_0 is taken to be $-d = -(\delta^{\mathbf{n}}, \delta^{\mathbf{n-1}}, \ldots, \delta)^T$ where δ is a sufficiently small positive number. It is not necessary to give δ a specific numerical value, but the algorithm can be executed leaving δ as a small positive parameter and remembering that $\delta^{\mathbf{r+1}} < \delta^{\mathbf{r}}$ for any nonnegative r, and that δ is smaller than any positive constant not involving δ. In this case, if D is any square matrix of order n, $Dd = D(\delta^{\mathbf{n}}, \delta^{\mathbf{n-1}}, \ldots, \delta)^T = \delta^{\mathbf{n}} D_{.1} + \delta^{\mathbf{n-1}} D_{.2} + \ldots + \delta D_{.n}$. Using this, it is possible to execute this algorithm without giving the small positive parameter δ any specific value, but using the equivalent lexicographic rules, hence the name.

2.3.5 Another Sufficient Condition for the Complementary Pivot Method to Process the LCP (q, M)

We will now discuss some results due to J. M. Evers [2.11] on another set of sufficient conditions under which the complementary pivot algorithm can be guaranteed to process the LCP (q, M). First, we discuss some lemmas. These lemmas are used later on in Theorem 2.6 to derive some conditions under which the complementary pivot algorithm can be guaranteed to solve the LCP (q, M) when M is a matrix of the form $E + N$ where E is a symmetric PSD matrix, and N is copositive.

Lemma 2.10 *Let $M = E + N$ where E is a symmetric PSD matrix and N is copositive. If the system*

$$\begin{aligned} (E+N)z &\geqq 0 \\ cz &> 0 \\ z^T(E+N)z &= 0 \\ z &\geqq 0 \end{aligned} \tag{2.8}$$

has a solution z, then the system

$$\begin{aligned} Ex - N^T y &\geqq c^T \\ y &\geqq 0 \end{aligned} \tag{2.9}$$

has no solution (x, y).

Proof. Let $\bar{z}$ be a feasible solution for (2.8). Since E is PSD and N is copositive, $\bar{z}^T(E+N)\bar{z} = 0$ implies that $\bar{z}^T E\bar{z} = \bar{z}^T N\bar{z} = 0$. Since E is symmetric, by Theorem 1.11, $\bar{z}^T E\bar{z} = 0$ implies that $E\bar{z} = 0$. So by (2.8), $N\bar{z} \geqq 0$. Let $(\bar{x}, \bar{y})$ be feasible to (2.9). So $0 \leqq \bar{y}^T N\bar{z} = -\bar{x}^T E\bar{z} + \bar{y}^T N\bar{z}$ (since $E\bar{z} = 0$) $= \bar{z}^T(-E\bar{x} + N^T\bar{y}) \leqq -c\bar{z} < 0$, a contradiction.

□

Lemma 2.11 *If the variant of the complementary pivot algorithm starting with an arbitrary positive vector d for the column of the artificial variable z_0 in the original tableau ends up in ray termination when applied on the LCP (q, M) in which M is copositive, there exists a $\bar{z}$ satisfying*

$$\begin{aligned} M\bar{z} &\geqq 0 \\ q^T\bar{z} &< 0 \\ \bar{z}^T M\bar{z} &= 0 \\ \bar{z} &\geqq 0 \end{aligned} \tag{2.10}$$

Proof. Let the terminal extreme half-line obtained in the algorithm be $\left\{ (w, z, z_0) = (w^k + \lambda w^h, z^k + \lambda z^h, z_0^k + \lambda z_0^h) : \quad \lambda \geqq 0 \right\}$ where (w^k, z^k, z_0^k) is the BFS of (2.7) and (w^h, z^h, z_0^h) is a homogeneous solution corresponding to (2.7), that is

$$\begin{aligned} w^h - Mz^h - dz_0^h &= 0 \\ w^h, z^h, z_0^h &\geqq 0 \\ (w^h, z^h, z_0^h) &\neq 0 \end{aligned} \tag{2.11}$$

and every point on the terminal extreme half-line satisfies the complementarity constraint, that is

$$(w^k + \lambda w^h)^T(z^k + \lambda z^h) = 0 \quad \text{for all } \lambda \geqq 0 \ . \tag{2.12}$$

Clearly $z^h \neq 0$ (otherwise the terminal extreme half-line is the intial one, a contradiction), so $z^h \geq 0$. By complementarity, we have $(w^h)^T z^h = 0$, from (2.11) this implies

that $(z^h)^T M z^h = -d^T z^h z_0^h \leqq 0$, (since $d > 0$, $z^h \geq 0$ implies that $d^T z^h > 0$) which implies by the copositivity of M, that $(z^h)^T M z^h = 0$ and $z_0^h = 0$. Using this in (2.11) we conclude that

$$M z^h = w^h \geqq 0 \ . \tag{2.13}$$

Since (w^k, z^k, z_0^k) is a BFS of (2.7) we have $w^k = Mz^k + dz_0^k + q$. Using this and (2.13) in (2.12) we get, for all $\lambda \geqq 0$, $(z^k + \lambda z^h)^T dz_0^k + (z^k + \lambda z^h)^T q = -(z^k + \lambda z^h)^T M (z^k + \lambda z^h) \leqq 0$ (since M is copositive and $z^k + \lambda z^h \geqq 0$). Make $\lambda > 0$, divide this inequality by λ and take the limit as λ tends to $+\infty$. This leads to

$$(z^h)^T dz_0^k + (z^h)^T q \leqq 0 \ . \tag{2.14}$$

But $z_0^k > 0$ (otherwise (w^k, z^k) will be a solution to the LCP (q, M), contradicting the hypothesis that the algorithm terminated with ray termination without leading to a solution of the LCP), $d > 0$, $z^h \geq 0$. Using these facts in (2.14) we conclude that $q^T z^h < 0$. All these facts imply that $z^h = \bar{z}$ satisfies (2.10).

□

Theorem 2.6 *Let $M = E + N$ where E is a symmetric PSD matrix and N is copositive. If the system (2.9) with $c^T = -q$ has a solution (x, y) there exists no secondary ray, and the complementary pivot algorithm terminates with a solution of the LCP (q, M).*

Proof. Follows from Lemma 2.10 and 2.11.

□

Corollary 2.1 *Putting $E = 0$ in Theorem 2.6, we conclude that if N is copositive, for every $u \geqq 0$, $v \geqq 0$ in $\mathbf{R}^n$, there exists $w, z \in \mathbf{R}^n$ satisfying*

$$\begin{aligned} Nz - w &= -N^T u - v \\ z, w &\geqq 0, \quad z^T w = 0 \ . \end{aligned}$$

2.3.6 Unboundedness of the Objective Function

Consider a mathematical program in which an objective function $f(x)$ is required to be minimized subject to constraints on the decision variables $x = (x_1, \ldots, x_n)^T$. This problem is said to be **unbounded below** if the set of feasible solutions of the problem is nonempty and $f(x)$ is not bounded below on it, that is, iff there exists an infinite sequence of feasible solutions $\{x^1, \ldots, x^r, \ldots\}$ such that $f(x^r)$ diverges to $-\infty$ as r goes to $+\infty$.

It is well known that if a linear program is unbounded below, there exists a feasible half-line (in fact an extreme half-line of the set of feasible solutions, see [2.26]) along which the objective function diverges to $-\infty$. This half-line is of the form

$\{ x^0 + \lambda x^1 : \lambda \geqq 0 \}$ satisfying the property that $x^0 + \lambda x^1$ is a feasible solution for all $\lambda \geqq 0$, and the objective value at $x^0 + \lambda x^1$ diverges to $-\infty$ as λ goes to $+\infty$. This property may not hold in general convex programming problems, that is, problems in which a convex function is required to be minimized over a closed convex set. Consider the following example due to R. Smith and K. G. Murty.

$$\begin{array}{rl} \text{Minimize} & -x_1 \\ \text{Subject to} & x_2 - x_1^{\mathbf{2}} \geqq 0 \ . \\ & x_1, x_2 \geqq 0 \end{array} \tag{2.15}$$

The set of feasible solutions of this problem is drawn in Figure 2.3.

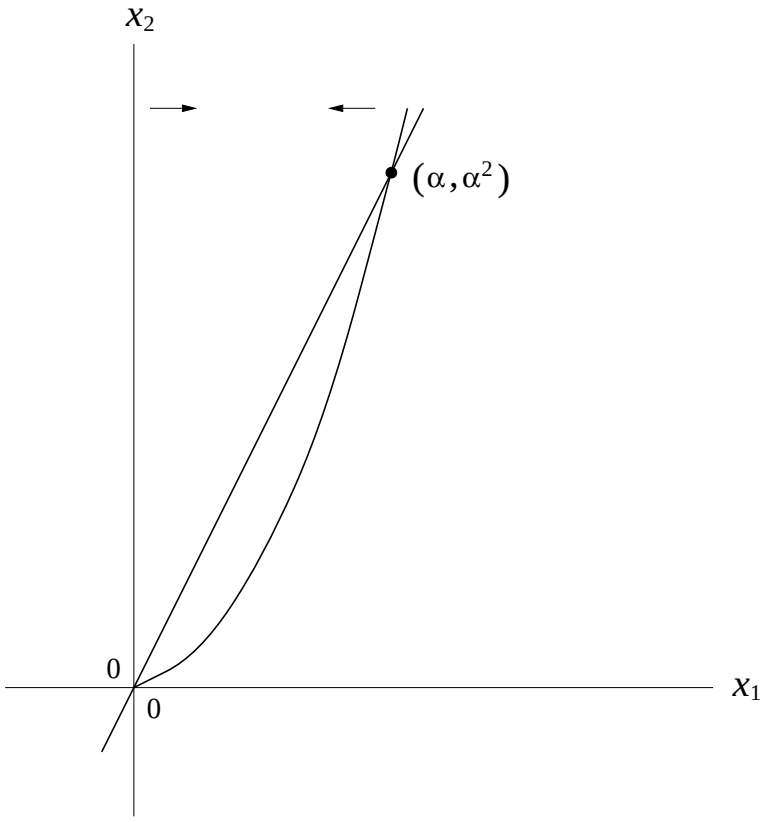


Figure 2.3 The feasible region for (2.15) is the area between the x_2 axis and the parabola. For every $\alpha > 0$, the straight line $x_2 - \alpha x_1 = 0$ intersects the parabola at exactly two points.

The equation $x_2 - x_1^{\mathbf{2}} = 0$ represents a parabola in the x_1, x_2-Cartesian plane. For every $\alpha > 0$, the straight line $x_2 - \alpha x_1 = 0$ intersects this parabola at the two points $(0, 0)$ and $(\alpha, \alpha^{\mathbf{2}})$. These facts clearly imply that even though $-x_1$ is unbounded below in (2.15), there exists no half-line in the feasible region along which $-x_1$ diverges to $-\infty$.

However, for convex quadratic programs (i.e., problems of the form (1.11) in which the matrix D is PSD) we have the following theorem.

Theorem 2.7 *Consider the quadratic program (1.11) in which D is* PSD *and symmetric. Suppose (1.11) is feasible and that $Q(x)$ is unbounded below in it. Then there exists a feasible half-line for (1.11) along which $Q(x)$ diverges to $-\infty$. Such a half-line can be constructed from the data in the terminal tableau obtained when the complementary pivot algorithm is applied to solve the corresponding LCP (1.19).*

Proof. For any positive integer r, let e_r denote the column vector in $\mathbf{R}^r$, all of whose entries are 1. By Theorem 2.1, when the complementary pivot algorithm is applied to solve (1.19) it must end in ray termination. When this happens, by the results established in the proof of Theorem 2.1, we get vectors (u^k, v^k, x^k, y^k) and (u^h, v^h, x^h, y^h) satisfying

$$\begin{pmatrix} u^k \\ v^k \end{pmatrix} - \begin{pmatrix} D & -A^T \\ A & 0 \end{pmatrix} \begin{pmatrix} x^k \\ y^k \end{pmatrix} - \begin{pmatrix} e_n \\ e_m \end{pmatrix} z_0^k = \begin{pmatrix} c^T \\ -b \end{pmatrix} \tag{2.16}$$

$$u^k, v^k, x^k, y^k \geqq 0, \quad (u^k)^T x^k = (v^k)^T y^k = 0, \quad z_0^k > 0 \ .$$

$$\begin{pmatrix} u^h \\ v^h \end{pmatrix} - \begin{pmatrix} D & -A^T \\ A & 0 \end{pmatrix} \begin{pmatrix} x^h \\ y^h \end{pmatrix} = 0 \tag{2.17}$$

$$u^h, v^h, x^h, y^h \geqq 0, \quad (u^h)^T x^h = (v^h)^T y^h = 0, \quad (x^h, y^h) \geq 0 \ .$$

$$(u^k)^T x^h = (v^k)^T y^h = (u^h)^T x^k = (v^h)^T y^k = 0 \ . \tag{2.18}$$

$$\left((x^h)^T, (y^h)^T\right) \begin{pmatrix} c^T \\ -b \end{pmatrix} < 0 \ . \tag{2.19}$$

So we have $v^h = Ax^h$ and $0 = (y^h)^T v^h = (y^h)^T A x^h$. We also have $u^h - Dx^h + A^T y^h = 0$, and hence $0 = (x^h)^T u^h = (x^h)^T D x^h - (x^h)^T A^T y^h = (x^h)^T D x^h$. Since D is PSD and symmetric by Theorem 1.11, this implies that $Dx^h = 0$. So $A^T y^h = -u^h \leqq 0$, that is $(y^h)^T A \leqq 0$. From (2.16), $-b = v^k - Ax^k - e_m z_0^k$, $z_0^k > 0$. So $(-b^T y^h) = (v^k)^T y^h - (x^k)^T A^T y^h - z_0^k e_m^T y^h = -(x^k)^T A^T y^h - z_0^k e_m^T y^h = -(x^k)^T(-u^h) - z_0^k e_m^T y^h = -z_0^k (e_m^T y^h) \leqq 0$ since $z_0^k > 0$ and $y^h \geqq 0$. So $b^T y^h = z_0^k (e_m^T y^h) \geqq 0$.

If $b^T y^h > 0$, (1.11) must be infeasible. To see this, suppose $\hat{x}$ is a feasible solution of (1.11). Then $A\hat{x} \geqq b$, $\hat{x} \geqq 0$. So $(y^h)^T A\hat{x} \geqq (y^h)^T b$. But it has been established earlier that $(y^h)^T A = -(u^h)^T \leqq 0$. Using this in the above, we have, $(y^h)^T b \leqq (y^h)^T A\hat{x} = -(u^h)^T \hat{x} \leqq 0$ (since both u^h and $\hat{x}$ are $\geqq 0$), and this contradicts the fact that $(y^h)^T b > 0$.

So, under the hypothesis that (1.11) is feasible, we must have $b^T y^h = 0$. In this case, from (2.19) we have $cx^h < 0$. From earlier facts we also have $Ax^h = v^h \geqq 0$, $x^h \geqq 0$ and $Dx^h = 0$. Let $\tilde{x}$ be any feasible solution to (1.11). These facts together imply that $\tilde{x} + \lambda x^h$ is also feasible to (1.11) for any $\lambda \geqq 0$ and $Q(\tilde{x} + \lambda x^h) = Q(\tilde{x}) + \lambda(cx^h)$ (this equation follows from the fact that $Dx^h = 0$) diverges to $-\infty$ as

λ tends to $+\infty$. Thus in this case, $\{ \tilde{x} + \lambda x^h : \quad \lambda \geqq 0 \}$ is a feasible half-line along which $Q(x)$ diverges to $-\infty$.

□

Since D is assumed to be PSD, we have $x^T Dx \geqq 0$ for all $x \in \mathbf{R}^n$. So, in this case, if $Q(x)$ is unbounded below in (1.11), the linear function cx must be unbounded below on the set of feasible solutions of (1.11), and this is exactly what happens on the half-line constructed above.

If ray termination occurs in the complementary pivot algorithm applied on (1.19) when D is PSD, we get the vectors satisfying (2.16), (2.17), (2.18) and (2.19) from the terminal tableau. If $b^T y^h > 0$, we have shown above that (1.11) must be infeasible. On the other hand, if $b^T y^h = 0$, $Q(x)$ is unbounded below in (1.11) if (1.11) is feasible. At this stage, whether (1.11) is feasible or not can be determined by using Phase I of the Simplex Method or some other algorithm to find a feasible solution of the system $Ax \geqq b$, $x \geqq 0$.

With a slight modification in the formulation of a convex quadratic program as an LCP, we can make sure that at termination of the complementary pivot algorithm applied to this LCP, if ray termination has occurred, then either a proof of infeasibility or a feasible extreme half-line along which the objective function is unbounded, are readily available, without having to do any additional work. See Section 2.9.2 for this version.

2.3.7 Some Results on Complementary BFSs

Theorem 2.8 *If the LCP (q, M) has a complementary feasible solution, then it has a complementary feasible solution which is a BFS of*

$$\begin{aligned} & w - Mz = q \\ & w \geqq 0, \quad z \geqq 0 \ . \end{aligned} \tag{2.20}$$

Proof. Let $(\bar{w}, \bar{z})$ be a complementary feasible solution for the LCP (q, M). So for each $j = 1$ to n, we have $\bar{w}_j \bar{z}_j = 0$. If $(\bar{w}, \bar{z})$ is a BFS of (2.20), we are done. Otherwise, using the algorithm discussed in Section 3.5.4 of [2.26], starting with $(\bar{w}, \bar{z})$, we can obtain a BFS $(\hat{w}, \hat{z})$ of (2.20) satisfying the property that the set of variables which have positive values in $(\hat{w}, \hat{z})$, is a subset of the set of variables which have positive values in $(\bar{w}, \bar{z})$. So $\hat{w}_j \hat{z}_j = 0$, for $j = 1$ to n. Hence $(\hat{w}, \hat{z})$ is a complementary feasible solution of the LCP (q, M) and it is also a BFS of (2.20).

□

Note 2.1 The above theorem does not guarantee that whenever the LCP (q, M) has a complementary feasible solution, there exists a complementary feasible basis for (2.20). See Exercise 1.10.

Theorem 2.9 *Suppose M is nondegenerate. If $(\bar{w}, \bar{z})$ is a complementary feasible solution for the LCP (q, M), the set of column vectors $\{ I_{.j} : j \text{ such that } \bar{w}_j > 0 \} \cup \{ -M_{.j} : j \text{ such that } \bar{z}_j > 0 \}$ is linearly independent. Also, in this case, define a vector of variables $y = (y_1, \ldots, y_n)$ by*

$$y_j = \begin{cases} w_j \ , & \text{if } \bar{w}_j > 0 \\ z_j \ , & \text{if } \bar{z}_j > 0 \\ \text{either } w_j \text{ or } z_j \text{ choosen arbitrarily} \ , & \text{if both } \bar{w}_j \text{ and } \bar{z}_j \text{ are } 0 \ . \end{cases}$$

Then y is a complementary feasible basic vector for (2.20).

Proof. From Corollary 3.1 of Chapter 3, when M is nondegenerate, every complementary vector is basic. Since $(\bar{w}, \bar{z})$ is a complementary feasible solution, this implies that the set $\{ I_{.j} : j \text{ such that } \bar{w}_j > 0 \} \cup \{ -M_{.j} : j \text{ such that } \bar{z}_j > 0 \}$ is linearly independent. Also from this result, y is a complementary basic vector, and the BFS of (2.20) with y, as the basic vector is $(\bar{w}, \bar{z})$, and hence y is a complementary feasible basic vector.

□

Theorem 2.10 *If M is* PSD *or copositive plus, and (2.20) is feasible, then there exists a complementary feasible basic vector for (2.20).*

Proof. When the complementary pivot algorithm is applied to solve the LCP (q, M), it terminates with a complementary feasible basic vector when M is copositive plus and (2.20) is feasible, by Theorem 2.1.

□

2.4 A METHOD OF CARRYING OUT THE COMPLEMENTARY PIVOT ALGORITHM WITHOUT INTRODUCING ANY ARTIFICIAL VARIABLES, UNDER CERTAIN CONDITIONS

Consider the LCP (q, M) of order n, suppose the matrix M satisfies the condition :

> there exists a column vector of M in which all the entries are strictly positive. (2.21)

Then a variant of the complementary pivot algorithm which uses no artificial variable at all, can be applied on the LCP (q, M). We discuss it here. The original tableau for

this version of the algorithm is :

w	z	
I	$-M$	q

$$w \geqq 0, \quad z \geqq 0 \tag{2.22}$$

As before, we assume that $q \not\geqq 0$. Let s be such that $M_{.s} > 0$. So the column vector associated with z_s is strictly negative in (2.22). Hence the variable z_s can be made to play the same role as that of the artificial variable z_0 in versions of the complementary pivot algorithm discussed earlier, and thus there is no need to introduce the artificial variable. Determine t to satisfy $\left(\frac{q_t}{m_{ts}}\right) =$ minimum $\left\{ \left(\frac{q_i}{m_{is}}\right) : i = 1 \text{ to } n \right\}$. Ties for t can be broken arbitrarily. When a pivot step is carried out in (2.22) with the column of z_s as the pivot column and row t as the pivot row, the right hand side constants vector becomes nonnegative after this pivot step (this follows because $-m_{is} < 0$ for all i and by the choice of t). Hence, $(w_1, \ldots, w_{t-1}, z_s, w_{t+1}, \ldots, w_n)$ is a feasible basic vector for (2.22), and if $s = t$, it is a complementary feasible basic vector and the solution corresponding to it is a solution of the LCP (q, M), terminate. If $s \neq t$, the feasible basic vector $(w_1, \ldots, w_{t-1}, z_s, w_{t+1}, \ldots, w_n)$ for (2.22) satisfies the following properties :

i) It contains exactly one basic variable from the complementary pair (w_i, z_i) for $n - 2$ values of i (namely $i \neq s$, t here).

ii) It contains both the variables from a fixed complementary pair (namely (w_s, z_s) here), as basic variables.

iii) There exists exactly one complementary pair both the variables in which are not contained in this basic vector (namely (w_t, z_t) here).

The complementary pair of variables identified by property (iii), both of which are not contained in the basic vector, is known as the **left out complementary pair of variables** in the present basic vector.

For carrying out this version of the complementary pivot algorithm, any feasible basic vector for (2.22) satisfying (i), (ii), (iii) is known as an **almost complementary feasible basic vector**. All the basic vectors obtained during this version of the algorithm, with the possible exception of the terminal one (which may be a complementary basic vector), will be such almost complementary feasible basic vectors, and the complementary pair in property (ii) both of whose variables are basic, will be the same for all of them.

In the canonical tableau of (2.22) with respect to the initial almost complementary feasible basic vector, the updated column vector of w_t can be verified to be strictly negative (because the pivot column in the original tableau, $-M_{.s}$, is strictly negative). Hence if w_t is selected as the entering variable into the initial basic vector, an almost complementary extreme half-line is generated. Hence the initial almost complementary BFS of (2.22) is at the end of an almost complementary ray.

The algorithm chooses z_t as the entering variable into the initial almost complementary feasible basic vector $(w_1, \ldots, w_{t-1}, z_s, w_{t+1}, \ldots, w_n)$. In all subsequent steps,

the entering variable is uniquely determined by the complementary pivot rule, that is, the entering variable in a step is the complement of the dropping variable in the previous step. The algorithm can terminate in two possible ways :

1. At some stage one of the variables form the complementary pair (w_s, z_s) (this is the pair specified in property (ii) of the almost complementary feasible basic vectors obtained during the algorithm) drops out of the basic vector, or becomes equal to zero in the BFS of (2.22). The BFS of (2.22) at that stage is a solution of the LCP (q, M).
2. At some stage of the algorithm both the variables in the complementary pair (w_s, z_s) may be strictly positive in the BFS, and the pivot column in that stage may turn out to be nonpositive, and in this case the algorithm terminates with another almost complementary ray. This is **ray termination**.

When ray termination occurs, the algorithm bas been unable to solve the LCP (q, M).

Example 2.13

Consider the LCP (q, M), where

$$M = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} \qquad q = \begin{pmatrix} -4 \\ -5 \\ -1 \end{pmatrix}$$

All the column vectors of M are strictly positive here. We will illustrate the algorithm on this problem using $s = 3$.

Original Tableau

w_1	w_2	w_3	z_1	z_2	z_3	q
1	0	0	−2	−1	−1	−4
0	1	0	−1	−2	**[−1]**	−5
0	0	1	−1	−1	−2	−1

$-M_{.3} < 0$. The minimum $\{ \frac{-4}{1}, \frac{-5}{1}, \frac{-1}{2} \} = -5$, and hence $t = 2$ here. So the pivot row is row 2, and the pivot element for the pivot operation to get the initial almost complementary feasible basic vector is inside a box in the original tableau. Applying the algorithm we get the following canonical tableaus:

Basic variables	w_1	w_2	w_3	z_1	z_2	z_3	q	Ratios
w_1	1	-1	0	-1	$\boxed{1}$	0	1	$\frac{1}{1}$ *Min.*
z_3	0	-1	0	1	2	1	5	$\frac{5}{2}$
w_3	0	-2	1	1	3	0	9	$\frac{9}{3}$
z_2	1	-1	0	-1	1	0	1	
z_3	-2	1	0	$\boxed{3}$	0	1	3	$\frac{3}{3}$ *Min.*
w_3	-3	1	1	4	0	0	6	$\frac{6}{4}$
z_2	$\frac{1}{3}$	$-\frac{2}{3}$	0	0	1	$\frac{1}{3}$	2	
z_1	$-\frac{2}{3}$	$\frac{1}{3}$	0	1	0	$\frac{1}{3}$	1	
w_3	$-\frac{4}{3}$	$-\frac{1}{3}$	1	0	0	$-\frac{4}{3}$	2	

So the solution of this LCP is $w = (w_1, w_2, w_3) = (0, 0, 2)$; $z = (z_1, z_2, z_3) = (1, 2, 0)$.

Exercise

2.3 Show that the version of the complementary pivot algorithm discussed in this section can be used to process all LCPs (q, M) in which M is copositive plus and at least one of its columns is strictly positive. In this case, prove that ray termination cannot occur, and that the algorithm will terminate with a complementary feasible basic vector for the problem.

2.5 TO FIND AN EQUILIBRIUM PAIR OF STRATEGIES FOR A BIMATRIX GAME USING THE COMPLEMENTARY PIVOT ALGORITHM

The LCP corresponding to the problem of finding an equilibrium pair of strategies in a bimatrix game is (1.42), where A, B^T are positive matrices. The original tableau for this problem is :

u	v	ξ	η	
I_m	0	0	$-A$	$-e_m$
0	I_N	$-B^T$	0	$-e_N$

$$u \geqq 0, \quad v \geqq 0, \quad \xi \geqq 0, \quad \eta \geqq 0 \tag{2.23}$$

where for any r, I_r denotes the identity matrix of order r. The complementary pairs of variables in this problem are (u_i, ξ_i), $i = 1$ to m, and (v_j, η_j), $j = 1$ to N.

We leave it to the reader to verify that when the complementary pivot algorithm discussed in Section 2.2 is applied on this problem, it ends up in ray termination right after obtaining the initial almost complementary feasible basic vector. However, it turns out that the variant of the complementary pivot algorithm discussed in Section 2.4 can be applied to this problem, and when it is applied it works. We discuss the application of this version of the algorithm here.

So, here, an **almost complementary feasible basic vector** for (2.23), is defined to be a feasible basic vector that contains exactly one basic variable from each complementary pair excepting two pairs. Both variables of one of these pairs are basic variables, and both variables in the other pair are nonbasic variables. These are the conditions for almost complementarity (i), (ii), (iii), discussed in Section 2.4.

The column vectors of the variables ξ_i, η_j, in (2.23) are all nonpositive, but none of them is strictly negative. But, because of their special structure, an almost complementary feasible basic vector for (2.23) can be constructed by the following special procedure.

Initially make the variable ξ_1 a basic variable and the variables $\xi_2, \ldots, \xi_m$ nonbasic variables. Make ξ_1 equal to ξ_1^0, the smallest positive number such that $v^0 = -e_N + (B^T)_{.1}\xi_1^0 \geq 0$. At least one of the components in v^0, say, v_r^0 is zero. Make v_r a nonbasic variable too. The complement of v_r is η_r. Make the value of η_r to be the smallest positive value, η_r^0, such that $u^0 = A_{.r}\eta_r^0 - e_m \geq 0$. At least one of the components in u^0, say u_s^0 is 0. If $s = 1$, the basic vector $(u_2, \ldots, u_m, v_1, \ldots, v_{r-1}, v_{r+1}, \ldots, v_N, \xi_1, \eta_r)$ is a complementary feasible basic vector, and the feasible solution corresponding to it is a solution of the LCP (1.42), terminate.

If $s \neq 1$, the basic vector, $(u_1, \ldots, u_{s-1}, u_{s+1}, \ldots, u_m, v_1, \ldots, v_{r-1}, v_{r+1}, \ldots, v_N, \xi_1, \eta_r)$ is a feasible basic vector. Both the variables in the complementary pair (u_1, ξ_1) are basic variables in it. Both variables in the complementary pair (u_s, ξ_s) are nonbasic variables. And this basic vector contains exactly one basic variable from every complementary pair in (2.23), excepting (u_1, ξ_1), (u_s, ξ_s). Hence this initial basic vector is an almost complementary feasible basic vector. All the basic vectors obtained during the algorithm (excepting the terminal complementary feasible basic vector) will be almost complementary feasible basic vectors containing both the variables in the pair (u_1, ξ_1) as basic variables.

When u_s is made as the entering variable into the initial basic vector, an almost complementary extreme half-line is generated. Hence the BFS of (2.23) with respect

to the initial basic vector is an almost complementary BFS at the end of an almost complementary extreme half-line.

The algorithm begins by taking ξ_s as the entering variable into the initial basic vector. In all subsequent steps, the entering variable is picked by the complementary pivot rule. The algorithm terminates when one of the variables in the pair (u_1, ξ_1) drops from the basic vector. It can be proved that temination occurs after at most a finite number of pivots. The terminal basis is a complementary feasible basis. In this algorithm if degeneracy is encountered, its should be resolved using the lexico minimum ratio rule (see Section 2.2.8).

Example 2.14

We will solve the LCP (1.43) corresponding to the 2 person game in Example 1.9. In tableau form it is

u_1	u_2	v_1	v_2	v_3	ξ_1	ξ_2	η_1	η_2	η_3	q
1	0	0	0	0	0	0	−2	−2	−1	−1
0	1	0	0	0	0	0	$\boxed{-1}$	−2	−2	−1
0	0	1	0	0	$\boxed{-1}$	−2	0	0	0	−1
0	0	0	1	0	−3	−1	0	0	0	−1
0	0	0	0	1	−2	−3	0	0	0	−1

$u, v, \xi, \eta \geqq 0$ and $u_1\xi_1 = u_2\xi_2 = v_1\eta_1 = v_2\eta_2 = v_3\eta_3 = 0$

Making $\xi_2 = 0$, the smallest value of ξ_1 that will yield nonnegative values to the v's is 1. When $\xi_2 = 0$, $\xi_1 = 1$ the value of v_1 is 0. Hence, v_1 will be made a nonbasic variable. The complement of v_1 is η_1. So make η_2 and η_3 nonbasic variables. The smallest value of η_1 that will make the u's nonnegative is $\eta_1 = 1$. When $\eta_1 = 1$ with $\eta_2 = \eta_3 = 0$, u_2 becomes equal to 0. So make u_2 a nonbasic variable. The canonical tableau with respect to the initial basic vector is therefore obtained as below by performing pivots in the columns of ξ_1 and η_1, with the elements inside a box as pivot elements.

Basic variables	u_1	u_2	v_1	v_2	v_3	ξ_1	ξ_2	η_1	η_2	η_3	q	Ratios
u_1	1	−2	0	0	0	0	0	0	2	3	1	
η_1	0	−1	0	0	0	0	0	1	2	2	1	
ξ_1	0	0	−1	0	0	1	2	0	0	0	1	$\frac{1}{2}$
v_2	0	0	−3	1	0	0	$\boxed{5}$	0	0	0	2	$\frac{2}{5}$ *Min.*
v_3	0	0	−2	0	1	0	1	0	0	0	1	$\frac{1}{1}$

The algorithm continues by selecting ξ_2, the complement of u_2, as the entering variable. v_2 drops from the basic vector.

Basic variables	u_1	u_2	v_1	v_2	v_3	ξ_1	ξ_2	η_1	η_2	η_3	q
u_1	1	-2	0	0	0	0	0	0	$\boxed{2}$	3	1
η_1	0	-1	0	0	0	0	0	1	2	2	1
ξ_1	0	0	$\frac{1}{5}$	$-\frac{2}{5}$	0	1	0	0	0	0	$\frac{1}{5}$
ξ_2	0	0	$-\frac{3}{5}$	$\frac{1}{5}$	0	0	1	0	0	0	$\frac{2}{5}$
v_3	0	0	$-\frac{7}{5}$	$-\frac{1}{5}$	1	0	0	0	0	0	$\frac{3}{5}$

Since v_2 has dropped from the basic vector, its complement η_2 is the next entering variable. There is a tie in the minimum ratio when η_2 is the entering variable, since it can replace either u_1 or η_1 from the basic vector. Such ties should be resolved by the lexico minimum ratio test, but in this case we will let u_1 drop from the basic vector, since that leads to a complementary feasible basis to the problem.

Basic variables	u_1	u_2	v_1	v_2	v_3	ξ_1	ξ_2	η_1	η_2	η_3	q
η_2	$\frac{1}{2}$	-1	0	0	0	0	0	0	1	$\frac{3}{2}$	$\frac{1}{2}$
η_1	-1	1	0	0	0	0	0	1	0	-1	0
ξ_1	0	0	$\frac{1}{5}$	$-\frac{2}{5}$	0	1	0	0	0	0	$\frac{1}{5}$
ξ_2	0	0	$-\frac{3}{5}$	$\frac{1}{5}$	0	0	1	0	0	0	$\frac{2}{5}$
v_3	0	0	$-\frac{7}{5}$	$-\frac{1}{5}$	1	0	0	0	0	0	$\frac{3}{5}$

The present basic vector is a complementary feasible basic vector. The solution $(u_1, u_2; v_1, v_2, v_3; \xi_1, \xi_2; \eta_1, \eta_2, \eta_3) = (0, 0; 0, 0, \frac{3}{5}; \frac{1}{5}, \frac{2}{5}; 0, \frac{1}{2}, 0)$ is a solution of the LCP. In this solution $\xi_1 + \xi_2 = \frac{3}{5}$ and $\eta_1 + \eta_2 + \eta_3 = \frac{1}{2}$. Hence the probability vector $x = \frac{\xi}{(\sum \xi_i)} = \left(\frac{1}{3}, \frac{2}{3}\right)^T$ and $y = \frac{\eta}{(\sum \eta_j)} = (0, 1, 0)^T$ constitute an equilibrium pair of strategies for this game.

Theorem 2.11 *If the lexicographic minimum ratio rule is used to determine the dropping variable in each pivot step (this is to prevent cycling under degeneracy) of the complementary pivot algorithm discussed above for solving (1.42), it terminates in a finite number of pivot steps with a complementary feasible solution.*

Proof. The original tableau for this problem is (2.23), in which $A > 0$, $B^T > 0$, by the manner in which the problem is formulated. In the algorithm discussed above

for this problem, both variables from exactly one complementary pair are nonbasic in every almost complementary feasible basic vector obtained, and this pair is known as the **left out complementary pair of variables**. The left out complementary pair may be different in the various almost complementary feasible basic vectors obtained during the algorithm, but the complementary pair both of whose variables are basic, remains the same in all of them.

Let $(u_1, \ldots, u_{s-1}, u_{s+1}, \ldots, u_m, v_1, \ldots, v_{r-1}, v_{r+1}, \ldots, v_N, \xi_1, \eta_r)$ be the initial almost complementary feasible basic vector obtained in the algorithm, by the special procedure discussed above. Let the initial tableau be the canonical tableau of (2.23) with respect to the initial almost complementary feasible basic vector. In this, the left out complementary pair is (u_s, ξ_s) both of which are nonbasic at present. Let

$$u^1 = (u^1_i), \quad u^1_i = -1 + \left(\frac{a_{ir}}{a_{sr}}\right), \quad \text{for } i \neq s, \quad u^1_s = 0 \ .$$

$$v^1 = (v^1_j), \quad v^1_j = -1 + \left(\frac{b_{1j}}{b_{1r}}\right), \quad \text{for } j \neq r, \quad v^1_r = 0 \ .$$

$$\xi^1 = \left(\frac{1}{b_{1r}}, 0, \ldots, 0\right)$$

$$\eta^1 = (\eta^1_j), \quad \eta^1_j = 0, \quad \text{for } j \neq r, \quad \eta^1_r = \frac{1}{a_{sr}} \ .$$

$$\bar{u}^h = (\bar{u}^h_i), \quad \bar{u}^h_i = \left(\frac{a_{is}}{a_{ir}}\right), \quad \text{for } i \neq s, \quad \bar{u}^h_s = 1 \ .$$

$$\bar{v}^h = 0, \quad \bar{\xi}^h = 0$$

$$\bar{\eta}^h = (\bar{\eta}^h_j), \quad \bar{\eta}^h_j = 0, \quad \text{for } j \neq r, \quad \bar{\eta}^h_r = \left(\frac{1}{a_{sr}}\right) \ .$$

The present BFS can be verified to be $(u^1, v^1, \xi^1, \eta^1)$. It can also be verified that $(\bar{u}^h, \bar{v}^h, \bar{\xi}^h, \bar{\eta}^h)$ is a homogeneous solution corresponding to the initial tableau, and that the initial almost complementary extreme half-line generated when u_s is brought into the basic vector in the initial tableau is $\left\{ (u^1, v^1, \xi^1, \eta^1) + \lambda(\bar{u}^h, \bar{v}^h, \bar{\xi}^h, \bar{\eta}^h) : \lambda \geqq 0 \right\}$.

The algorithm begins by bringing the nonbasic variable ξ_s into the basic vector in the initial tableau, and continues by using the complementary pivot rule to choose the entering variable and the lexico-minimum ratio rule to choose the dropping variable in each step.

Let B be the basis consisting of the columns of the basic variables in the initial tableau (not the original tableau), in a step of this procedure and let $\beta = (\beta_{ij}) = B^{-1}$. Let $\bar{q}$ be the updated right hand side constants vector in this step. If u_1 or ξ_1, is eligible to be a dropping variable in this step by the usual minimum ratio test, it is choosen as the dropping variable, and the pivot step is carried out, leading to a complementary feasible basic vector for the problem. If both u_1 and ξ_1 are ineligible to be dropping variables in this step, the lexico minimum ratio rule chooses the dropping variable so that the pivot row corresponds to the row which is the lexico minimum $\left\{ \frac{(\bar{q}_i, \beta_i.)}{p_{it}} : i \text{ such that } p_{it} > 0 \right\}$ where $p = (p_{1t}, \ldots, p_{m+N,t})^T$ is the pivot column (updated column of the entering variable) in this step. This lexico minimum ratio rule determines the dropping variable uniquely and unambiguously in each pivot step.

In each almost complementary feasible basic vector, obtained during the algorithm, there is exactly one left out complementary pair of variables; and hence it can have at most two adjacent almost complementary feasible basic vectors, that can be obtained by bringing one variable from the left out complementary pair into it.

The left out complementary pair in the initial almost complementary feasible basic vector is (u_s, ξ_s), and when u_s is brought into the initial almost complementary feasible basic vector, we obtain the initial almost complementary extreme half-line. So the only manner in which the almost complementary path can be continued from the initial almost complementary BFS is by bringing ξ_s into the basic vector. The updated column of ξ_s in the initial tableau can be verified to contain at least one positive entry. Hence when ξ_s is brought into the initial basic vector, we get an adjacent almost complementary feasible basic vector, and the almost complementary path continues uniquely and unambiguously from there. Each almost complementary feasible basic vector has at most two adjacent ones, from one of them we arrive at this basic vector; we move to the other when we leave this basic vector. These facts, and the perturbation interpretation of the lexico minimum ratio rule imply that an almost complementary feasible basic vector obtained in the algorithm can never reappear later on. Since there are at most a finite number of almost complementary feasible basic vectors, the algorithm must terminate in a finite number of pivot steps. If it terminates by obtaining a complementary feasible basic vector, the BFS corresponding to it is a solution of the LCP (1.42) and we are done. The only other possibility in which the algorithm can terminate is if the updated column vector of the entering variable in some step has no positive entries in it, in which case we get a terminal almost complementary extreme half-line (this is the **ray termination** discussed earlier). We will now show that this second possibility (ray termination) cannot occur in this algorithm.

Suppose ray termination occurs in pivot step k. Let the almost complementary BFS in this step be $(u^k, v^k, \xi^k, \eta^k)$ and let the terminal extreme half-line be: $\{ (u^k, v^k, \xi^k, \eta^k) + \lambda(u^h, v^h, \xi^h, \eta^h) : \lambda \geqq 0 \}$. From this and from the almost compelementary property being maintained in the algorithm, we have :

$$\begin{pmatrix} u^k + \lambda u^h \\ v^k + \lambda v^h \end{pmatrix} - \begin{pmatrix} 0 & A \\ B^T & 0 \end{pmatrix} \begin{pmatrix} \xi^k + \lambda \xi^h \\ \eta^k + \lambda \eta^h \end{pmatrix} = \begin{pmatrix} -e_m \\ -e_N \end{pmatrix} \tag{2.24}$$

$$(u_i^k + \lambda u_i^h)(\xi_i^k + \lambda \xi_i^h) = 0 \quad \text{for all } i \neq 1 \tag{2.25}$$

$$(v_j^k + \lambda v_j^h)(\eta_j^k + \lambda \eta_j^h) = 0 \quad \text{for all } j \tag{2.26}$$

$$u^k, v^k, \xi^k, \eta^k, u^h, v^h, \xi^h, \eta^h \geqq 0 \tag{2.27}$$

for all $\lambda \geqq 0$. $(u^h, v^h, \xi^h, \eta^h)$ is a homogeneous solution satisfying the nonnegativity restrictions and

$$\begin{pmatrix} u^h \\ v^h \end{pmatrix} - \begin{pmatrix} 0 & A \\ B^T & 0 \end{pmatrix} \begin{pmatrix} \xi^h \\ \eta^h \end{pmatrix} = 0 \ .$$

That is :

$$\begin{aligned} u^h &= A\eta^h \\ v^h &= B^T \xi^h \ . \end{aligned} \tag{2.28}$$

We also have $(u^h, v^h, \xi^h, \eta^h) \neq 0$, which implies by (2.27) and (2.28) that $(\xi^h, \eta^h) \geq 0$. Now suppose $\xi^h \neq 0$. So $\xi^h \geq 0$. Since $B > 0$, this implies by (2.28) that $v^h = B^T \xi^h > 0$. From (2.26) this implies that $\eta_j^k + \lambda \eta_j^h = 0$ for all j and for all $\lambda \geqq 0$. From (2.24) this implies that $(u^k + \lambda u^h) = -e_m < 0$, a contradiction.

Suppose we have $\xi^h = 0$ but $\eta^h \neq 0$. So $\eta^h \geq 0$ and since $A > 0$, $u^h = A\eta^h > 0$.So from (2.25), we must have $\xi_i^k = 0$ for all $i \neq 1$. Since $\xi^h = 0$, by (2.28) $v^h = 0$. So from (2.24)

$$v^k = -e_N + B^T \xi^k \tag{2.29}$$

and since $\xi_i^k = 0$ for all $i \neq 1$, v^k is obtained by the same procedure as v^1, the value of v in the initial BFS (since ξ_1^k must be the smallest value that makes v^k nonnegative in (2.29) in order to get an extreme point solution). So v^k is the same as v^1 in the initial BFS in (2.23). By our discussion earlier, this implies that $v_j^k > 0$ for all $j \neq r$, and $v_r^k = 0$. By (2.26) this implies that $\eta_j^k + \lambda \eta_j^h = 0$ for all $\lambda \geqq 0$ and $j \neq r$. These facts clearly imply that $(u^k, v^k, \xi^k, \eta^k)$ is the same as the initial BFS obtained for (2.23). This is a contradiction, since a BFS obtained in a step of the algorithm cannot reappear later on, along the almost complementary path.

These facts imply that ray termination cannot occur. So the algorithm must terminate in a finite number of steps by obtaining a complementary feasible basic vector, and the terminal BFS is therefore a solution of the LCP (1.42).

□

Comments 2.1 The complementary pivot algorithm for computing equilibrium strategies in bimatrix games is due to C. E. Lemke and J. T. Howson [1.18]. C. E. Lemke [2.21] extended this into the complementary pivot algorithm for LCPs discussed in Section 2.2. The proof of Theorem 2.1 is from the paper of R. W. Cottle and G. B. Dantzig [1.3] which also discusses various applications of the LCP and some principal pivoting methods for solving it. C. E. Lemke was awarded the ORSA/TIMS John Von Neumann Theory Prize in 1978 for his contributions to this area. The citation of the award says "Nash's equilibrium proofs were nonconstrutive, and for many years it seemed that the nonlinearity of the problem would prevent the actual numerical solution of any but the simplest noncooperative games. The breakthrough came in 1964 with an ingenious algorithm for the bimatrix case devised by Carlton Lemke and L. T. Howson Jr. It provided both a constructive existence proof and a practical means of calculation. The underlying logic, involving motions on the edges of an appropriate polyhedron, was simple and elegant yet conceptually daring in an epoch when such motions were typically contemplated in the context of linear programming. Lemke took the lead in exploiting the many ramifications and applications of this procedure, which range from the very basic linear complementary problem of mathematical programming to the problem of calculating fixed points of continuous, nonlinear mappings arising in various contexts. A new chapter in the theory and practice of mathematical programming was thereby opened which quickly became a very active and well-populated area of research...".

The geometric interpretation of the LCP using complementary cones was initiated in K. G. Murty [3.47, 1.26].

2.6 A VARIABLE DIMENSION ALGORITHM

We consider the LCP (q, M) which is to find $w, z \in \mathbf{R}^n$ satisfying

$$w - Mz = q \tag{2.30}$$

$$w, z \geqq 0 \tag{2.31}$$

$$\text{and} \quad w^T z = 0 \tag{2.32}$$

Definition : Principal Subproblem
Let $\mathbf{J} \subset \{1, \ldots, n\}$. *Denote* $w_{\mathbf{J}} = (w_j : j \in \mathbf{J})$, $z_{\mathbf{J}} = (z_j : j \in \mathbf{J})$, $q_{\mathbf{J}} = (q_j : j \in \mathbf{J})$, *and the principal submatrix of* M *corresponding to* $\mathbf{J}$, $M_{\mathbf{JJ}} = (m_{ij} : i, j \in \mathbf{J})$. *The* **principal subproblem** *of the LCP (2.30)–(2.32) in the variables* $w_{\mathbf{J}}$, $z_{\mathbf{J}}$ *(or the principal subproblem of the LCP (2.30)–(2.32) associated with the subset* $\mathbf{J}$*) is the LCP* $(q_{\mathbf{J}}, M_{\mathbf{JJ}})$ *of order* $|\mathbf{J}|$, *the complementary pairs of variables in it are* $\{w_j, z_j\}$ *for* $j \in \mathbf{J}$ *and it is: find* $w_{\mathbf{J}}$, $z_{\mathbf{J}}$ *satisfying*

$$\begin{aligned} &w_{\mathbf{J}} - M_{\mathbf{JJ}} z_{\mathbf{J}} = q_{\mathbf{J}} \\ &w_{\mathbf{J}}, z_{\mathbf{J}} \geqq 0 \\ &w_{\mathbf{J}}^T z_{\mathbf{J}} = 0 \ . \end{aligned}$$

This principal subproblem is therefore obtained from (2.30)–(2.32) by striking off the columns of all the variables w_j, z_j *for* $j \notin \mathbf{J}$ *and the equation in (2.30) corresponding to* $j \notin \mathbf{J}$.

Let $\mathbf{J} = \{1, \ldots, n\} \setminus \{i\}$, $\omega = (w_1, \ldots, w_{i-1}, w_{i+1}, \ldots, w_n)^T$, $\xi = (z_1, \ldots, z_{i-1}, z_{i+1}, \ldots, z_n)^T$. The following results follow by direct verification.

Results 2.2 If $(\hat{w} = (\hat{w}_1, \ldots, \hat{w}_n)^T; \hat{z} = (\hat{z}_1, \ldots, \hat{z}_n)^T)$ is a solution of the LCP (q, M) and $\hat{z}_i = 0$, then $(\hat{\omega} = (\hat{w}_1, \ldots, \hat{w}_{i-1}, \hat{w}_{i+1}, \ldots, \hat{w}_n)^T; \hat{\xi} = (\hat{z}_1, \ldots, \hat{z}_{i-1}, \hat{z}_{i+1}, \ldots, \hat{z}_n)^T)$ is a solution of its principal subproblem in the variables ω, ξ.

Results 2.3 Suppose that $(\tilde{\omega} = (\tilde{w}_1, \ldots, \tilde{w}_{i-1}, \tilde{w}_{i+1}, \ldots, \tilde{w}_n)^T$; $\tilde{\xi} = (\tilde{z}_1, \ldots, \tilde{z}_{i-1}, \tilde{z}_{i+1}, \ldots, \tilde{z}_n)^T)$ is a solution of the principal subproblem of the LCP (q, M) in the variables ω, ξ. Define $\tilde{z}_i = 0$ and let $\tilde{z} = (\tilde{z}_1, \ldots, \tilde{z}_{i-1}, \tilde{z}_i, \tilde{z}_{i+1}, \ldots, \tilde{z}_n)^T$. If $q_i + M_i.\tilde{z} \geqq 0$, define $\tilde{w}_i = q_i + M_i.\tilde{z}$, and let $\tilde{w} = (\tilde{w}_1, \ldots, \tilde{w}_{i-1}, \tilde{w}_i, \tilde{w}_{i+1}, \ldots, \tilde{w}_n)^T$, then $(\tilde{w}, \tilde{z})$ is a solution of the original LCP (q, M).

Example 2.15

Consider the following LCP (q, M)

w_1	w_2	w_3	z_1	z_2	z_3	q
1	0	0	2	0	−3	4
0	1	0	−1	−4	−3	−14
0	0	1	1	2	−2	13

$w_j \geqq 0$, $z_j \geqq 0$, $w_j z_j = 0$ for all $j = 1$ to 3

Let $\omega = (w_1, w_2)^T$, $\xi = (z_1, z_2)^T$. Then the principal subproblem of this LCP in the variable ω, ξ is

w_1	w_2	z_1	z_2	γ
1	0	2	0	4
0	1	−1	−4	−14

$w_j \geqq 0$, $z_j \geqq 0$, $w_j z_j = 0$ for $j = 1, 2$

$(\hat{w} = (0, 0, 5)^T; \hat{z} = (2, 3, 0)^T)$ is a solution of the original LCP and $\hat{z}_3$ is equal to zero in this solution. This implies that $(\hat{\omega} = (0, 0); \hat{\xi} = (2, 3))$ is a solution of this principal subproblem which can easily be verified. Also, $(\tilde{\omega} = (4, 0)^T; \tilde{\xi} = (0, \frac{14}{4})^T)$ is another solution of the principal subproblem. Defining $\tilde{z}_3 = 0$, $\tilde{z} = (0, \frac{14}{4}, 0)^T$, we verify that $q_3 + M_{3.}\tilde{z} = 13 + (-1, -2, 2)(0, \frac{14}{4}, 0)^T = 6 > 0$. Hence, define $\tilde{w}_3 = 6$, and $\tilde{w} = (4, 0, 6)^T$. It can be verified that $(\tilde{w} = (4, 0, 6)^T; \tilde{z} = (0, \frac{14}{4}, 0)^T)$ is another solution of the original LCP.

We now discuss a variable dimension algorithm for the LCP (q, M) due to L. Van der Heyden [2.38]. If $q \geqq 0$, $(w = q, z = 0)$ is a readily available solution. So we assume that $q \ngeqq 0$. The method proceeds by solving a sequence of principal subproblems of (2.30), (2.31), (2.32) always associated with subsets of the form $J = \{1, \ldots, k\}$ (this problem is called the k-problem), for some k satisfying $1 \leqq k \leqq n$. When the method is working on the k-problem, the bottom $n-k$ constraints in (2.30) as well as the columns of variables w_j, z_j for $j > k$ can be ignored, hence the reason for the name. All the intermediate solutions for (2.30) obtained during the method (with the exception of the terminal solution which is a complementary feasible solution satisfying (2.30), (2.31), (2.32)) are of two types called **position 1** and **position 2 solutions** defined below.

Position 1 Solution : This is a solution $(\tilde{w}, \tilde{z})$ for (2.30) satisfying the following properties :

i) there exists an index k such that $\tilde{z}_k = 0$ and $\tilde{w}_k < 0$.

ii) $\tilde{z}_j = 0$ for $j > k$.

iii) if $k > 1$, $\tilde{w}^{(k-1)} = (\tilde{w}_1, \ldots, \tilde{w}_{k-1})$, $\tilde{z}^{(k-1)} = (\tilde{z}_1, \ldots, \tilde{z}_{k-1})$ is a solution for the principal subproblem of (2.30), (2.31), (2.32) determined by the subset $\{1, \ldots, k-1\}$, that is, $\tilde{w}^{(k-1)} \geqq 0$, $\tilde{z}^{(k-1)} \geqq 0$ and $(\tilde{w}^{(k-1)})^T \tilde{z}^{(k-1)} = 0$.

From the definition, a position 1 solution $(\bar{w}, \bar{z})$ always satisfies $\bar{w}^T \bar{z} = 0$, it is complementary (but infeasible) and it will be a complementary basic solution associated with a complementary (but infeasible in the same sense that the solution violates (2.31)) basic vector for (2.30).

Position 2 Solution : This is a solution $(\hat{w}, \hat{z})$ for (2.30) satisfying the following properties :

a) there exists an index k such that $\hat{z}_k > 0$, $\hat{w}_k < 0$.

b) $\hat{z}_j = 0$ for $j > k$.

c) there is a $u < k$ such that both $\hat{z}_u$ and $\hat{w}_u$ are zero.

d) $\hat{w}^{(k-1)} = (\hat{w}_1, \ldots, \hat{w}_{k-1})^T \geqq 0$, $\hat{z}^{(k-1)} = (\hat{z}_1, \ldots, \hat{z}_{k-1})^T \geqq 0$ and $(\hat{w}^{(k-1)})^T \hat{z}^{(k-1)} = 0$.

From the definition, a position 2 solution discussed above is an almost complementary solution (not feasible, since some of the variables are < 0) of the type discussed in Section 2.4, it satisfies $\hat{w}^T \hat{z} = \hat{w}_k \hat{z}_k$. It will be an almost complementary basic solution associated with an almost complementary basic vector for (2.30) which has both w_k, z_k as basic variables, and contains exactly one basic variable from the complementary pair (w_j, z_j) for each $j \neq k$ or u (both variables w_u, z_u are out of this almost complementary basic vector, so the complementary pair (w_u, z_u) is the left out complementary pair in this basic vector). This almost complementary basic vector has w_j as a basic variable for all $j > k$. All intermediate (i. e., except the initial and terminal) solutions obtained by the method when it is working on the k-problem wiil be position 2 solutions of (2.30) as defined above.

Note 2.1 As mentioned above, all the solutions obtained during the algorithm will be basic solutions of (2.30).The definitions given above for positions 1, 2 solutions are under the assumption that q is nondegenerate in the LCP (q, M) (i. e., that every solution to (2.30) has at least n nonzero variables). In the general case when q may be degenerate, the algorithm perturbs q by adding the vector $(\varepsilon, \varepsilon^{\mathbf{2}}, \ldots, \varepsilon^{\mathbf{n}})^T$ to it, where ε is treated as a sufficiently small positive number without giving any specific value to it (see Section 2.1, 2.2.2, 2.2.8), and all the inequalities for the signs of the variables should be understood in the usual lexico sense.

The Algorithm

The algorithm takes a path among basic vectors for (2.30) using pivot steps. All basic vectors obtained will be almost complementary basic vectors as defined in Section 2.4, or complementary basic vectors.

Initial Step: STEP 0 : The initial basic vector is $w = (w_1, \ldots, w_n)$. The initial solution is the Position 1 basic solution of (2.30) corresponding to it, define $k =$ minimum $\{i : q_i < 0\}$. Begin with the k-problem, by making a type 1 pivot step to increase the value of the nonbasic variable z_k from 0, as described below.

STEP 1 : Type 1 Pivot Step, to increase the Value of a Nonbasic Variable from Zero. Let $(y_1, \ldots, y_k, w_{k+1}, \ldots, w_n)$ be the basic vector in some stage of working for the k-problem. If this is the initial basic vector, $(y_1, \ldots, y_k)$ will be a complementary basic vector for the principal subproblem of (2.30), (2.31), (2.32) defined by the subset $\{1, \ldots, k\}$. Except possibly at termination of work on the k-problem, y_k will always be w_k; $y_1, \ldots, y_{k-1}$ will all be w_j or z_j for $j \leqq k-1$. This type of pivot step occurs when the value of a nonbasic variable, say v, selected by the rules specified in the algorithm, is to be increased from its present value of zero. The variable v will be either w_j or z_j for some $j \leqq k$. Let the canonical tableau for (2.30) with respect to the present basic vector be

Tableau 2.7 Canonical Tableau

$y_1 \ldots y_k \quad w_{k+1} \ldots w_n$	$\ldots v \ldots$	
	a_1	$\bar{q}_1$
I	$\ldots \vdots \ldots$	$\vdots$
	a_n	$\bar{q}_n$

While working on the k-problem, in all the canonical tableaus, we will have $\bar{q}_1, \ldots, \bar{q}_{k-1} \geqq 0$ and $\bar{q}_k < 0$ (and $y_k = w_k$). Let $\beta = B^{-1}$ be the inverse of the present basis. The algorithm always maintains $(\bar{q}_i, \beta_i.) \succ 0$ for $i = 1$ to $k-1$. Let λ denote the nonnegative value given to the nonbasic variable v. The new solution as a function of λ is

$$\begin{aligned} &\text{all nonbasic variables other than } v \text{ are } 0 \\ &v = \lambda \\ &y_i = \bar{q}_i - \lambda a_i, \quad i = 1 \text{ to } k \\ &w_j = \bar{q}_j - \lambda a_j, \quad j = k+1 \text{ to } n \end{aligned} \tag{2.33}$$

We will increase the value of λ from 0 until one of the variables y_i for $i = 1$ to k, changes its value from its present to zero in (2.33), and will change sign if λ increases any further. This will not happen if the updated column of the entering variable v satisfies

$$a_i \leqq 0, \quad i = 1, \ldots, k-1 \quad \text{and} \quad a_k \geqq 0 \tag{2.34}$$

If condition (2.34) is satisfied, the method is unable to proceed further and termination occurs with the conclusion that the method is unable to process this LCP. If condition (2.34) is not satisfied, define

$$\theta = \text{Max}\Big\{\frac{\bar{q}_i}{a_i} : \text{ Over } 1 \leqq i \leqq k-1 \text{ such that } a_i > 0 \text{ ; and } \frac{\bar{q}_k}{a_k}, \text{ if } a_k < 0\Big\} \tag{2.35}$$

Let $\mathbf{\Delta}$ be the set of all i between 1 to k which tie for the maximum in (2.35). If $\mathbf{\Delta}$ is a singleton set, let r be the element in it. Otherwise let r be the element which attains the lexicomaximum in lexicomaximum $\left\{ \left(\frac{(\bar{q}_i, \beta_i.)}{a_i} \right) : i \in \mathbf{\Delta} \right\}$. If $r = k$, v replaces $y_k (= w_k)$ from the basic vector. After this pivot step we are lead to the basic vector $(y_1, \ldots, y_{k-1}, v, w_{k+1}, \ldots, w_n)$ which will be a complementary basic vector for (2.30) (except that the variables $y_1, \ldots, y_{k-1}, v$ may have to be rearranged so that the j^{th} variable here is from the j^{th} complementary pair), and $(y_1, \ldots, y_{k-1}, v)$ is a complementary lexico feasible basic vector for the k-problem (except for the rearrangement of the basic variables as mentioned above). If $(y_1, \ldots, y_{k+1}, v, w_{k+1}, \ldots, w_n)$ is feasible to (2.30) (this happens if the updated right hand side constants vector is $\geqq 0$ after the pivot step of replacing y_k by v), it is a complementary feasible basic vector for (2.30), the method terminates with the basic solution corresponding to it as being a solution for (2.30), (2.31), (2.32). On the other hand, if $(y_1, \ldots, y_{k-1}, v, w_{k+1}, \ldots, w_n)$ is not a feasible basic vector for (2.30), the k-problem has just been solved and the method moves to another principal subproblem with index greater than k (this is called a **forward move**), go to Step 3.

If $r < k$, v replaces y_r from the basic vector, leading to the new basic vector $(y_1, \ldots, y_{r-1}, v, y_{r+1}, \ldots, y_k, w_{k+1}, \ldots, w_n)$. Two things can happen now. If $y_r = z_k$, then this new basic vector is a complementary basic vector for (2.30) (except for rearrangement of the variables as mentioned above), but $(y_1, \ldots, y_{r-1}, v, y_{r+1}, \ldots, y_k)$ is not lexico feasible for the k-problem. In this case the method moves to make a type 2 pivot step (discussed next) leading to a principal subproblem with index less than k (this is called a **regressive move**, moving to a smaller principal subproblem already solved earlier). The next steps of the algorithm will be concerned with finding yet another solution for this smaller principal subproblem. Go to Step 2.

The second possibility is that $y_r \neq z_k$. In this case the basic vector $(y_1, \ldots, y_{r-1}, v, y_{r+1}, \ldots, y_k, w_{k+1}, \ldots, w_n)$ is another almost complementary basic vector, the basic solution of (2.30) associated with which is another position 2 solution. In this case, the method continues the work on the k-problem by making a type 1 pivot step next, to increase the value of the complement of y_r from zero.

STEP 2 : Type 2 Pivot Step to Decrease the Value of a Nonbasic Variable w_g from Zero. This pivot step will be made whenever we obtain a complementary basic vector $(y_1, \ldots, y_k, w_{k+1}, \ldots, w_n)$ after doing some work on the k-problem, with $y_k = w_k$. Let Tableau 2.7 be the canonical tableau with respect to this complementary basic vector. We will have $\bar{q}_i \geqq 0$, $i = 1$ to $k-1$ and $(\bar{q}_k, \beta_k.) \prec 0$ at this stage ($\beta_k.$ is the k^{th} row of the present basis inverse). Let g be the maximum j such that $y_j = z_j$. Now the algorithm decreases the value of the nonbasic variable w_g from zero. Letting $v = w_g$, and giving this variable a value λ (we want to make $\lambda \leqq 0$), the new solution obtained is of the same form as in (2.33). We will decrease the value of λ from 0 until one of the variables y_i for $i = 1$ to g, changes its value from its present to zero in (2.33), and will change sign if λ decreases any further. This will not happen if the updated column of the entering variable v satisfies

$$a_i \geqq 0, \quad i = 1 \text{ to } g \tag{2.36}$$

in which case termination occurs with the conclusion that the method is unable to process this LCP. If (2.36) is not satisfied, define

$$\theta = \text{ Minimum } \left\{ -\left(\frac{\bar{q}_i}{a_i}\right) : 1 \leqq i \leqq g, \quad i \text{ such that } a_i < 0 \right\} . \tag{2.37}$$

Let $\mathbf{\Delta}$ be the set of all i between 1 to g which tie for the minimum in (2.37). If $\mathbf{\Delta}$ is a singleton set let r be the element in it. Otherwise let r be the element which attains the lexicominimum in lexico minimum $\left\{ -\frac{(\bar{q}_i, \beta_i .)}{a_i} : i \in \mathbf{\Delta} \right\}$. Replace y_r in the present basic vector by v ($= w_g$ here) and move over to the g-problem after this pivot step, by going to Step 1 to increase the value of the complement of y_r from 0.

STEP 3 : We move to this step when we have solved a k-problem after performing a type 1 pivot step on it in Step 1. Let $(y_1, \ldots, y_k, w_{k+1}, \ldots, w_n)$ be the complementary basic vector at this stage with $y_j \in \{w_j, z_j\}$ for $j = 1$ to k. Let $\bar{q} = (\bar{q}_1, \ldots, \bar{q}_n)^T$ be the current updated right hand side constants vector. Since $(y_1, \ldots, y_k)$ is a complementary feasible basic vector for the k-problem, we have $\bar{q}_i \geqq 0$ for $i = 1$ to k. If $\bar{q}_i \geqq 0$ for $i = k+1$ to n also, this basic vector is complementary feasible to the original problem (2.30), (2.31), (2.32), and we would have terminated. So $\bar{q}_i < 0$ for at least one i between $k+1$ to n. Let u be the smallest i for which $\bar{q}_i < 0$, replace k by u and go back to Step 1 to increase the value of z_k from zero.

Numerical Example 2.16

We provide here a numerical example for this algorithm from the paper [2.38] of L. Van der Heyden. Consider the LCP (q, M) where

$$q = \begin{pmatrix} -1 \\ -2 \\ -10 \end{pmatrix} \quad , \qquad M = \begin{pmatrix} 1 & 1 & 1 \\ 3 & 1 & 1 \\ 2 & 2 & 1 \end{pmatrix}$$

Since $q_1 < 0$, the algorithm begins with $k = 1$, on the 1-problem. Pivot elements are inside a box.

Basic Vector	w_1	w_2	w_3	z_1	z_2	z_3		
w_1	1	0	0	$\boxed{-1}$	-1	-1	-1	$k = 1$. Increase z_1. In this
w_2	0	1	0	-3	-1	-1	-2	type 1 pivot step, w_1
w_3	0	0	1	-2	-2	-1	-10	drops from basic vector.
z_1	-1	0	0	1	1	1	1	$k = 3$.
w_2	-3	1	0	0	2	$\boxed{2}$	1	Increase z_3.
w_3	-2	0	1	0	0	1	-8	w_2 drops.

Basic Vector	w_1	w_2	w_3	z_1	z_2	z_3		
z_1	$\frac{1}{2}$	$-\frac{1}{2}$	0	1	0	0	$\frac{1}{2}$	$k=3$. Increase z_2
z_3	$-\frac{3}{2}$	$\frac{1}{2}$	0	0	$\boxed{1}$	1	$\frac{1}{2}$	(complement of w_2).
w_3	$-\frac{1}{2}$	$-\frac{1}{2}$	1	0	-1	0	$-\frac{17}{2}$	z_3 drops.
z_1	$\frac{1}{2}$	$\boxed{-\frac{1}{2}}$	0	1	0	0	$\frac{1}{2}$	Need a type 2 pivot step.
z_2	$-\frac{3}{2}$	$\frac{1}{2}$	0	0	1	1	$\frac{1}{2}$	Decrease w_2.
w_3	-2	0	1	0	0	1	-8	z_1 drops.
w_2	$\boxed{-1}$	1	0	-2	0	0	-1	$k=1$. Increase w_1 (compl.
z_2	-1	0	0	1	1	1	1	of z_1 that just dropped)
w_3	-2	0	1	0	0	1	-8	w_2 drops.
w_1	1	-1	0	2	0	0	1	$k=3$.
z_2	0	-1	0	3	1	$\boxed{1}$	2	Increase z_3.
w_3	0	-2	1	4	0	1	-6	z_2 drops.
w_1	1	-1	0	2	0	0	1	Increase w_2 (complement
z_3	0	-1	0	3	1	1	2	of z_2 that just dropped).
w_3	0	$\boxed{-1}$	1	1	-1	0	-8	w_3 drops.
w_1	1	0	-1	1	1	1	9	Complementary
z_3	0	0	-1	2	2	1	10	feasible
w_2	0	1	-1	-1	1	0	8	basic vector

Thus $(w_1, w_2, w_3; z_1, z_2, z_3) = (9, 8, 0; 0, 0, 10)$ is a complementary feasible solution of this problem.

Conditions Under Which the Algorithm is Guaranteed to Work

Theorem 2.12 *For every* $\mathbf{J} \subset \{1, \ldots, n\}$, *if the principal submatrix* $M_{\mathbf{JJ}}$ *of* M *associated with* $\mathbf{J}$ *satisfies the property that there exists no positive vector* $z_{\mathbf{J}}$ *such that the last component of* $M_{\mathbf{JJ}} z_{\mathbf{J}}$ *is nonpositive and the other components are zero, the termination criteria (2.34) or (2.36) will never be satisfied and the algorithm terminates*

with a complementary feasible basic vector for the LCP (q, M) after a finite number of steps.

Proof. When (2.34) or (2.36) is satisfied, we have a solution of the type given in equation (2.33), which we denote by $(w(\lambda), z(\lambda)) = (\bar{w} + \lambda w^h, \bar{z} + \lambda z^h)$ satisfying the property that for $\lambda > 0$, there exists a k such that $w_k(\lambda) < 0$, $z_k(\lambda) > 0$, $z_j(\lambda) = 0$ for $j > k$, and if $k > 1$ the vectors $w^{(k-1)}(\lambda) = (w_1(\lambda), \ldots, w_{k-1}(\lambda))$, $z^{(k-1)}(\lambda) = (z_1(\lambda), \ldots, z_{k-1}(\lambda))$ are nonnegative and complementary. Let $\mathbf{J} = \{1, \ldots, k\}$, $w^h_{\mathbf{J}} = (w^h_1, \ldots, w^h_k)^T$, $z^h_{\mathbf{J}} = (z^h_1, \ldots, z^h_k)^T$. Then

$$\begin{aligned} w^h_{\mathbf{J}} - M_{\mathbf{JJ}} z^h_{\mathbf{J}} &= 0 \\ z^h_{\mathbf{J}} &\geqq 0 \\ w^h_k &\leqq 0 \end{aligned} \tag{2.38}$$

and if $k > 1$, $(w^h_1, \ldots, w^h_{k-1}) \geqq 0$, and $w^h_j z^h_j = 0$ for $j = 1$ to $k - 1$. Let $\mathbf{P} = \{ j : 1 \leqq j \leqq k$, and $z^h_j > 0 \}$. Clearly $\mathbf{P} \neq \emptyset$, otherwise $(w^h_{\mathbf{J}}, z^h_{\mathbf{J}}) = 0$. Letting $z^h_{\mathbf{P}} = (z^h_j : j \in \mathbf{P})$, all the components of $M_{\mathbf{PP}} z^h_{\mathbf{P}}$ are zero except possibly the last one because of (2.38) and the fact that $w^h_j z^h_j = 0$ for $j = 1$ to $k - 1$. Also, the last component of $M_{\mathbf{PP}} z^h_{\mathbf{P}}$ is $\leqq 0$ because of (2.38). And since $z^h_{\mathbf{P}} > 0$, this contradicts the hypothesis in the theorem.

The finiteness of the algorithm follows from the path argument used in Sections 2.2, 2.3, the argument says that the algorithm never returns to a previous position as this situation implies the existence of a position with three adjacent positions, a contradiction. Since there are only a finite number of positions we must terminate with a solution for the original LCP.

□

Corollary 2.2 *If M has the property that for every $\mathbf{J} \subset \{1, \ldots, n\}$, the corresponding submatrix $M_{\mathbf{JJ}}$ of M satisfies the property that the system*

$$\begin{aligned} M_{\mathbf{JJ}} z_{\mathbf{J}} &\leqq 0 \\ z_{\mathbf{J}} &\geqq 0 \end{aligned}$$

has the unique solution $z_{\mathbf{J}} = 0$, then the variable dimension algorithm discussed above will terminate with a solution of the LCP (q, M) for any $q \in \mathbf{R}^n$.

Proof. Follows from Theorem 2.12.

□

R. W. Cottle [3.9] has shown that the class of matrices M satisfying the hypothesis in Theorem 2.12 or Corollary 2.2, is the strictly semi-monotone matrices defined later on in Section 3.4, which is the same as $\bar{Q}$ (completely Q-matrices, that is, matrices all of whose principal submatrices are Q-matrices). This class includes all P-matrices and positive or strictly copositive matrices.

By the results discussed in Chapter 3, the LCP (q, M) has a unique solution when M is a P-matrix. So if M is s P-matrix and the LCP (q, M) is solved by the variable dimension algorithm, type 2 pivot steps will never have to be performed.

M. J. Todd [2.35, 2.36] has shown that when q is nondegenerate in (2.30) and M is a P-matrix, the variable dimension algorithm discussed above corresponds to the lexicographic Lemke algorithm discussed in Section 2.3.4.

Now consider the LCP (q, M) of order n. Let e_n denote the column vector of all 1's in $\mathbf{R}^n$. Introduce the artificial variable z_0 associated with the column vector $-e_n$, as in the complementary pivot algorithm (see equation (2.3)). Introduce an additional artificial variable w_0, which is the complement of z_0, and the artificial constraint "$w_0 - e_n^T z = q_0$", where q_0 is treated as a large positive number, without giving it any specific value. This leads to an LCP of order $n+1$, in which the variables are $(w_0, w_1, \ldots, w_n)$, $(z_0, z_1, \ldots, z_n)$ and the data is

$$M^\star = \begin{pmatrix} 0 & -e_n^T \\ e_n & M \end{pmatrix} \quad , q^\star = \begin{pmatrix} q_0 \\ q \end{pmatrix} \quad .$$

Since q_0 is considered as a large positive parameter, $w_0 > 0$ and $z_0 = 0$ in any complementary solution of this larger dimensional LCP $(q^\star, M^\star)$, and hence if $((\bar{w}_0, \bar{w}), (\bar{z}_0, \bar{z}))$ is a solution of this LCP, then $(\bar{w}, \bar{z})$ is a solution of the original LCP (q, M).

Essentially by combining the arguments in Theorems 2.1 and 2.12, L. Van der Heyden [2.39] has shown that if M is a copositive plus matrix and the system "$w - Mz = q$, $w \geqq 0$, $z \geqq 0$" has a feasible solution, when the variable dimension algorithm is applied on the LCP $(q^\star, M^\star)$, it will terminate with a complementary feasible solution $((\bar{w}_0, \bar{w}), (\bar{z}_0, \bar{z}))$ in a finite number of steps. This shows that the variable dimension algorithm will process LCP's associated with copositive plus matrices, by introducing an artificial dimension and by applying the variable dimension algorithm to the enlarged LCP.

2.7 EXTENSIONS TO FIXED POINT COMPUTING, PIECEWISE LINEAR AND SIMPLICIAL METHODS

It has also been established that the arguments used in the complementary pivot algorithm can be generalized, and these generalizations have led to algorithms that can compute approximate Brouwer and Kakutani fixed points! Until now, the greatest single contribuition of the complementarity problem is probably the insight that it has provided for the development of fixed point computing algorithms. In mathematics, fixed point theory is very higly developed, but the absence of efficient algorithms for computing these fixed points has so far frustrated all attempts to apply this rich theory to real life problems. With the development of these new algorithms, fixed point theory is finding numerous applications in mathematical programming, in mathematical economics, and in various other areas. We present one of these fixed point computing

algorithms, and some of its applications, in this section. We show that the problem of computing a KKT point for an NLP can be posed as a fixed point problem and solved by these methods.

The algorithms that are discussed later in this section trace a path through the simplices of a triangulation in $\mathbf{R}^n$, that is why they are called **simplicial methods**. Since they use piecewise linear approximations of maps, these methods are also called **piecewise linear methods**. Since the path traced by these methods has exactly the same features as that of the complementary pivot algorithm (see Sections 2.2.5, 2.2.6) these methods are also called **complementary pivot methods**.

2.7.1 Some Definitions

Let $g(x)$ be a real valued function defined over a convex subset $\bar{\Gamma} \subset \mathbf{R}^n$. We assume that the reader is familiar with the definition of continuity of $g(x)$ at a point $x^0 \in \bar{\Gamma}$, and the definition of the vector of partial derivatives of $g(x)$ at x^0, $\nabla g(x^0) = \left(\frac{\partial g(x^0)}{\partial x_1}, \ldots, \frac{\partial g(x^0)}{\partial x_n}\right)$, when it exists. The function $g(x)$ is said to be differentiable at x^0 if $\nabla g(x^0)$ exists, and for any $y \in \mathbf{R}^n$, $\frac{1}{\alpha}\left(g(x^0 + \alpha y) - g(x^0) - \alpha\left(\nabla g(x^0)\right)y\right)$ tends in the limit to zero as α tends to zero. If $g(x)$ is differentiable at x^0, for any $y \in \mathbf{R}^n$, we can approximate $g(x^0 + \alpha y)$ by $g(x^0) + \alpha\left(\nabla g(x^0)\right)y$ for values of α for which $|\alpha|$ is small. This is the **first order Taylor series expansion** for $g(x + \alpha y)$ at $x = x^0$. If $g(x)$ is differentiable at x^0, the partial derivative vector $\nabla g(x^0)$ is known as the **gradient vector** of $g(x)$ at x^0.

When the second order partial derivatives of $g(x)$ exist at x^0, we denote the $n \times n$ matrix of second order partial derivatives $\left(\frac{\partial^2 g(x^0)}{\partial x_i \partial x_j}\right)$ by the symbol $H(g(x^0))$. It is called the **Hessian matrix** of $g(x)$ at x^0.

Let $g_1(x), \ldots, g_m(x)$ be m real valued convex functions defined on the convex subset $\bar{\Gamma} \subset \mathbf{R}^n$. For each $x \in \bar{\Gamma}$, define $s(x) = \text{Maximum}\{ g_1(x), \ldots, g_m(x) \}$. The function $s(x)$ is known as the **pointwise supremum** or **maximum** of $\{g_1(x), \ldots, g_m(x)\}$. It is also convex on $\bar{\Gamma}$. See Figure 2.4 where we illustrate the pointwise supremum of several affine functions defined on the real line.

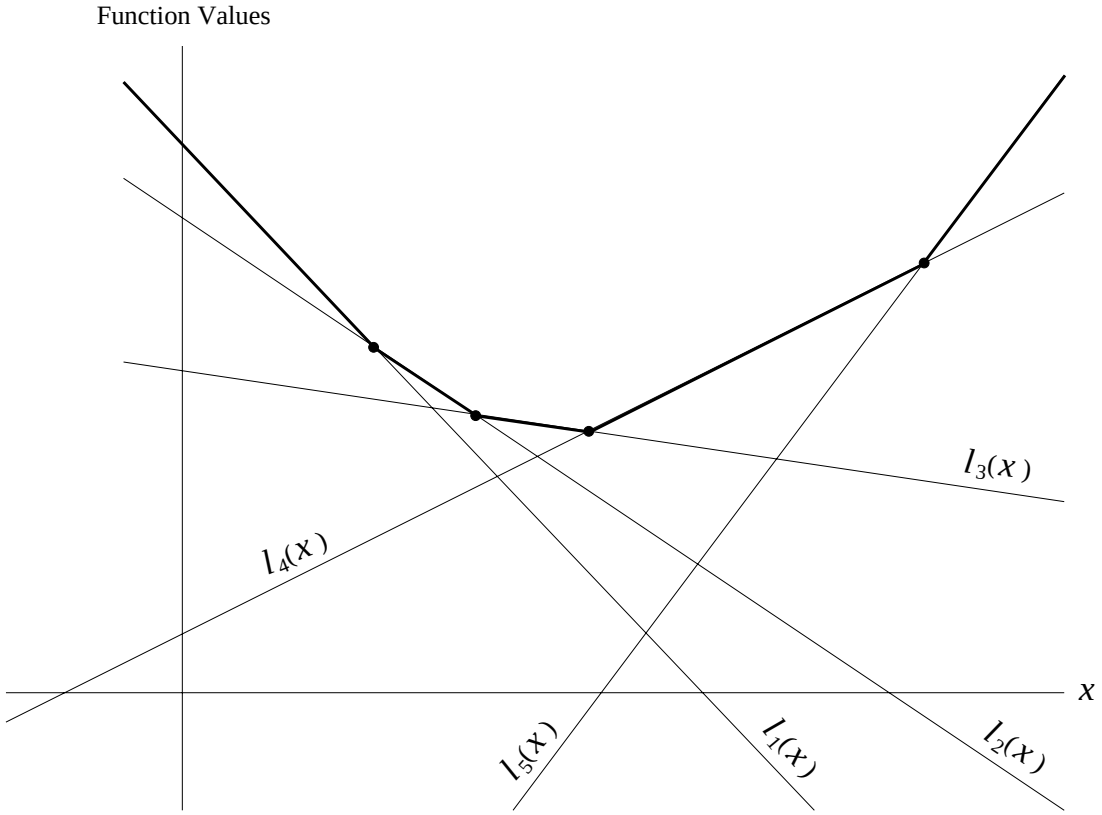


Figure 2.4 $l_1(x)$ to $l_5(x)$ are five affine functions defined on $\mathbf{R}^1$. Function values are plotted on the vertical axis. Their pointwise maximum is the function marked with thick lines here.

Subgradients and Subdifferentials of Convex Functions

Let $g(x)$ be a real valued convex function defined on $\mathbf{R}^n$. Let $x^0 \in \mathbf{R}^n$ be a point where $g(x^0)$ is finite. The vector $d = (d_1, \ldots, d_n)^T$ is said to be a subgradient of $g(x)$ at x^0 if

$$g(x) \geqq g(x^0) + d^T(x - x^0), \quad \text{for all } x \in \mathbf{R}^n \ . \tag{2.39}$$

Notice that the right hand side of (2.39) is $l(x) = (g(x^0) - d^t x^0) + d^T x$, is an affine function in x; and we have $g(x^0) = l(x^0)$. One can verify that $l(x)$ is the first order Taylor expansion for $g(x)$ around x^0, constructed using the vector d in place of the gradient vector of $g(x)$ at x^0. So d is a subgradient of $g(x)$ at x^0, iff this modified Taylor approximation is always an underestimate for $g(x)$ at every x.

Example 2.17

Let $x \in \mathbf{R}^1$, $g(x) = x^2$. $g(x)$ is convex. Consider the point $x^0 = 1$, $d = 2$. It can be verified that the inequality (2.39) holds in this case. So $d = (2)$ is a subgradient for $g(x)$ at $x^0 = 1$ in this case. The affine function $l(x)$ on the right hand side of (2.39) in this case is $1 + 2(x - 1) = 2x - 1$. See Figures 2.5, 2.6 where the inequality (2.39) is illustrated.

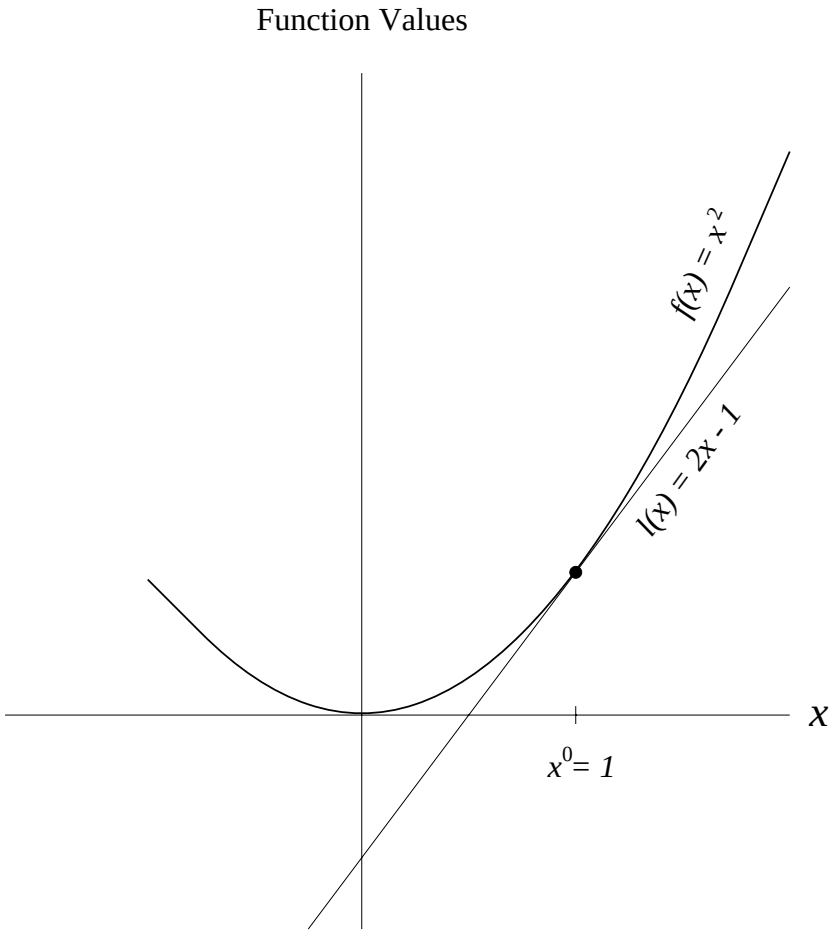


Figure 2.5 A Convex Function, and the Affine Function Below it Constructed Using a Subgradient for it at the Point x^0.

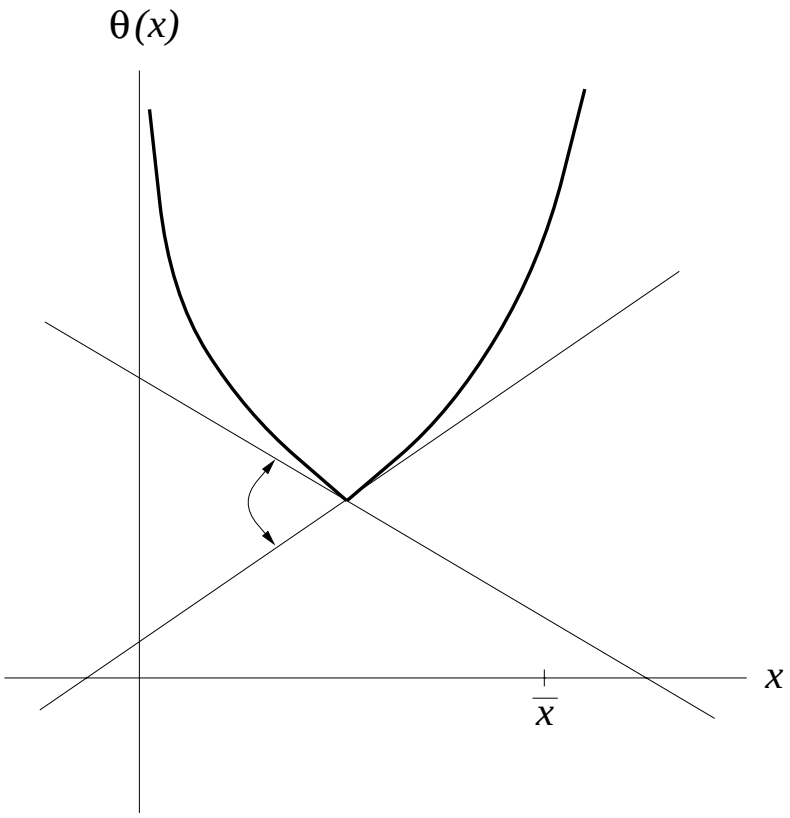


Figure 2.6 The subdifferential to $\theta(x)$ at $\bar{x}$ is the set of slope vectors of all lines in the cone marked by the angle sign.

The set of all subgradients of $g(x)$ at x^0 is denoted by the symbol $\partial g(x^0)$, and called the **subdifferential set** of $g(x)$ at x^0. It can be proved that if $g(x)$ is differentiable at x^0, then its gradient $\nabla g(x^0)$ is the unique subgradient of $g(x)$ at x^0. Conversely if $\partial g(x^0)$ contains a single vector, then $g(x)$ is differentiable at x^0 and $\partial g(x^0) = \{\nabla g(x^0)\}$. See references [2.92–2.94] for these and other related results.

Subgradients of Concave Functions

Let $h(x)$ be a concave function defined on a convex subset $\mathbf{\Gamma} \subset \mathbf{R}^n$. In defining a subgradient vector for $h(x)$ at a point $x^0 \in \mathbf{\Gamma}$, the inequality in (2.39) is just reversed; in other words, d is a subgradient for the concave function $h(x)$ at x^0 if $h(x) \leqq h(x^0) + d^T(x - x^0)$ for all x. With this definition, all the results stated above also hold for concave functions.

Computing a Subgradient

Let $\theta(x)$ be a convex function defined on $\mathbf{R}^n$. Let $\bar{x} \in \mathbf{R}^n$, if $\theta(x)$ is differentiable at $\bar{x}$, then the gradient vector $\nabla\theta(\bar{x})$ is the only subgradient of $\theta(x)$ at $\bar{x}$. If $\theta(x)$ is not differentiable at $\bar{x}$, in general, the computation of a subgradient for $\theta(x)$ at $\bar{x}$ may be hard. However, if $\theta(x)$ is the pointwise supremum of a finite set of differentiable convex functions, say

$$\theta(x) = \text{ Maximum } \{ g_1(x), \ldots, g_m(x) \}$$

where each $g_i(x)$ is differentiable and convex, then the subdifferential of $\theta(x)$ is easily obtained. Let

$$\mathbf{J}(\bar{x}) = \{ i : \quad \theta(\bar{x}) = g_i(\bar{x}) \}$$

the the subdifferential of $\theta(x)$ at $\bar{x}$,

$$\partial\theta(\bar{x}) = \text{ convex hull of } \{ \nabla g_i(\bar{x}) : i \in \mathbf{J}(\bar{x}) \} \ .$$

See references [2.92–2.94].

2.7.2 A Review of Some Fixed Point Theorems

Let $\mathbf{\Gamma} \subset \mathbf{R}^n$ be a compact convex subset with a nonempty interior. Let $f(x) : \mathbf{\Gamma} \to \mathbf{\Gamma}$ be a single valued map, that is, for each $x = (x_1, \ldots, x_n)^T \in \mathbf{\Gamma}$, $f(x) = (f_1(x), \ldots, f_n(x))^T \in \mathbf{\Gamma}$, which is continuous. We have the following celebrated theorem.

Theorem 2.13 : Brouwer's Fixed Point Theorem *If $f(x) : \mathbf{\Gamma} \to \mathbf{\Gamma}$ is continuous, it has a fixed point, that is, the system*

$$f(x) - x = 0 \tag{2.40}$$

which is a system of n equations in n unknowns, has a solution $x \in \mathbf{\Gamma}$.

See references [2.48, 2.50, 2.68, 2.69, 2.72] for proofs of this theorem. We now provide an illustration of this theorem.

Example 2.18

Consider $n = 1$. Let $\mathbf{\Gamma} = \{ x : \ x \in \mathbf{R}^1, \ 0 \leqq x \leqq 1 \}$ denoted by $[0, 1]$. Consider the continuous function $f(x) : [0, 1] \to [0, 1]$. We can draw a diagram for $f(x)$ on the two dimensional Cartesian plane by plotting x on the horizontal axis, and the values of $f(x)$ along the vertical axis, as in Figure 2.7. Since $f(x)$ is defined on $[0, 1]$ the curve of $f(x)$ begins somewhere on the thick vertical line $x = 0$, and goes all the way to the thick vertical line $x = 1$, in a continuous manner. Since $f(x) \in [0, 1]$, the curve for $f(x)$ lies between the two thin horizontal lines $f(x) = 0$ and $f(x) = 1$. The dashed

diagonal line is $f(x) - x = 0$. It is intuitively clear that the curve of $f(x)$ must cross the diagonal of the unit square, giving a fixed point for $f(x)$.

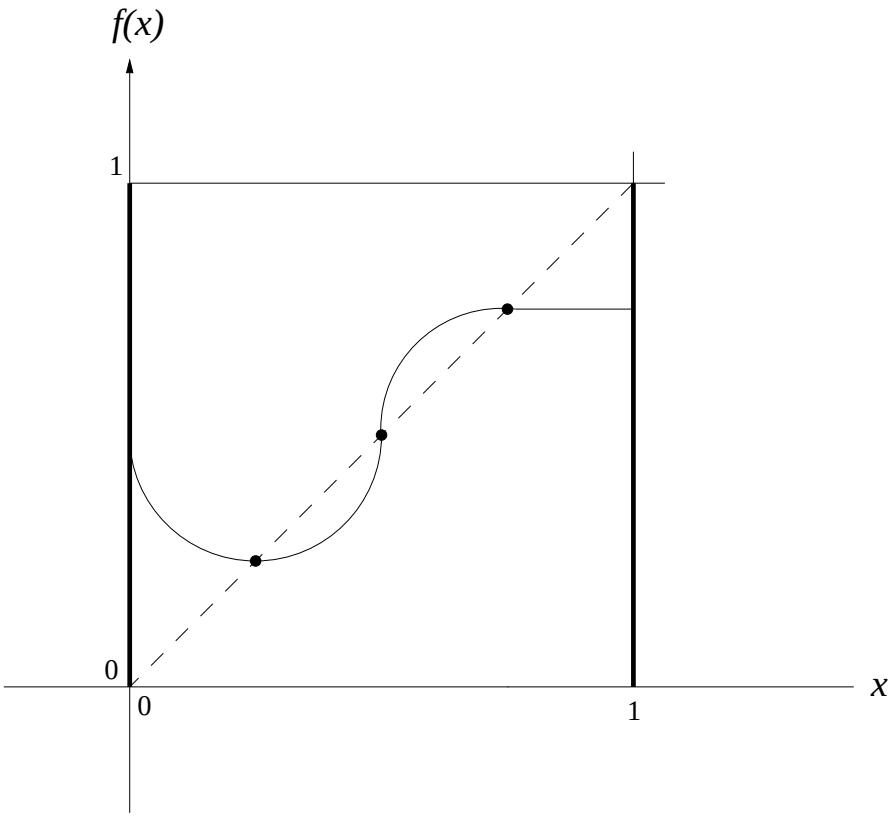


Figure 2.7 The curve of $f(x) : [0,1] \to [0,1]$. Points of intersection of the curve with the dashed diagonal line are the Brouwer fixed points of $f(x)$.

Example 2.19

This example illustrates the need for convexity in Theorem 2.13. Let $n = 2$. Let $\mathbf{K}$ denote the dotted ring in Figure 2.8 between two concentric circles. Let $f(x)$ denote the continuous mapping $\mathbf{K} \to \mathbf{K}$ obtained by rotating the ring through a specified angle θ in the anti-clockwise direction. Clearly this $f(x)$ has no fixed points in $\mathbf{K}$.

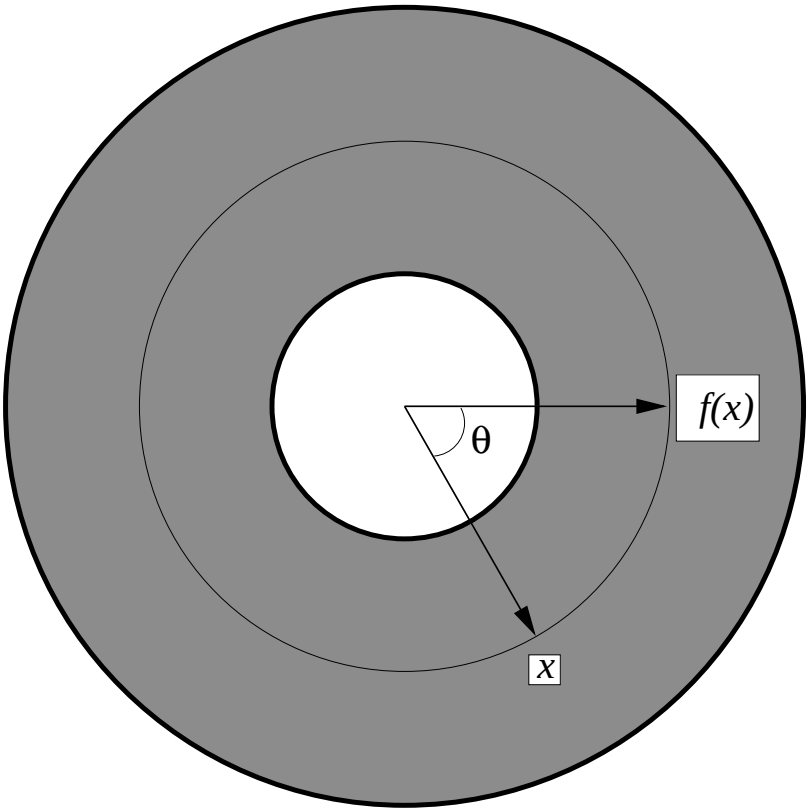


Figure 2.8 The need of convexity for the validity of Brouwer's fixed point theorem.

The need for the boundedness of the set $\mathbf{\Gamma}$ for the validity of Theorem 2.13 follows from the fact that the mapping $f(x) = x + a$ for each $x \in \mathbf{R}^n$, where $a \neq 0$ is a specified point in $\mathbf{R}^n$, has no fixed points. The need for the closedness of the set $\mathbf{\Gamma}$ for the validity of Theorem 2.13 follows from the fact that the mapping $f(x) = \frac{1}{2}(x+1)$ from the set $\{x : 0 \leqq x < 1\}$ into itself has no fixed point in the set.

The system (2.40) is a system of n equality constraints in n unknowns. An effort can be made to solve (2.40) using methods for solving nonlinear equations.

A Monk's Story

The following story of a monk provides a nice intuitive justification for the concept and the existence of a fixed point. A monk is going on a pilgrimage to worship in a temple at the top of a mountain. He begins his journey on Saturday morning at 6:00 AM promptly. The path to the temple is steep and arduous and so narrow that trekkers on it have to go in a single file. Our monk makes slow progress, he takes several breaks on the way to rest, and at last reaches the temple by evening. He spends the night worshipping at the temple. Next morning, he begins his return trip from the temple exactly at 6:00 AM, by the same path. On the return trip, since the path is downhill, he makes fast progress and reaches the point from where he started his hjourney on Saturday morning, well before the evening.

Suppose we call a point (or spot or location) on the path, a fixed point, if the monk was exactly at that spot at precisely the same time of the day on both the forward and return trips.

The existence of a fixed point on the path can be proved using Brouwer's fixed point theorem, but there is a much simpler and intuitive proof for its existence (see A. Koestler, *The Act of Creation*, Hutchinson, 1976, London). Imagine that on Saturday morning exactly at 6:00 AM, a duplicate monk starts from the temple, down the mountain, proceeding at every point of time at exactly the same rate that the original monk would on Sunday. So, at any point of time of the day on Saturday, the duplicate monk will be at the same location on the path as the original monk will be at the time on Sunday. Since the path is so narrow that both cannot pass without being in each other's way, the two monks must meet at some time during the day, and the spot on the path where they meet is a fixed point.

Successive Substitution Method for Computing a Brouwer's Fixed Point

One commonly used method to compute a Brouwer's fixed point of the single valued map $f(x) : \mathbf{\Gamma} \to \mathbf{\Gamma}$ is an iterative method that begins with an arbitrary point $x^0 \in \mathbf{\Gamma}$, and obtains a sequence of points $\{x^r : r = 0, 1, \ldots\}$ in $\mathbf{\Gamma}$ using the iteration

$$x^{r+1} = f(x^r) \ .$$

The sequence so generated, converges to a Brouwer's fixed point of $f(x)$ if $f(x)$ satisfies the **contraction property**, that is, if there exists a constant ν satisfying $0 \leqq \nu < 1$ such that for every $x, y \in \mathbf{\Gamma}$, we have

$$\|f(x) - f(y)\| \leqq \nu \|x - y\| \ . \tag{2.41}$$

If the map $f(x)$ satisfies the contraction proprety, this successsive substituitions method is a very convenient method for computing a Brouwer's fixed point of $f(x)$. Unfortunately, the contraction property is a strong property and does not usually hold in many practical applications.

Newton-Raphson Method for Solving a System of n Equations in n Unknowns

The system (2.40) is a system of n equations in n unknowns, and we can try to solve it using approaches for solving nonlinear equations of this type, like **Newton-Raphson method**, which we now present. The method is also called Newton's method often in the literature, or Newton's method for solving equations. Consider the system

$$g_i(x) = 0 \quad i = 1 \text{ to } n \tag{2.42}$$

where each $g_i(x)$ is a real valued function defined on $\mathbf{R}^n$. Assume that each function $g_i(x)$ is differentiable. Let $\nabla g_i(x)$ be the row vector of partial derivatives and let the **Jacobian** be

$$\nabla g(x) = \begin{pmatrix} \nabla g_1(x) \\ \vdots \\ \nabla g_n(x) \end{pmatrix}$$

in which the i^{th} row vector is the partial derivative vector of $g_i(x)$ written as a row.

To solve (2.42) the Newton-Raphson method begins with an arbitrary point x^0 and generates a sequence of points $\{x^0, x^1, x^2, \ldots\}$. Given x^r in the sequence, the method approximates (2.42) by its first order Taylor approximation around x^r leading to

$$g(x^r) + \nabla g(x^r)(x - x^r) = 0$$

whose solution is $x^r - (\nabla g(x^r))^{-1} g(x^r)$, which is taken as the next point in the sequence. This leads to the iteration

$$x^{r+1} = x^r - (\nabla g(x^r))^{-1} g(x^r) \ .$$

If the Jacobian is nonsingular, the quantity $y = (\nabla g(x^r))^{-1} g(x^r)$ can be computed efficiently by solving the system of linear equations

$$(\nabla g(x^r))y = g(x^r)$$

If the Jacobian $\nabla g(x^r)$ is singular, the inverse $(\nabla g(x^r))^{-1}$ does not exist and the method is unable to proceed further. Several modifications have been proposed to remedy this situation, see references [10.9, 10.13, 10.33]. Many of these modifications are based on the applications of Newton's method for unconstrained minimization or a modified version of it (see Sections 10.8.4, 10.8.5) to the least squares formulation of (2.42) leading to problem of finding the unconstrained minimum of

$$\sum_{i=1}^{n} (g_i(x))^{\mathbf{2}} \ .$$

As an example, consider the system

$$g_1(x) = x_1^{\mathbf{2}} + x_2^{\mathbf{2}} - 1 = 0$$
$$g_2(x) = x_1^{\mathbf{2}} - x_2 = 0$$

The Jacobian matrix is

$$\begin{pmatrix} 2x_1 & 2x_2 \\ 2x_1 & -1 \end{pmatrix}$$

Let $x^0 = (1, 0)^T$ be the initial point. So $g(x^0) = (0, 1)^T$. The Jacobian matrix at x^0 is $\begin{pmatrix} 2 & 0 \\ 2 & -1 \end{pmatrix}$. This leads to the next point $x^1 = (1, 1)^T$. It can be verified that

$x^2 = \left(\frac{5}{6}, \frac{2}{3}\right)^T$, and so on. The actual solution in this example can be seen from Figure 2.9.

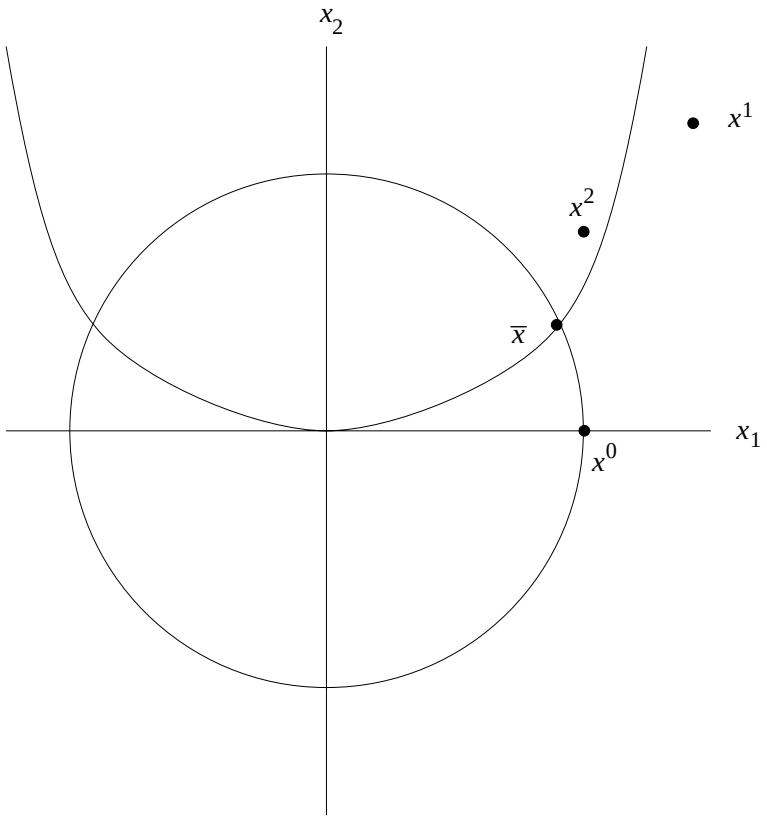


Figure 2.9 The circle here is the set of all points (x_1, x_2) satisfying $x_1^2 + x_2^2 - 1 = 0$. The parabola is the set of all points satisfying $x_1^2 - x_2 = 0$. The two intersect in two points (solutions of the system) one of which is $\bar{x}$. Beginning with x^0, the Newton-Raphson method obtains the sequence x^1, $x^2, \ldots$ converging to $\bar{x}$.

In order to solve (2.40) by Newton-Raphson method or some modified versions of it, the map $f(x)$ must satisfy strong properties like being differentiable etc., which do not hold in may many practical applications. Thus, to use Brouwer's fixed point theorem in practical applications we should devise methods for solving (2.40) without requiring the map $f(x)$ to satisfy any conditions besides continuity. In 1967 H. Scarf in a pioneering paper [2.68] developed a method for finding an approximate solution of (2.40) using a triangulation of the space, that walks through the simplices of the triangulation along a path satisfying properties similar to the one traced by the complementary pivot algorithm for the LCP. This method has the advantage that it works without requiring any conditions on the map $f(x)$ other than those required by Brouwer's theorem for the existence of the fixed point (i. e., continuity). Subsequently vastly improved versions of these methods have been developed by many researches. We will discuss one of these methods in detail.

Approximate Brouwer Fixed Points

Let $f(x) : \mathbf{\Gamma} \rightarrow \mathbf{\Gamma}$ be continuous as defined in Theorem 2.13. A true Brouwer fixed point of $f(x)$ is a solution of (2.40). However, in general, we may not be able to compute an exact solution of (2.40) using finite precision arithmetic. In practice, we attempt to compute an approximate Brouwer fixed point. There are two types of approximate Brouwer fixed points, we define them below.

Type 1: A point $\bar{x} \in \mathbf{\Gamma}$ is said to be an approximate Brouwer fixed point of $f(x)$ of Type 1 if

$$||\bar{x} - f(\bar{x})|| < \varepsilon$$

for some user selected tolerance ε (a small positive quantity).

Type 2: A point $x^* \in \mathbf{\Gamma}$ is said to be an approximate Brouwer fixed point of Type 2 if there exists an exact solution y of (2.40) such that

$$||x^* - y|| < \varepsilon \ .$$

In general, a Type 1 approximate Brouwer fixed point $\bar{x}$ may not be a Type 2 approximate Brouwer fixed point, that is, $\bar{x}$ may be far away from any exact solution of (2.40). If some strong conditions hold (such as: $f(x)$ is continuously differentiable in the interior of $\mathbf{\Gamma}$ and all the derivatives are Lipschitz continuous, or $f(x)$ is twice continuously differentiable in the interior of $\mathbf{\Gamma}$) a Type 1 approximate Brouwer fixed point can be shown to be also a Type 2 approximate Brouwer fixed point with a modified tolerance. At any rate, the algorithms discussed in the following sections are only able to compute approximate Brouwer fixed points of Type 1.

Kakutani Fixed Points

In many applications, the requirement that $f(x)$ be a point-to-point map is itself too restrictive. In 1941 S. Kakutani generalized Theorem 2.13 to point-to-set maps. As before, let $\mathbf{\Gamma}$ be a compact convex subset of $\mathbf{R}^n$. Let $\mathbf{F}(x)$ be a point-to-set map on $\mathbf{\Gamma}$, that is, for each $x \in \mathbf{\Gamma}$, $\mathbf{F}(x)$ is itself a specified subset of $\mathbf{\Gamma}$.

Example 2.20

Let $n = 1$. Let $\mathbf{\Gamma} = \{ x \in \mathbf{R}^1 : 0 \leqq x \leqq 1 \}$. For each $x \in \mathbf{\Gamma}$, suppose $\mathbf{F}(x) = \{ y : x \leqq y \leqq 1 \} = [x, 1]$. See Figure 2.10.

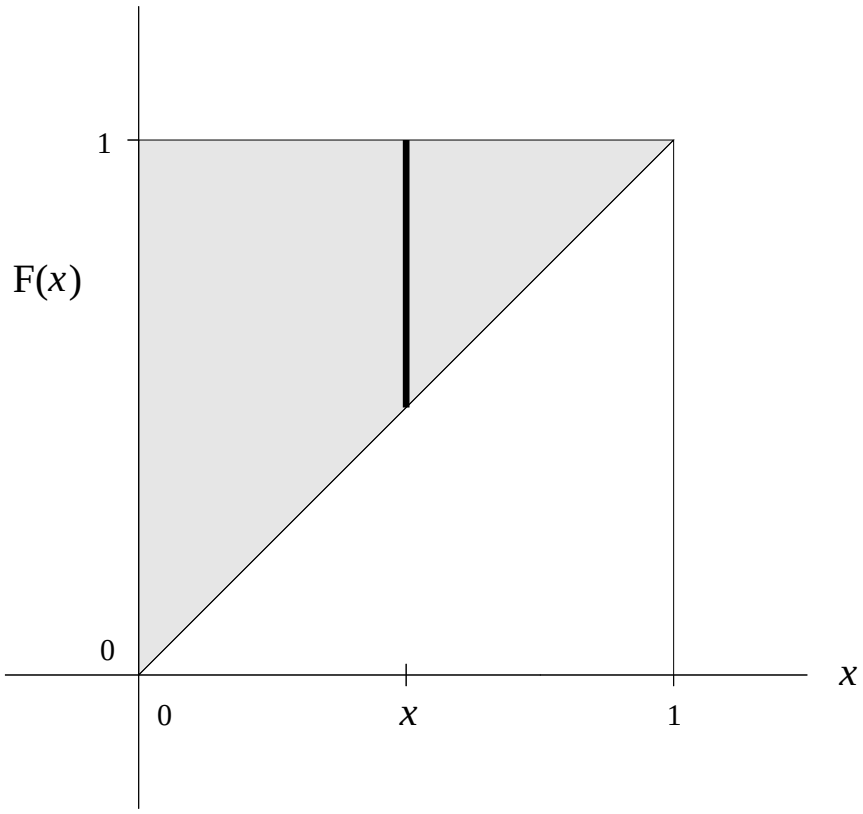


Figure 2.10 For each $x \in \mathbf{\Gamma}$, $\mathbf{F}(x)$ is the closed interval $[x, 1]$.

We consider only maps in which $\mathbf{F}(x)$ is a compact convex subset of $\mathbf{\Gamma}$ for each $x \in \mathbf{\Gamma}$. The point-to-set map $\mathbf{F}(x)$ is said to be an USC (Upper Semi-Continuous) map if it satisfies the following properties. Let $\{ x^k : k = 1, 2, \ldots \}$ be any sequence of points in $\mathbf{\Gamma}$ converging to a point $x^\star \in \mathbf{\Gamma}$. For each k, suppose y^k is an arbitrary point selected from $\mathbf{F}(x^k)$, $k = 1, 2, \ldots$. Suppose that the sequence $\{ y^k : k = 1, 2, \ldots \}$ converges to the point $y^\star$. The requirement for the upper semi-continuity of the point-to-set map $\mathbf{F}(x)$ is that these conditions imply that $y^\star \in \mathbf{F}(x^\star)$.

It can be verified that the point-to-set map $\mathbf{F}(x)$ given in Figure 2.8 satisfies this USC property.

Theorem 2.14 Kakutani's Fixed Point Theorem

If $\mathbf{F}(x)$ is a USC point-to-set map defined on the compact convex subset $\mathbf{\Gamma} \subset \mathbf{R}^n$, there exists a point $x \in \mathbf{\Gamma}$ satisfying

$$x \in \mathbf{F}(x) \ . \tag{2.43}$$

Any point satisfying (2.43) is known as a **Kakutani's fixed point** of the point-to-set map $\mathbf{F}(x)$. To prove his theorem, Kakutani used the fundamental notion of a **piecewise linear approximation** to the map $\mathbf{F}(x)$. The same picewise linear approximation scheme is used in the method discussed later on for computing fixed points. See reference [2.50] for the proof of Kakutani's theorem.

For each $x \in \mathbf{\Gamma}$, if $\mathbf{F}(x)$ is a singleton set (i. e., a set containing only a single element) $\{f(x)\} \subset \mathbf{\Gamma}$, it can be verified that this $\mathbf{F}(x)$ is USC iff $f(x)$ is continuous. Thus the USC property of point-to-set maps is a generalization of the continuity property of point-to-point maps. Also, every Brouwer fixed point of the point-to-point map $f(x)$ can be viewed as a Kakutani fixed point of $\mathbf{F}(x) = \{f(x)\}$.

Approximate Kakutani Fixed Points

Given the USC point-to-set map $\mathbf{F}(x)$ as defined in Theorem 2.14, a Kakutani fixed point is a point $x \in \mathbf{\Gamma}$ satisfying (2.43). As under the Brouwer fixed point case, using finite precision arithmetic, we may not be able to find $x \in \mathbf{\Gamma}$ satisfying (2.43) exactly. We therefore attempt to compute an approximate Kakutani fixed point. Again, there are two types of approximate Kakutani fixed points, we define them below

Type 1: A point $\bar{x} \in \mathbf{\Gamma}$ is said to be an approximate Kakutani fixed point of $\mathbf{F}(x)$ of Type 1 if there exists a $z \in \mathbf{F}(\bar{x})$ satisfying

$$||\bar{x} - z|| < \varepsilon$$

for some user selected tolerance ε (a small positive quantity).

Type 2: A point $x^* \in \mathbf{\Gamma}$ is said to be an approximative Kakutani fixed point of $\mathbf{F}(x)$ of Type 2 if there exists a y satisfying (2.43) and

$$||x^* - y|| < \varepsilon \ .$$

The algorithms discussed in the following sections are only able to compute Type 1 approximate Kakutani fixed points.

Use in Practical Applications

In pratical applications we have to deal with either point-to-point or point-to-set maps defined over the whole space $\mathbf{R}^n$, not necessarily on only a compact convex subset of $\mathbf{R}^n$. Also, it is very hard, if not computationally impossible, to check whether properties like USC etc. hold for our maps. For such maps, the existence of a fixed point is not guaranteed. Because of this, the algorithms that we discuss for computing fixed points may not always work on these problems. Also, it is impossible for us to continue the computation indefinitely, we have to terminate after a finite number of steps. In practice, from the path traced by the algorithm, it will be clear whether it seems to be converging, or running away. If it seems to be converging, from the point

obtained at termination, an approximate solution of the problem can be obtained. If the algorithm seems to be running away, either we can conclude that the algorithm has failed to solve the problem, or an effort can be made to run the algorithm again with different initial conditions. Before discussing the algorithm, we will now discuss some standard applications of fixed point computing.

2.7.3 Applications in Unconstrained Optimization

Let $\theta(x)$ be a real valued function defined on $\mathbf{R}^n$ and suppose it is required to solve the problem

$$\begin{array}{ll} \text{minimize} & \theta(x) \\ \text{over} & x \in \mathbf{R}^n \end{array} \tag{2.44}$$

If $\theta(x)$ is differentiable, a necessary condition for a point $x \in \mathbf{R}^n$ to be a local minimum for (2.44) is

$$\nabla\theta(x) = 0 \tag{2.45}$$

which is a system of n equations in n unknowns. Define $f(x) = x - (\nabla\theta(x))^T$. Then every Brouwer fixed point of $f(x)$ is a solution of (2.45) and vice versa. Hence every fixed point of $f(x)$ satisfies the first order necessary optimality conditions for (2.44). If $\theta(x)$ is convex, every solution of (2.45) is a global minimum for (2.44) and vice versa, and hence in this case (2.44) can be solved by computing a fixed point for $f(x)$ defined above. However, if $\theta(x)$ is not convex, there is no guarantee that a solution of (2.45), (i.e., a fixed point of $f(x) = x - (\nabla\theta(x))^T$) is even a local minimum for (2.44) (it could in fact be a local maximum). So, after obtaining an approximate fixed point, $\bar{x}$, of $f(x)$, one has to verify whether it is a local minimum or not. If $\theta(x)$ is twice continuously differentiable, a sufficient condition for a solution of (2.45) to be a local minimum for (2.44) is that the Hessian matrix $H(\theta(\bar{x}))$ be positive definite.

If $\theta(x)$ is not differentiable at some points, but is convex, then the subdifferential set $\partial\theta(x)$ exists for all x. In this case define $\mathbf{F}(x) = \{ x - y : y \in \partial\theta(x) \}$. Then every Kakutani fixed point of $\mathbf{F}(x)$ is a global minimum for (2.44) and vice versa.

One strange feature of the fixed point formulation for solving (2.45) is worth mentioning. Define $\mathbf{G}(x) = \{ x + y : y \in \partial\theta(x) \}$. Clearly, every Kakutani fixed point of $\mathbf{G}(x)$ also satisfies the necessary optimality conditions for (2.44). Mathematically, the problems of finding a Kakutani fixed point of $\mathbf{F}(x)$ or $\mathbf{G}(x)$ are equivalent, but the behavior of the fixed point computing algorithm discussed in Section 2.7.8 on the two problems could be very different. This is discussed later on under the subsection entitled, "Sufficient Conditions for Finite Termination" in Section 2.7.8. In practical applications, one might try computing the Kakutani fixed point of $\mathbf{F}(x)$ using the algorithm discussed in Section 2.7.8, and if its performance is not satisfactory switch over and use the same algorithm on $\mathbf{G}(x)$ instead.

2.7.4 Application to Solve a System of Nonlinear Inequalities

Consider the system

$$g_i(x) \leqq 0, \quad \text{for } i = 1 \text{ to } m \tag{2.46}$$

where each $g_i(x)$ is a real valued convex function defined on $\mathbf{R}^n$. Define the pointwise supremum function $s(x) = \text{Maximum } \{ g_1(x), \ldots, g_m(x) \}$. As discussed earlier, $s(x)$ is itself convex, and $\partial s(x) \supset \bigcup_{i \in \mathbf{J}(x)} \partial g_i(x)$, where $\mathbf{J}(x) = \{ i : g_i(x) = s(x) \}$. If each $g_i(x)$ is differentiable, then $\partial s(x) =$ convex hull of $\{ \nabla g_i(x) : i \in \mathbf{J}(x) \}$. If (2.46) has a feasible solution $\bar{x}$, then $s(\bar{x}) \leqq 0$, and conversely every point x satisfying $s(x) \leqq 0$ is feasible to (2.46). So the problem of finding a feasible solution of (2.46) can be tackled by finding the unconstrained minimum of $s(x)$, which is the same as the problem of finding a Kakutani fixed point of $\mathbf{F}(x) = \{ x - y : y \in \partial s(x) \}$ as discussed in Section 2.7.3. If $\bar{x}$ is a Kakutani fixed point of this map and $\bar{s}(x) > 0$, (2.46) is infeasible. On the other hand if $s(\bar{x}) \leqq 0$, $\bar{x}$ is a feasible solution of (2.46).

2.7.5 Application to Solve a System of Nonlinear Equations

Consider the system of equations

$$h_i(x) = 0, \quad i = 1 \text{ to } r \tag{2.47}$$

where each $h_i(x)$ is a real valued function defined on $\mathbf{R}^n$. Let $h(x) = (h_1(x), \ldots, h_r(x))^T$. If $r > n$, (2.47) is said to be an **overdetermined system**. In this case there may be no solution to (2.47), but we may be interested in finding a point $x \in \mathbf{R}^n$ that satisfies (2.47) as closely as possible. The **least squares approach** for finding this is to look for the unconstrained minimum of $\sum_{i=1}^{r} (h_i(x))^{\mathbf{2}}$, which can be posed as a fixed point problem as in Section 2.7.3.

If $r < n$, (2.47) is known as an **underdetermined system**, and it may have many solutions. It may be possible to develop additional $n - r$ equality constraints which when combined with (2.47) becomes a system of n equations in n unknowns. Or the least squares method discussed above can be used here also.

Assume that $r = n$. In this case define $f(x) = x - h(x)$. Then every Brouwer fixed point of $f(x)$ solves (2.47) and vice versa. As mentioned in Section 2.7.3, it may be worthwhile to also consider the equivalent problem of computing the fixed point of $d(x) = x + h(x)$ in this case.

2.7.6 Application to Solve the Nonlinear Programming Problem

Consider the nonlinear program

$$\begin{array}{ll} \text{Minimize} & \theta(x) \\ \text{subject to} & g_i(x) \leqq 0, \quad i = 1 \text{ to } m \end{array} \tag{2.48}$$

where $\theta(x)$, $g_i(x)$ are real valued functions defined over $\mathbf{R}^n$. We will assume that each of these functions is convex, and continuously differentiable. We make an additional assumption that if (2.48) is feasible (i. e., the set $\{ x : g_i(x) \leqq 0, i = 1 \text{ to } m \} \neq \emptyset$), then there exists an $x \in \mathbf{R}^n$ satisfying $g_i(x) < 0$, for each $i = 1$ to m. This assumption is known as a **constraint qualification**. As before, let $s(x)$ be the pointwise supremum function, maximum $\{ g_1(x), \ldots, g_m(x) \}$. Then (2.48) is equivalent to

$$\begin{array}{ll} \text{Minimize} & \theta(x) \\ & s(x) \leqq 0 \end{array} \tag{2.49}$$

By our assumption, and the results discussed earlier, $s(x)$ is also convex and $\partial s(x) =$ convex hull of $\{ \nabla g_i(x) : i \in \mathbf{J}(x) \}$, where $\mathbf{J}(x) = \{ i : s(x) = g_i(x) \}$. Consider the following point-to-set mapping defined on $\mathbf{R}^n$.

$$\mathbf{F}(x) = \begin{cases} \{ x - (\nabla\theta(x))^T \}, & \text{if } s(x) < 0 \ , \\ \{ x - y : \ y \in \text{ convex hull of } \{\nabla\theta(x), \partial s(x)\} \}, & \text{if } s(x) = 0 \ , \\ \{ x - y : \ y \in \partial s(x) \}, & \text{if } s(x) > 0 \ . \end{cases} \tag{2.50}$$

Under our assumptions of convexity and differentiability, it can be verified that $\mathbf{F}(x)$ defined in (2.50) is USC. Let $\bar{x}$ be a Kakutani fixed point of $\mathbf{F}(x)$. If $s(\bar{x}) < 0$, then $0 = \nabla\theta(\bar{x})$, and thus $\bar{x}$ is a global minimum for $\theta(x)$ over $\mathbf{R}^n$ and is also feasible to (2.48), and therefore solves (2.48). If $s(\bar{x}) > 0$, then $0 \in \partial s(\bar{x})$, thus 0 is a global minimum of $s(x)$, and since $s(\bar{x}) > 0$, (2.48) has no feasible solution. If $s(\bar{x}) = 0$, then $0 \in$ convex hull of $\{\nabla\theta(\bar{x}), \partial s(\bar{x}) \} =$ convex hull of $\{\nabla\theta(\bar{x}), \nabla g_i(\bar{x})$ for $i \in \mathbf{J}(\bar{x})\}$, so there exists nonnegative numbers λ_0, λ_i for $i \in \mathbf{J}(\bar{x})$ satisfying

$$\begin{gathered} \lambda_0 \nabla\theta(\bar{x}) + \sum_{i \in \mathbf{J}(\bar{x})} \lambda_i \nabla g_i(\bar{x}) = 0 \\ \lambda_0 + \sum_{i \in \mathbf{J}(\bar{x})} \lambda_i = 1 \\ \lambda_0, \ \lambda_i \geqq 0 \quad \text{for all } i \in \mathbf{J}(\bar{x}) \end{gathered} \tag{2.51}$$

If $\lambda_0 = 0$, (2.51) implies that $0 \in \partial s(\bar{x})$ and so $s(\bar{x})$ is a global minimizer of $s(\bar{x})$, $\bar{x}$ is feasible to (2.48) since $s(\bar{x}) = 0$, and these facts lead to the conclusion that $\{ x : g_i(x) \leqq 0, \text{ for } i = 1 \text{ to } m \} \neq \emptyset$ and yet there exists no x satisfying $g_i(x) < 0$ for all $i = 1$ to m, violating our constraint qualification assumption. So $\lambda_0 > 0$ in (2.51).

So if we define $\bar{\pi}_i = \frac{\lambda_i}{\lambda_0}$ if $i \in \mathbf{J}(\bar{x})$, $= 0$ otherwise, then from (2.51) we conclude that $\bar{x}$, $\bar{\pi}$ together satisfy the Karush-Kuhn-Tucker necessary conditions for optimality for (2.48), and our convexity assumption imply that $\bar{x}$ is the global minimum for (2.48).

Thus solving (2.48) is reduced to the problem of finding a Kakutani fixed point of the mapping $\mathbf{F}(x)$ defined in (2.50).

Example 2.21

Consider the problem :

$$\begin{array}{ll} \text{minimize} & \theta(x) = x_1^2 + x_2^2 - 2x_1 - 3x_2 \\ \text{subject to} & g_1(x) = x_1 + x_2 \leqq 1 \end{array} \tag{2.52}$$

Clearly $\nabla\theta(x) = (2x_1-2, 2x_2-3)$, $\nabla g_1(x) = (1,1)$. The mapping $\mathbf{F}(x)$ for this problem is

$$\mathbf{F}(x) = \begin{cases} \{ -x_1 + 2, -x_2 + 3)^T \}, & \text{if } x_1 + x_2 < 1 \ , \\ \text{Convex hull of } \{(-x_1 + 2, -x_2 + 3)^T, (x_1 - 1, x_2 - 1)^T\}, & \text{if } x_1 + x_2 = 1 \ , \\ \{ (x_1 - 1, x_2 - 1)^T \}, & \text{if } x_1 + x_2 > 1 \ . \end{cases}$$

It can be verified that $\bar{x} = (\frac{1}{4}, \frac{3}{4})^T$ is a Kakutani fixed point of this mapping $\mathbf{F}(x)$, and that $\bar{x}$ is the global optimum solution of the nonlinear program (2.52).

If $\theta(x)$, $g_i(x)$ are all continuously differentiable, but not necessarily convex, we can still define the point-to-set mapping $\mathbf{F}(x)$ as in (2.50) treating $\partial s(x) =$ convex hull of $\{ \nabla g_i(x) : i \in \mathbf{J}(x) \}$. In this general case, any Kakutani fixed point $\bar{x}$ of $\mathbf{F}(x)$ satisfies the first order necessary optimality conditions for (2.48), but these conditions are not sufficient to guarantee that $\bar{x}$ is a global or even a local minimum for (2.48), see Section 10.2 for definitions of a global minimum, local minimum. One can then try to check whether $\bar{x}$ satisfies some sufficient condition for being a local minimum for (2.48) (for example, if all the functions are twice continuously differentiable, a sufficient condition for $\bar{x}$ to be a local minimum for (2.48) is that the Hessian matrix of the Lagrangian with respect to x is positive definite at $\bar{x}$. See references [10.2, 10.3, 10.13, 10.17, A8, A12]). If these sufficient optimality conditions are not satisfied, it may be very hard to verify whether $\bar{x}$ is even a local minimum for (2.48). As an example, consider the problem: minimize $x^T Dx$, subject to $x \geqq 0$. The point $0 \in \mathbf{R}^n$ is a global minimum for this problem if D is PSD. If D is not PSD, 0 is a local minimum for this problem iff D is a copositive matrix. Unfortunately, there are as yet no efficient methods known for checking whether a matrix which is not PSD, is copositive. See Section 2.9.3.

Thus, in the general nonconvex case, the fixed point approach for (2.48) finds a point satisfying the first order necessary optimality conditions for (2.48), by computing a Kakutani fixed point of $\mathbf{F}(x)$ defined in (2.50). In this general case, many of the other solution techniques of nonlinear programming for solving (2.48) (see Chapter 10) are usually based on **descent methods**. These techniques generate a sequence of points

$\{ x^r : r = 0, 1, \ldots \}$. Given x^r, they generate a $y^r \neq 0$ such that the direction $x^r + \lambda y^r$, $\lambda \geqq 0$, is a **descent direction**, that is, it is either guaranteed to decrease the objective value or a measure of the infeasibility of the current solution to the problem or some criterion function which is a combination of both. The next point in the sequence x^{r+1} is usually taken to be the point which minimizes the criterion function on the half line $\{ x^r + \lambda y^r : \lambda \geqq 0 \}$ obtained by using some one dimensional (λ is the only variable to be determined in this problem) line minimization algorithm. And the whole process is then repeated with the new point. On general problems, these methods suffer from the same difficulties, they cannot theoretically guarantee that the point obtained at termination is even a local minimum. However, these descent methods do seem to have an edge over the fixed point method presented above in the general case. In the absence of convexity, one has more confidence that a solution obtained through a descent process is likely to be a local minimum, than a solution obtained through fixed point computation which is based purely on first order necessary conditions for optimality.

The approach for solving the nonlinear program (2.48) using the fixed point transformation has been used quite extensively, and seems to perform satisfactorily. See references [2.40, 2.58, 2.59].

Many practical nonlinear programming models tend to be nonconvex. The fixed point approach outlined above, provides additional arsenal in the armory for tackling such general problems.

Now consider the general nonlinear programming problem in which there are both equality and inequality constraints.

$$\begin{array}{lll} \text{minimize} & \theta(x) & \\ \text{subject to} & g_i(x) \leqq 0, & i = 1 \text{ to } m \\ & h_t(x) = 0, & t = 1 \text{ to } p \end{array} \tag{2.53}$$

The usual approach for handling (2.53) is the **penality function method** which includes a term with a large positive coefficient corresponding to a measure of violation of the equality constraints in the objective function. One such formulation leads to the problem

$$\begin{array}{lll} \text{minimize} & \theta(x) + \alpha \sum_{t=1}^{p} (h_t(x))^{\mathbf{2}} & \\ \text{subject to} & g_i(x) \leqq 0, & i = 1 \text{ to } m \end{array} \tag{2.54}$$

In (2.54), α, a large positive number, is the **penalty parameter**. If (2.53) has a feasible solution, every optimum solution of (2.54) would tend to satisfy $h_t(x) = 0$, $t = 1$ to p as α becomes very large, and thus would also be optimal to (2.53). When α is fixed to be a large positive number, (2.54) is in the same form as (2.48), and can be tackled through a fixed point formulation as discussed above.

Advantages and Disadvantages of this Approach

In the NLP (2.48) there may be several constraints (i, e., m may be large) and the problem difficulty can be expected to increase with the number of constraints. The

fixed point approach for solving (2.48), first transforms (2.48) into the equivalent (2.49), which is an NLP in which there is only a single constraint. The fact that (2.49) is a single constraint problem is definitely advantageous.

The original problem (2.48) is a smooth problem since the objective and constraint functions are all assumed to be continuously differentiable. Eventhough $g_i(x)$ are continuously differentiable for all i, there may be points x where $s(x)$ is not differentiable. However, $s(x)$ is differentiable almost everywhere and so (2.49) is a nonsmooth NLP. That this approach transforms a nice smooth NLP into a nonsmooth NLP is a disadvantage. But, because of the special nature of the function $s(x)$, for any x, we are able to compute a point in the subdifferential set $\partial s(x)$ efficiently, as discussed above. For computing a fixed point of the map $\mathbf{F}(x)$ defined in (2.50), the algorithms discussed in the following sections need as inputs only subroutines to compute $\nabla\theta(x)$, or a point from $\partial s(x)$ for any given x, which are easy to provide. Thus, eventhough (2.49) is a nonsmooth NLP, the fixed point approach is able to handle it efficiently. Practical computational experience with this approach is quite encouraging.

The fixed point approach solves NLPs using only the first order necessary conditions for optimality. The objective value is never computed at any point. This is a disadvantage in this approach. In nonconvex NLPs, a solution to the first order necessary conditions for optimality, may not even be a local minimum. Since the objective value is not used or even computed in this approach, we lack the circumstantial evidence, or the neighborhood information about the behaviour of objective values, to conclude that the final solution obtained is at least likely to be a local minimum.

2.7.7 Application to Solve the Nonlinear Complementarity Problem

As discussed in Section 1.6, the nonlinear complementary problem (NLCP) is the following. Given $g(x) = (g_1(x), \ldots, g_n(x))^T : \mathbf{R}^n_+ \to \mathbf{R}^n$, where $\mathbf{R}^n_+$ is the nonnegative orthant of $\mathbf{R}^n$, find $x \geqq 0$ satisfying $g(x) \geqq 0$, $x^T g(x) = 0$.

Define $\psi(x) =$ Maximum $\{-x_1, \ldots, -x_n\}$. So $\partial\psi(x) =$ convex hull of $\{-I_{.j} : j$ such that $-x_j \geqq -x_i$ for all $i = 1$ to n in $x\}$. Define the point-to-set map on $\mathbf{R}^n$,

$$\mathbf{F}(x) = \begin{cases} \{\, x - y : y \in \partial\psi(x) \,\}, & \text{if } \psi(x) > 0 \ , \\ \{\, x - y : y \in \text{ convex hull of } \{g(x), \partial\psi(x)\} \,\}, & \text{if } \psi(x) = 0 \ , \\ \{\, x - g(x)\}, & \text{if } \psi(x) < 0 \ . \end{cases} \tag{2.55}$$

It can be verified that every Kakutani fixed point of $\mathbf{F}(x)$ defined here is a solution of the NLCP and vice versa. Thus the NLCP can be solved by computing a Kakutani fixed point of $\mathbf{F}(x)$.

2.7.8 Merrill's Algorithm for Computing a Kakutani Fixed Point

Let $\mathbf{F}(x)$ be a point-to-set map defined on $\mathbf{R}^n$. We describe in this section, Merrill's method for computing a Kakutani fixed point of $\mathbf{F}(x)$.

Data Requirements of the Algorithm

If the algorithm requires the storage of the complete set $\mathbf{F}(x)$ for any x, it will not be practically useful. Fortunately, this algorithm does not require the whole set $\mathbf{F}(x)$ for even one point $x \in \mathbf{R}^n$. It only needs a computational procedure (or a subroutine), which, for any given $x \in \mathbf{R}^n$, outputs one point from the set $\mathbf{F}(x)$. The algorithm will call this subroutine a finite number of times. Thus the data requirements of the algorithm are quite modest, considering the complexity of the problem being attempted, and it can be implemented for the computer very efficiently. Also, the primary computational step in the algorithm is the pivot step, which is the same as that in the simplex method for linear programs.

n-Dimensional Simplex

The points $v_1, \ldots, v_r$ in $\mathbf{R}^n$ are the vertices of an $(r-1)$ dimensional simplex if the set of column vectors $\left\{ \begin{pmatrix} 1 \\ v_1 \end{pmatrix}, \ldots, \begin{pmatrix} 1 \\ v_r \end{pmatrix} \right\}$ in $\mathbf{R}^{n+1}$ form a linearly independent set. The simplex itself is the convex hull of its vertices and will be denoted by the symbol $\langle v_1, \ldots, v_r \rangle$. Given the simplex with vertices $v_1, \ldots, v_r$, the convex hull of any subset of its vertices is a **face** of the simplex. An n-dimensional simplex has $(n+1)$ vertices. See Figure 2.11. Clearly a 1-dimensional simplex is a line segment of positive length joining two distinct points, a 2-dimensional simplex is the triangle enclosed by three points which are not collinear, etc.

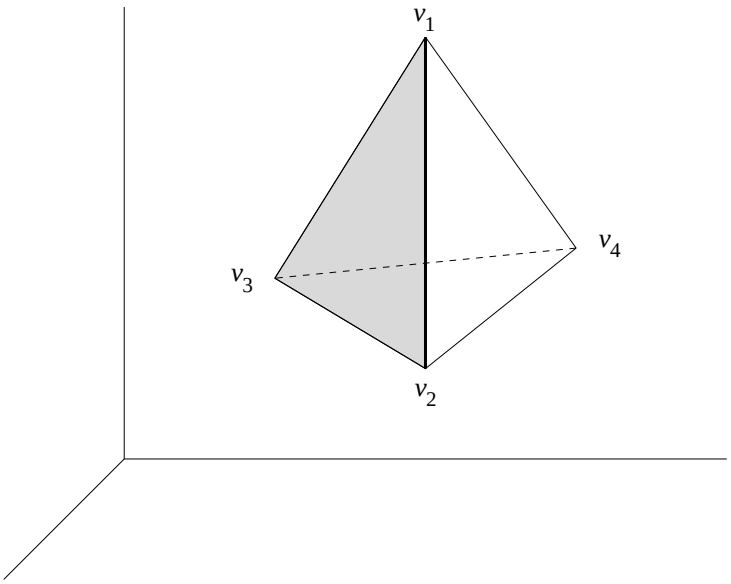


Figure 2.11 The tetrahedron which is the convex hull of vertices $\{v_1, v_2, v_3, v_4\}$ is a 3-dimensional simplex. Its vertices v_1, v_2, v_3, v_4 are its 0-dimensional faces. Its 6 edges, of which the thick line segment joining v_1 and v_2 is one, are its 1-dimensional faces. The dashed 2-dimensional simplex which is the convex hull of $\{v_1, v_2, v_3\}$ is one of the four 2-dimensional faces of the tetrahedron.

Triangulations

Let $\mathbf{K}$ be either $\mathbf{R}^n$ or a convex polyhedral subset of $\mathbf{R}^n$ of dimension n. A **triangulation** of $\mathbf{K}$ is a partition of $\mathbf{K}$ into simplexes satisfying the following properties

i) the simplexes cover $\mathbf{K}$,

ii) if two simplexes meet, their intersection is a common face,

iii) each point $x \in \mathbf{K}$ has a neighborhood meeting only a finite number of the simplexes,

iv) each $(n-1)$ dimensional simplex in the triangulation is the face of either two n-dimensional simplexes (in which case, the $(n-1)$ dimensional simplex is said to be an interior face in the triangulation) or exactly one n-dimensional simplex (in this case the $(n-1)$ dimensional simplex is said to be a boundary face in the triangulation),

v) for every point $x \in \mathbf{K}$ there exists a unique least dimension simplex, say σ, in the triangulation, containing x. If dimension of σ is $< n$, σ may be a face of several simplexes in the triangulation of dimension $>$ dimension of σ, and x is of course contained on the boundary of each of them. There exists a unique expression for x as a convex combination of vertices of σ, and this is the same expression for x as the convex combination of the vertices of any simplex in the triangulation containing x.

Example 2.22

In Figure 2.12 we give a triangulation of the unit cube in $\mathbf{R}^2$. The two 2-dimensional simplexes in this triangulation are the convex hulls of $\{v_0, v_1, v_2, \}$, $\{v_0, v_3, v_2\}$. The thick line segments in Figure 2.12 are the 1-dimensional simplexes in this triangulation which are the faces of exactly one two dimensional simplex. These 1-dimensional simplexes are the boundary faces in this triangulation. The thin diagonal line segment joining vertices v_0 and v_2 is the face of exactly two 2-dimensional simplexes, and hence is an interior face in this triangulation.

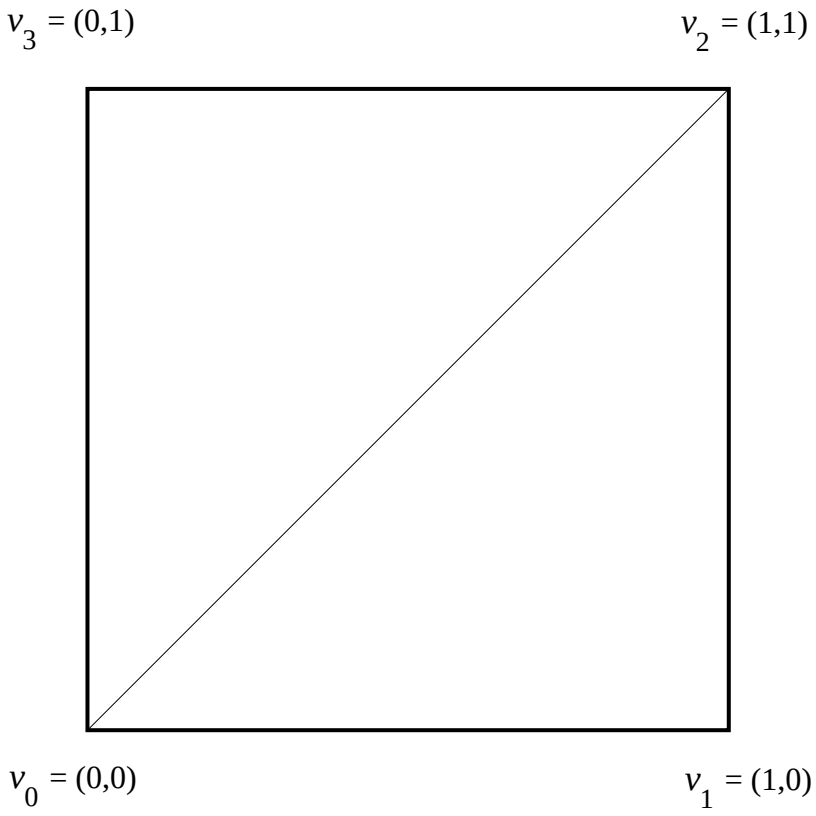


Figure 2.12 Triangulation K_1 of the unit square in $\mathbf{R}^2$.

Example 2.23

Consider the partition of the unit square in $\mathbf{R}^2$ into simplexes in Figure 2.13. It is not a triangulation since the two simplexes $\langle v_1, v_2, v_3 \rangle$ and $\langle v_3, v_4, v_5 \rangle$ intersect in $\langle v_3, v_5 \rangle$ which is a face of $\langle v_3, v_4, v_5 \rangle$ but not a face of $\langle v_1, v_2, v_3 \rangle$ (it is a proper subset of the face $\langle v_2, v_3 \rangle$ of $\langle v_1, v_2, v_3 \rangle$). So the partition of the unit square in $\mathbf{R}^2$ in Figure 2.13 into simplexes violates property (ii) given above, and is therefore not a triangulation.

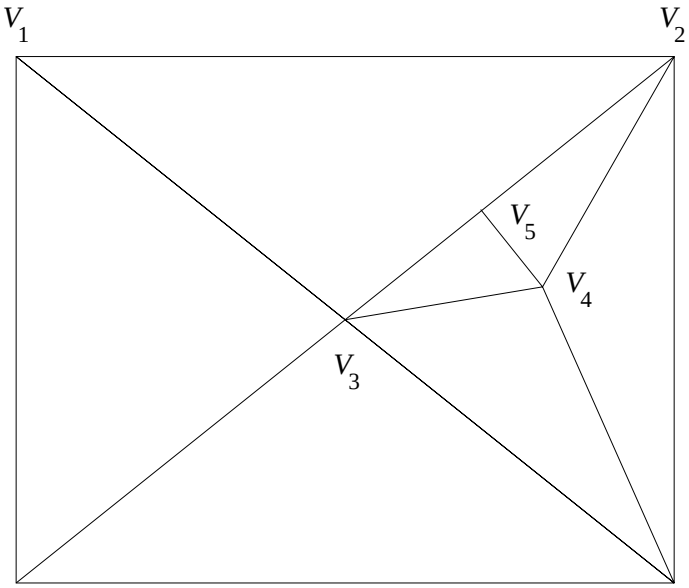


Figure 2.13 A partition of the unit cube in $\mathbf{R}^2$ into simplexes which is not a triangulation.

The triangulation for the unit cube in $\mathbf{R}^2$ given in Example 2.22 can be generalized to a triangulation of the unit cube in $\mathbf{R}^n$ which we call triangulation K_1, discussed by Freudenthal in 1942. The vertices of the simplexes in this triangulation are the same as the vertices of the unit cube. There are $n!$ n-dimensional simplexes in this triangulation. Let $v_0 = 0 \in \mathbf{R}^n$. Let $p = (p_1, \ldots, p_n)$ be any permutation of $\{1, \ldots, n\}$. Each of the $n!$ permutations p leads to an n-dimensional simplex in this triangulation. The n-dimensional simplex associated with the permutation p, denoted by (v_0, p), is $\langle v_0, v_1, \ldots, v_n \rangle$ where

$$v_i = v_{i-1} + I_{.p_i}, \quad i = 1 \text{ to } n \ . \tag{2.56}$$

In (2.56), I is the unit matrix of order n. For example, for $n = 2$, $p = (1, 2)$, we get the simplex $\langle v_0 = (0, 0)^T, v_1 = (1, 0)^T, v_2 = (1, 1)^T \rangle$. See Figure 2.12. See reference [2.72] for a proof that this does provide a triangulation of the unit cube of $\mathbf{R}^n$.

In this representation (v_0, p) for the simplex discussed above, v_0 is known as the **initial** or the 0^{th} **vertex** of this simplex. The other vertices of this simplex are obtained recursively as in (2.56). The vertex v_i is called the i^{th} **vertex** of this simplex for $i = 1$ to n.

This triangulation can be extended to provide a triangulation for the whole space $\mathbf{R}^n$ itself, which we call triangulation K_1 (it has been called by other symbols like K, I, etc., in other references) by first partitioning $\mathbf{R}^n$ into unit cubes using the integer points in $\mathbf{R}^n$, and then triangulating each unit cube as above. The vertices in this triangulation are all the points with integer coordinates in $\mathbf{R}^n$. Let $\bar{v}$ be any such vertex, and let $p = (p_1, \ldots, p_n)$ be any permutation of $\{1, \ldots, n\}$. Define $v_0 = \bar{v}$, and obtain v_i for $i = 1$ to n as in (2.56). Let $(\bar{v}, p)$ denote the simplex $\langle v_0, v_1, \ldots,$

$v_n\rangle$. The set of all such simplexes as $\bar{v}$ ranges over all points with integer coordinates in $\mathbf{R}^n$, and p ranges over all the permutations of $\{1, \ldots, n\}$ is the collection of all the n-dimensional simplexes in this triangulation K_1. Again see reference [2.72] for a proof that this is indeed a triangulation of $\mathbf{R}^n$. See Figure 2.14.

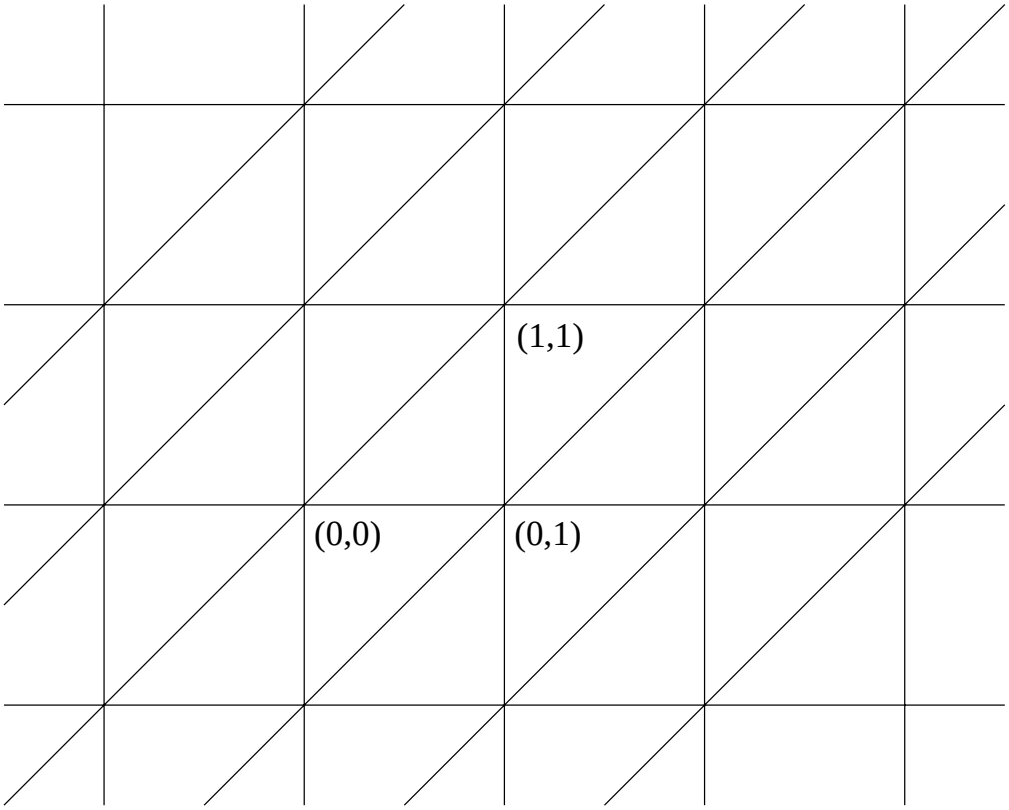


Figure 2.14 A partition of the unit cube in $\mathbf{R}^2$ into simplexes which is not a triangulation.

The **mesh** of a triangulation is defined to be the maximum Euclidean distance between any two points in a simplex in the triangulation. Clearly the mesh of triangulation K_1 of $\mathbf{R}^n$ is $\sqrt{n}$.

We can get versions of triangulation K_1 with smaller mesh by scaling the variables appropriately. Also the origin can be translated to any specified point. Let $x^0 \in \mathbf{R}^n$ be any specified point and δ a positive number. Let $\mathbf{J} = \{ x : x = (x_j) \in \mathbf{R}^n, x_j - x_j^0$ is an integer multiple of δ for all $j = 1$ to $n \}$. For any $v_0 \in \mathbf{J}$, and $p = (p_1, \ldots, p_n)$, a permutation of $\{1, \ldots, n\}$, define

$$v_i = v_{i-1} + \delta I_{.p_i}, \quad i = 1 \text{ to } n \ . \tag{2.57}$$

Let (v_0, p) denote the simplex $\langle v_0, v_1, \ldots, v_n\rangle$. The set $\mathbf{J}$ are the vertices, and the set of all simplexes (v_0, p) as v_0 ranges over $\mathbf{J}$ and p ranges over all the permutations of $\{1, 2, \ldots, n\}$ are the n-dimensional simplexes, in the triangulation of $\mathbf{R}^n$. We denote this triangulation by the symbol $\delta K_1(x^0)$. Its mesh is $\delta\sqrt{n}$.

How is the Triangulation used by the Algorithm ?

The algorithm traces a path. Each step in the path walks from one $(n-1)$-dimensional face of an n-dimensional simplex in the triangulation, to another $(n-1)$-dimensional face of the same simplex, and continues this way. See Figure 2.15. The path traced is unambiguous once it is started, and is similar to the one in the ghost story mentioned earlier, or the path traced by the complementary pivot method for the LCP. Computationally, the algorithm associates a column vector in $\mathbf{R}^n$ to each vertex in the triangulation. At each stage, the columns associated with the vertices of the current $(n-1)$-dimensional simplex form a basis, and the inverse of this basis is maintained. A step in the algorithm corresponds to the pivot step of entering the column associated with a new entering vertex into the basis. The path never returns to a simplex it has visited earlier.

To execute the path, one may consider it convenient to store all the simplexes in the triangulation explicitly. If this is necessary, the algorithm will not be practically useful. For practical efficiency the algorithm stores the simplexes using the mathematical formulae given above, which are easily programmed for the computer. The current simplex is always maintained by storing its 0^{th} vertex and the permutation corresponding to it. To proceed along the path efficiently, the algorithm provides very simple rules for termination once a desirable $(n-1)$-dimensional simplex in the triangulation is reached (this is clearly spelled out later on). If the termination condition is not satisfied, a mathematical formula provides the entering vertex. A minimum ratio procedure is then carried out to determine the dropping vertex, and another mathematical formula then provides the 0^{th} vertex and the permutation corresponding to the new simplex. All these procedures make it very convenient to implement this algorithm for the computer.

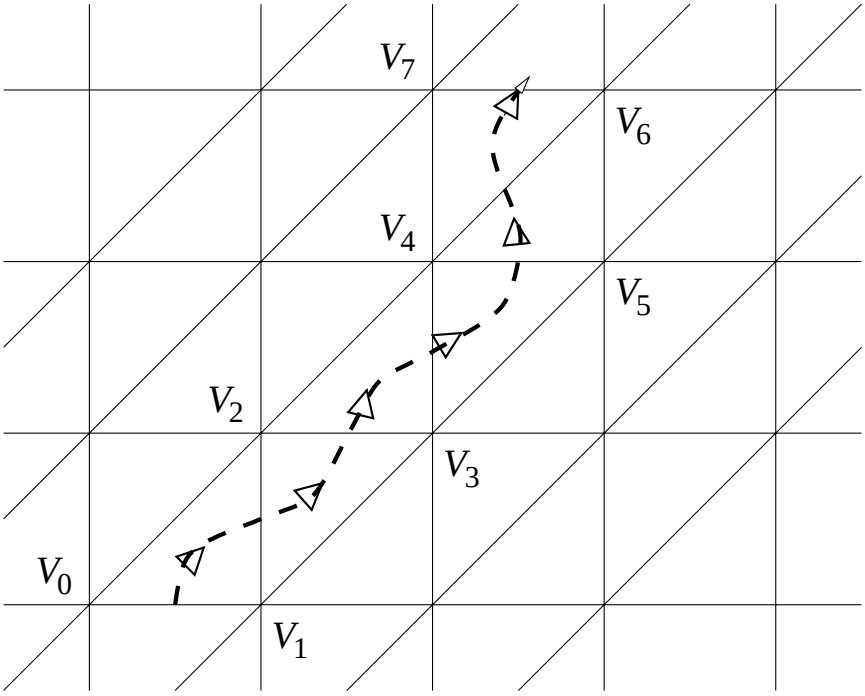


Figure 2.15 A path traced by the algorithm through the simplexes in the triangulation.

Special Triangulations of $\mathbf{R}^n \times [0,1]$

For computing a Kakutani fixed point of $\mathbf{F}(x)$ defined on $\mathbf{R}^n$, Merrill's algorithm uses a triangulation of $\mathbf{R}^n \times [0,1]$, which is a restriction of triangulation K_1 for $\mathbf{R}^{n+1}$ to this region, known as the special triangulation $\widetilde{K}_1$.

We will use the symbol $X = \begin{pmatrix} x \\ x_{n+1} \end{pmatrix}$ with $x \in \mathbf{R}^n$, to denote points in $\mathbf{R}^n \times [0,1]$. The set of vertices $\mathbf{J}$ in the special triangulation K_1 of $\mathbf{R}^n \times [0,1]$ are all the points $X = \begin{pmatrix} x \\ x_{n+1} \end{pmatrix} \in \mathbf{R}^{n+1}$ with x an integer vector in $\mathbf{R}^n$ and $x_{n+1} = 0$ or 1. The set of these vertices of the form $\begin{pmatrix} v \\ 1 \end{pmatrix}$ is denoted by $\mathbf{J}_1$, and the set of vertices of the form $\begin{pmatrix} v \\ 0 \end{pmatrix}$ is denoted by $\mathbf{J}_0$. $\mathbf{J} = \mathbf{J}_0 \cup \mathbf{J}_1$. The boundary of $\mathbf{R}^n \times [0,1]$ corresponding to $x_{n+1} = 1$ is known as the **top layer**, and the boundary corresponding to $x_{n+1} = 0$ is called the **bottom layer**. So $\mathbf{J}_1$, $\mathbf{J}_0$ are respectively the points with integer coordinates in the top and bottom layers. The $(n+1)$-dimensional simplexes in the special triangulation $\widetilde{K}_1$ of $\mathbf{R}^n \times [0,1]$ are those of the form (V_0, P) where $P = (p_1, \ldots, p_{n+1})$ is a permutation of $\{1, \ldots, n+1\}$ and $V_0 \in \mathbf{J}_0$, and $(V_0, P) = \langle V_0, V_1, \ldots, V_{n+1} \rangle$ where

$$V_i = V_{i-1} + I_{.p_i}, \quad i = 1 \text{ to } n+1 \ . \tag{2.58}$$

In (2.58), I is the unit matrix of order $n+1$. It can be verified that the set of all simplexes of the form (V_0, P) as V_0 ranges over $\mathbf{J}_0$, and P ranges over all permutations

of $\{1, 2, \ldots, n+1\}$ forms a triangulation of $\mathbf{R}^n \times [0,1]$. V_0 is the 0^{th} **vertex** and for $i = 1$ to $n+1$, the vertex V_i determined as in (2.58) in the i^{th} **vertex** of the $(n+1)$-dimensional simplex denoted by (V_0, P). The following properties should be noted.

Property 1 : In the representation (V, P) for an $(n+1)$-dimensional simplex in the special triangulation $\widetilde{K}_1$ of $\mathbf{R}^n \times [0,1]$, the 0^{th} vertex V is always an integer point in the bottom layer, that is, belongs to $\mathbf{J}_0$.

Property 2 : In the representation (V, P) for an $(n+1)$-dimensional simplex in the special triangulation $\widetilde{K}_1$ of $\mathbf{R}^n \times [0,1]$, there exists a positive integer r such that for all $i \leqq r-1$, the i^{th} vertex of (V, P) belongs to the bottom layer; and for all $i \geqq r$, the i^{th} vertex of (V, P) belongs to the top layer. The i here is the index satisfying the property that if the permutation $P = (p_1, \ldots, p_{n+1})$, then $p_i = n+1$. This property follows from the fact that the vertices of the simplex (V, P) are obtained by letting $V_0 = V$, and using (2.58) recursively.

Two $(n+1)$-dimensional simplexes in the special triangulation $\widetilde{K}_1$ are said to be adjacent, if they have a common n-dimensional simplex as a face (i. e., if $(n+1)$ of their vertices are the same). Merrill's algorithm generates a sequence of $(n+1)$-dimensional simplexes σ_1, σ_2, σ_3, ... of $\widetilde{K}_1$ in which every pair of consecutive simplexes are adjacent. So, given σ_j, σ_{j+1} is obtained by dropping a selected vertex V^- of σ_j and adding a new vertex V^+ in its place. The rules for obtaining σ_{j+1} given σ_j and V^- are called the **entering vertex choice rules** of the algorithm. These rules are very simple, they permit the generation of vertices as they are needed. We provide these rules here.

Let $\sigma_j = (V, P)$, where $P = (p_1, \ldots, p_{n+1})$ is a permutation of $\{1, \ldots, n+1\}$. The vertices of σ_j are $V_0 = V, V_1, \ldots, V_{n+1}$, as determined by (2.58). Let V^- be the dropping vertex. So V^- is V_i for some $i = 0$ to $n+1$. There are several cases possible which we consider separately.

Case 1 : $\{V_0, V_1, \ldots, V_{n+1}\} \setminus \{V^-\} \subset \mathbf{J}_1$. By property 2, this can only happen if $V^- = V_0$ and $V_1 \in \mathbf{J}_1$, that is, $p_1 = n+1$. The face of σ_j obtained by dropping the vertex V^-, is the n-dimensional simplex $\langle V_1, \ldots, V_{n+1} \rangle$ in the top layer, and hence is a boundary face. $\langle V_1, \ldots, V_{n+1} \rangle$ is the face of exactly one $(n+1)$ dimensional simplex in the triangulation $\widetilde{K}_1$, σ_j, and the algorithm terminates when this happens.

Case 2 : $\{V_0, V_1, \ldots, V_{n+1}\} \setminus \{V^-\} \subset \mathbf{J}_0$. By property 2, this implies that $V^- = V_{n+1}$ and $V_n \in \mathbf{J}_0$, that is $p_{n+1} = n+1$. We will show that this case cannot occur in the algorithm. So whenever $V^- = V_{n+1}$, we will have $p_{n+1} \neq n+1$ in the algorithm.

Case 3 : $\{V_0, V_1, \ldots, V_{n+1}\} \setminus \{V^-\}$ contains vertices on both the top and bottom layers. So the convex hull of $\{V_0, V_1, \ldots, V_{n+1}\} \setminus \{V^-\}$ is an n-dimensional simplex in the triangulation $\widetilde{K}_1$ which is an interior face, and hence is a face of exactly two $(n+1)$ dimensional simplexes in the triangulation, one is the present σ_j. The other σ_{j+1} is $(\widehat{V}, \widehat{P})$ as given below (V^+ given below is the new vertex in σ_{j+1} not in σ_j, it is the entering vertex that replaces the dropping vertex V^-).

V^-	V^+ (entering vertex)	$\widehat{V}$	$\widehat{P}$
$V^- = V_0$ $p_1 \neq n+1$ (see Case 1)†	$V_{n+1} + I_{.p_1}$	$V_0 + I_{.p_1}$	$(p_2, \ldots, p_{n+1}, p_1)$
$V^- = V_i$ $0 < i < n+1$	$V_{i-1} + I_{.p_{i+1}}$	V_0	$(p_1, \ldots, p_{i-1}, p_{i+1},$ $p_i, p_{i+2}, \ldots, p_{n+1})$
$V^- = V_{n+1}$ $p_{n+1} \neq n+1$ (see Case 2)*	$V_0 - I_{.p_{n+1}}$	$V_0 - I_{.p_{n+1}}$	$(p_{n+1}, p_1, \ldots, p_n)$

It can be verified that if $(\widehat{V}, \widehat{P})$ is defined as above, then $\widehat{V} \in \mathbf{J}_0$ (since $V \in \mathbf{J}_0$ where V is the 0^{th} vertex of σ_j) and so $(\widehat{V}, \widehat{P})$ is an $(n+1)$-dimensional simplex in the special triangulation, and that $(\widehat{V}, \widehat{P})$ and (V, P) share $(n+1)$ common vertices, so they are adjacent $(n+1)$ dimensional simplexes in this triangulation. See Figure 2.16 for an illlustration of the special triangulation $\widetilde{K}_1$ of $\mathbf{R}^1 \times [0,1]$.

The restriction of the special triangulation $\widetilde{K}_1$ of $\mathbf{R}^n \times [0,1]$ to either the top layer (given by $x_{n+1} = 1$) or the bottom layer (given by $x_{n+1} = 0$) in the same as the triangulation K_1 of $\mathbf{R}^n$. The mesh of the special triangulation $\widetilde{K}_1$ of $\mathbf{R}^n \times [0,1]$ is defined to be the mesh of the triangulation of $\mathbf{R}^n$ on either the top and bottom layer, and hence it is $\sqrt{n}$.

We can get special triangulation of $\mathbf{R}^n \times [0,1]$ of smaller mesh by scaling the variables in $\mathbf{R}^n$ appropriately. Also, the origin in the $\mathbf{R}^n$ part can be translated to any specified point in $\mathbf{R}^n$. Let $x^0 \in \mathbf{R}^n$ be a specified point and δ a positive number. Let $\mathbf{J}(x^0, \delta) = \left\{ \begin{pmatrix} x \\ x_{n+1} \end{pmatrix} : x = (x_j) \in \mathbf{R}^n, x_j - x_j^0 \text{ is an integer multiple of } \delta \text{ for each } j = 1 \text{ to } n, \ x_{n+1} = 0 \text{ or } 1 \right\}$. Then the points is $\mathbf{J}(x^0, \delta)$ are the vertices of the special triangulation of $\mathbf{R}^n \times [0,1]$ denoted by $\delta\widetilde{K}_1(x^0)$. $\mathbf{J}_0(x^0, \delta) = \left\{ X = \begin{pmatrix} x \\ x_{n+1} \end{pmatrix} : X \in \mathbf{J}(x^0, \delta), x_{n+1} = 0 \right\}$, $\mathbf{J}_1(x^0, \delta) = \left\{ \begin{pmatrix} x \\ x_{n+1} \end{pmatrix} : x \in \mathbf{J}(x^0, \delta), \ x_{n+1} = 1 \right\}$. For any $V \in \mathbf{J}_0(x^0, \delta)$, and $P = (p_1, \ldots, p_{n+1})$ a permutation of $\{1, \ldots, n+1\}$ define

$$\begin{aligned} V_0 &= V \\ V_i &= V_{i-1} + \delta I_{.p_i}, \quad i = 1 \text{ to } n+1 \end{aligned} \tag{2.59}$$

and let $(V, P) = \langle V_0, V_1, \ldots, V_{n+1} \rangle$. The set of all $(n+1)$ dimensional simplexes (V, P) given by (2.59) with $V \in \mathbf{J}_0(x^0, \delta)$ and P ranging over all the permutation of

† In this case, if $p_1 = n+1$, as discussed in Case 1 above, the algorithm terminates. So the algorithm continues only if $p_1 \neq n+1$ when this case occurs.

* In this case, we cannot have $p_{n+1} = n+1$, as discussed in Case 2 above. So, whenever this case occurs in the algorithm, we will have $p_{n+1} \neq n+1$.

$\{1, \ldots, n+1\}$ are the $(n+1)$-dimensional simplexes in the special triangulation $\delta \widetilde{K}_1(x^0)$ of $\mathbf{R}^n \times [0,1]$. Its mesh in $\delta\sqrt{n}$. In this triangulation, the vertex $\begin{pmatrix} x^0 \\ 0 \end{pmatrix}$ plays the same role as the origin $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ in the triangulation $\widetilde{K}_1$.

The Piecewise Linear Approximation and a Linear Approximate fixed Point of $F(x)$

Consider the special triangulation $\widetilde{K}_1$ of $\mathbf{R} \times [0,1]$ defined above, and let $\mathbf{J}_0$, $\mathbf{J}_1$ be the vertices in this triangulation on the bottom and top layers respectively. On the top layer, we define a a piecewise linear map $f(X)$ known as a piecewise linear approximation of $\mathbf{F}(x)$ relative to the present triangulation. For each $V = \begin{pmatrix} v \\ 1 \end{pmatrix} \in \mathbf{J}_1$ define $f(V) = \begin{pmatrix} f(v) \\ 1 \end{pmatrix}$, where $f(v) \in \mathbf{F}(v)$. The point $f(v)$ can be selected from the set $\mathbf{F}(v)$ arbitrarily, in fact it can be determined using the subroutine for finding a point from the set $\mathbf{F}(v)$, which was pointed out as a required input for this algorithm. Any nonvertex point $X = \begin{pmatrix} x \\ 1 \end{pmatrix}$ on the top layer must lie in an n-dimensional simplex in the triangulation on this layer. Suppose the vertices of this simplex are $V_i = \begin{pmatrix} v_i \\ 1 \end{pmatrix}$, $i = 1$ to $n+1$. Then x can be expressed as a convex combinations of $v_1, \ldots, v_{n+1}$ in a unique manner. Suppose this expression is $\alpha_1 v_1 + \ldots + \alpha_{n+1} v_{n+1}$ where $\alpha_1 + \ldots + \alpha_{n+1} = 1$, $\alpha_1, \ldots, \alpha_{n+1} \geqq 0$. Then define $f(x) = \alpha_1 f(v_1) + \ldots + \alpha_{n+1} f(v_{n+1})$. $f(x)$ is the piecewise linear approximation of $\mathbf{F}(x)$ defined on the top layer relative to the present triangulation. For $X = \begin{pmatrix} x \\ 1 \end{pmatrix}$ define $f(X) = \begin{pmatrix} f(x) \\ 1 \end{pmatrix}$. In each n-dimensional simplex in the top layer in this triangulation $f(x)$ is linear. So $f(x)$ is a well defined piecewise linear continuous function defined on the top layer. Remember that the definition of $f(x)$ depends on the choice of $f(v)$ from $\mathbf{F}(v)$ for $V = \begin{pmatrix} v \\ 1 \end{pmatrix} \in \mathbf{J}_1$.

The point $x \in \mathbf{R}^n$ is said to be a linear approximate fixed point of $\mathbf{F}(x)$ relative to the present piecewise linear approximation if

$$x = f(x) \ . \tag{2.60}$$

The n-dimensional simplex $\left\langle V_i = \begin{pmatrix} v_i \\ 1 \end{pmatrix} : i = 1 \text{ to } n+1 \right\rangle$ on the top layer contains a fixed point of the piecewise linear map $f(x)$ iff the system

$$\begin{array}{|cccc|c|} \hline \lambda_1 & & \ldots & \lambda_{n+1} & \\ \hline 1 & & \ldots & 1 & 1 \\ f(v_1) - v_1 & & \ldots & f(v_{n+1}) - v_{n+1} & 0 \\ \hline \multicolumn{5}{|l|}{\lambda_i \geqq 0, \quad i = 1 \text{ to } n+1} \\ \hline \end{array} \tag{2.61}$$

has a feasible solution. Thus the problem of finding a fixed point of the piecewise linear approximation $f(x)$ boils down to the problem of finding an n-dimensional simplex on the top layer whose vertices are such that (2.61) is feasible.

For each vertex $V = \begin{pmatrix} v \\ 1 \end{pmatrix}$ in the top layer associate the column vector $\begin{pmatrix} 1 \\ f(v) - v \end{pmatrix} \in \mathbf{R}^{n+1}$, which we denote by $A_{.V}$ and call the label of the vertex V. The coefficient matrix in (2.61) whose columns are the labels of the vertices of the simplex is called the **label matrix** corresponding to the simplex. Because of the nature of the labels used on the vertices, this is called a **vector labelling method**.

An n-dimensional simplex on the top layer is said to be a **completely labelled simplex** if the system (2.61) corresponding to it has a nonnegative solution, that is, if it contains a fixed point of the current piecewise linear approximation.

Let $V_0 = \begin{pmatrix} v_0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, $P = (1, \ldots, n+1)$ and let $(V_0, P) = \langle V_0, V_1, \ldots, V_{n+1} \rangle$, where $V_i = \begin{pmatrix} v_i \\ 0 \end{pmatrix}$, $i = 1$ to n. Then $\langle V_0, V_1, \ldots, V_n \rangle$ is the n-dimensional face of (V_0, P) in the bottom layer. Let $W = \begin{pmatrix} w \\ o \end{pmatrix}$ be an arbitrary point in the interior of this n-dimensional simplex $\langle V_0, \ldots, V_n \rangle$, for example, $w = \frac{(v_0 + \ldots + v_n)}{(n+1)}$. For every vertex $V = \begin{pmatrix} v \\ 0 \end{pmatrix} \in \mathbf{J}_0$ in the bottom layer, define $f(V) = \begin{pmatrix} f(v) \\ 0 \end{pmatrix} = \begin{pmatrix} w \\ 0 \end{pmatrix}$. For any nonvertex X in $\mathbf{R}^n \times [0, 1]$, X must lie in some $(n+1)$-dimensional simplex in the present triangulation, say $\langle V_0^1, V_1^1, \ldots, V_{n+1}^1 \rangle$. So there exist unique numbers $\alpha_0, \ldots, \alpha_{n+1} \geqq 0$ such that $\alpha_0 + \alpha_1 + \ldots + \alpha_{n+1} = 1$, $X = \alpha_0 V_0^1 + \alpha_1 V_1^1 + \ldots + \alpha_{n+1} V_{n+1}^1$. Then define $f(X) = \alpha_0 f(V_0^1) + \ldots + \alpha_{n+1} f(V_{n+1}^1)$. The map $f(X)$ is thus a continuous piecewise linear map defined on $\mathbf{R}^n \times [0, 1]$. In each $(n+1)$ dimensional simplex in the present triangulation, $f(X)$ is linear. Also, under this map, every point in the bottom layer maps into the point W. Define the label of any vertex $V = \begin{pmatrix} v \\ 0 \end{pmatrix} \in J_0$ to be the column vector $A_{.V} = \begin{pmatrix} 1 \\ w - v \end{pmatrix} \in \mathbf{R}^{n+1}$.

Let $\langle V_0, V_1, \ldots, V_n \rangle$ be the n-dimensional simplex in the bottom layer, from the interior of which we selected the point W. Since W is in the interior of this simplex, B_1, the $(n+1) \times (n+1)$ label matrix corresponding to this simplex is nonsingular. Let $b = (1, 0, 0, \ldots, 0)^T \in \mathbf{R}^{n+1}$. Then the system corresponding to (2.61) for this simplex is

$$\begin{array}{|c|c|} \hline \lambda & \\ \hline B_1 & b \\ \hline \end{array} \qquad \lambda \geqq 0 \tag{2.62}$$

This system has the unique positive solution $\lambda = \bar{b} = B_1^{-1} b > 0$, since W is in the interior of this simplex. Incidentally, this $\langle V_0, V_1, \ldots, V_n \rangle$ is the only n-dimensional simplex in the bottom layer whose label matrix leads to a nonnegative solution to the system like (2.61). The reason for it is that since W is in the interior of $\langle V_0, V_1, \ldots,$

$V_n\rangle$, W is not contained in any other simplex in the triangulation in the bottom layer. Also, since $\bar{b} > 0$, the $n \times (n+1)$ matrix $\left(\bar{b} \ \vdots \ B_1^{-1}\right)$ has all rows lexicopositive. The inverse tableau corresponding to the initial system (2.62) is

basic vector	basis inverse	
λ	B_1^{-1}	$\bar{b}$

(2.63)

The initial simplex $\langle V_0, V_1, \ldots, V_n\rangle$ in the bottom layer is an n-dimensional face of the unique $(n+1)$-dimensional simplex $\langle V_0, V_1, \ldots, V_n, V_{n+1}\rangle$ in the present triangulation $\widetilde{K}_1$. Introduce a new variable, say λ_{n+1}, in (2.62) with its column vector equal to the label of this new vertex V_{n+1}, and bring this variable into the present basic vector. The pivot column for this pivot operation is $B_1^{-1}A_{.V_{n+1}}$. If this pivot column is nonpositive, it would imply that the set of feasible solutions of this augmented system (2.62) with this new variable is unbounded, which is impossible since the first constraint in the system says that the sum of all the variables is 1, and all the variables are nonnegative. So, the pivot column contains at least one positive entry, and it is possible to bring the new variable into the present basic vector. The dropping variable is determined by the usual lexico minimum ratio test of the primal simplex algorithm, this always determines the dropping variable uniquely and unambiguously and maintains the system lexico feasible. If the label of V_i is the dropping column, the next basis is the label matrix of the n-dimensional simplex $\langle V_0, \ldots, V_{i-1}, V_{i+1}, \ldots, V_{n+1}\rangle$. The inverse tableau corresponding to this new basis is obtained by entering the pivot column by the side of the present inverse tableau in (2.63) and performing a pivot step in it, with the row in which the dropping variable λ_i is basic, as the pivot row.

By the properties of the triangulation, the new n-dimensional simplex $\langle V_0, \ldots, V_{i-1}, V_{i+1}, \ldots, V_{n+1}\rangle$ is the face of exactly one or two $(n+1)$ dimensional simplexes in the triangulation. One is the simplex $\langle V_0, \ldots V_{n+1}\rangle$. If there is another, it must be a simplex of the form $\langle Y, V_0, \ldots, V_{i-1}, V_{i+1}, \ldots, V_{n+1}\rangle$. Then bring the column $A_{.Y}$ into the basis next. Continuing in this manner, we generate a unique path of the form S_1^n, S_1^{n+1}, S_2^n, $S_2^{n+1}, \ldots$. Here S_k^n, S_k^{n+1} represent the k^{th} n-dimensional simplex and $(n+1)$-dimensional simplex respectively in this path. Termination can only occur if at some stage the basis corresponds to an n-dimensional simplex S_r^n all of whose vertices are on the top layer. Each n-dimensional simplex in this path is the face of at most two $(n+1)$-dimensional simplexes, we arrive at this face through one of these $(n+1)$-dimensional simplexes, and leave it through the other. The initial n-dimensional simplex in the bottom layer is a boundary face, and hence is the face of a unique $(n+1)$-dimensional simplex in the triangulation. So the path continues in a unique manner and it cannot return to the initial n-dimensional simplex again. Also, since the initial n-dimensional simplex is the only n-dimensional simplex in the bottom layer for which the system corresponding to (2.61) is feasible, the path will never pass through any other n-dimensional simplex in the bottom layer after the first step. Any n-dimensional simplex obtained on the path whose vertices belong to both the bottom and top layers is an interior face, so it is incident to two $(n+1)$-dimensional simplexes,

we arrive at this n-face through one of these $(n+1)$-dimensional simplexes and leave it through the other, and the algorithm continues. The reader can verify that the properties of the path generated are very similar to the almost complementary basic vector path traced by the complementary pivot algorithm for the LCP. Thus we see that the path continues uniquely and unambiguously and it can only terminate when the columns of the current basis are the labels of vertices all of whom belong to the top layer. When it terminates, from the final BFS we get a fixed point of the current piecewise linear approximation.

Example 2.24

Consider $n = 1$. We consider a single-valued map from $\mathbf{R}^1$ to $\mathbf{R}^1$, $\mathbf{F}(x) = \{x^2 - 5x + 9\}$, $x \in \mathbf{R}^1$. The special triangulation of $\mathbf{R}^1 \times [0,1]$ is given in Figure 2.16.

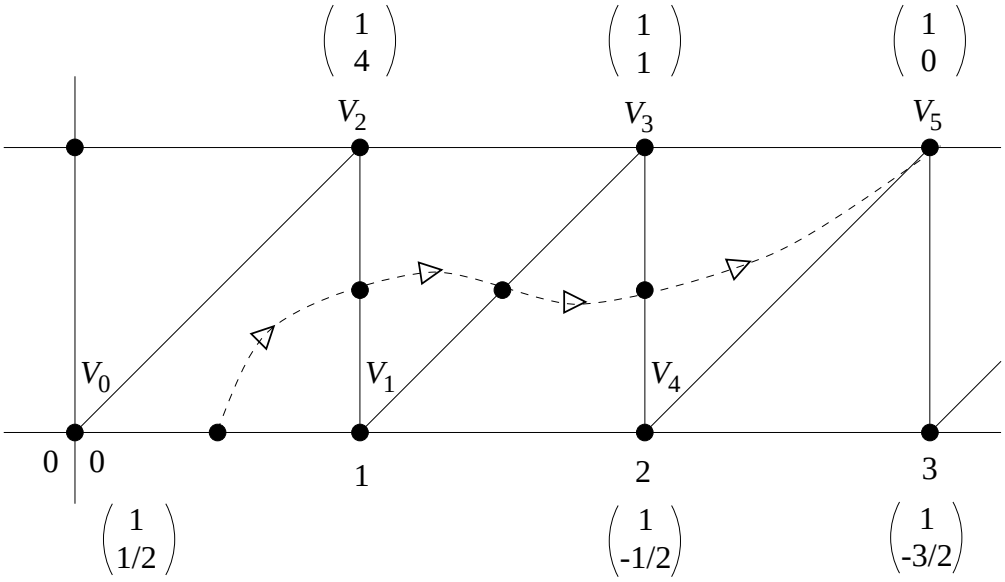


Figure 2.16 The column vector by the side of a vertex is its vector label.

The vertices for the triangulation are all the points with integer coordinates in $\mathbf{R}^1 \times [0,1]$. For each $V = \begin{pmatrix} v \\ 1 \end{pmatrix}$ on the top layer with v integer, we define $f(v) = v^2 - 5v + 9$. We take the initial 1-dimensional simplex on the bottom layer to be $\langle V_0, V_1 \rangle$ and the point W to be the interior point $(w,0)^T = \left(\frac{1}{2}, 0\right)^T$ in it. For each $V = \begin{pmatrix} v \\ 0 \end{pmatrix}$ in the bottom layer, define $f(V) = W = \left(\frac{1}{2}, 0\right)^T$. The label of the vetex $V = \begin{pmatrix} v \\ x_{n+1} \end{pmatrix}$ is $\begin{pmatrix} 1 \\ f(v) - v \end{pmatrix}$ if $x_{n+1} = 1$, or $\begin{pmatrix} 1 \\ w - v \end{pmatrix}$ if $x_{n+1} = 0$. The labels of some of the vertices are entered in Figure 2.16. The initial system corresponding to (2.62) here is

λ_0	λ_1	
1	1	1
$\frac{1}{2}$	$-\frac{1}{2}$	0
λ_0, λ_1	$\geqq$	0

The feasible solution of this system and the basis inverse are given below.

Basic variable	Basis Inverse		$\bar{b}$	Pivot Column $\bar{A}_{.V_2}$	Ratios
λ_0	$\frac{1}{2}$	1	$\frac{1}{2}$	$\boxed{\frac{9}{2}}$	$\frac{1}{2}/\frac{9}{2}$ *Min.*
λ_1	$\frac{1}{2}$	-1	$\frac{1}{2}$	$-\frac{7}{2}$	

The initial simplex $\langle V_0, V_1 \rangle$ is the face of the unique 2-dimensional simplex $\langle V_0, V_1, V_2 \rangle$ in the triangulation. So we associate the label of V_2 with a variable λ_2 and bring it into the basic vector. The pivot column is

$$\begin{pmatrix} \frac{1}{2} & 1 \\ \frac{1}{2} & -1 \end{pmatrix} \quad \begin{pmatrix} 1 \\ 4 \end{pmatrix} \quad = \quad \begin{pmatrix} \frac{9}{2} \\ -\frac{7}{2} \end{pmatrix} \quad = \quad \bar{A}_{.V_2}$$

and this is entered on the inverse tableau. The dropping variable is λ_0 and the pivot element is inside a box. Pivoting leads to the next inverse tableau. For ease in understanding, the vertices are numbered as V_i, $i = 0, 1, \ldots$ in Figure 2.16 and we will denote the variable in the system associated with the label of the vertex V_i by λ_i.

Basic variable	Basis Inverse		$\bar{b}$	Pivot Column $\bar{A}_{.V_3}$	Ratios
λ_2	$\frac{1}{9}$	$\frac{2}{9}$	$\frac{1}{9}$	$\boxed{\frac{3}{9}}$	$\frac{1}{3}$ *Min.*
λ_1	$\frac{8}{9}$	$-\frac{2}{9}$	$\frac{8}{9}$	$\frac{6}{9}$	$\frac{8}{6}$

The current 1-simplex $\langle V_2, V_1 \rangle$ is the face of $\langle V_0, V_1, V_2 \rangle$ and $\langle V_3, V_1, V_2 \rangle$. We came to the present basic vector through $\langle V_0, V_1, V_2 \rangle$, so we have to leave $\langle V_2, V_1 \rangle$ through the 2-simplex $\langle V_3, V_1, V_2 \rangle$. Hence the updated column of the label of V_3, $\bar{A}_{.V_3}$, is the entering column. It is already entered on the inverse tableau. The dropping variable is λ_2. Continuing, we get the following

Basic variable	Basis Inverse		$\bar{b}$	pivot Column $\bar{A}_{.V_4}$	Ratios
λ_3	$\frac{1}{3}$	$\frac{2}{3}$	$\frac{1}{3}$	$-\frac{2}{3}$	
λ_1	$\frac{2}{3}$	$-\frac{2}{3}$	$\frac{2}{3}$	$\boxed{\frac{5}{3}}$	$\frac{2}{5}$ *Min.*
				$\bar{A}_{.V_5}$	
λ_3	$\frac{3}{5}$	$\frac{2}{5}$	$\frac{3}{5}$	$\frac{3}{5}$	1
λ_4	$\frac{2}{5}$	$-\frac{2}{5}$	$\frac{2}{5}$	$\boxed{\frac{2}{5}}$	1
λ_3	0	1	0		
λ_5	1	-1	1		

In the basic vector λ_3, λ_4, there is a tie for the dropping variable by the usual primal simplex minimum ratio test, and hence the lexico minimum ratio test was used in determining the dropping variable. The algorithm terminates with the basic vector (λ_3, λ_5) since the corresponding vertices V_3, V_5 are both in the top layer. The fixed point of the piecewise linear approximation is $0 \times v_3 + 1 \times v_5 = 0 \times 2 + 1 \times 3 = 3$, from the terminal BFS. It can be verified that $x = 3$ is indeed a fixed point of $\mathbf{F}(x)$, since $\mathbf{F}(3) = \{3\}$.

Sufficient Conditions for Finite Termination with a Linear Approximate Fixed Point

Once the triangulation of $\mathbf{R}^n \times [0, 1]$ and the piecewise linear approximation are given, the path generated by this algorithm either terminates with an n-dimensional simplex on the top layer (leading to a fixed point of the present piecewise linear approximation) after a finite number of pivot steps, or continues indefinitely. Sufficient conditions to guarantee that the path terminates after a finite number of steps are discussed in [2.58], where the following theorem is proved.

Theorem 2.15 *Given $\hat{x} \in \mathbf{R}^n$ and $\alpha > 0$ let $\mathbf{B}(\hat{x}, \alpha) = \{ x : x \in \mathbf{R}^n$ satisfying $\|x - \hat{x}\| \leqq \alpha \}$. Suppose there are fixed positive numbers ν and γ and a point $\bar{x} \in \mathbf{R}^n$ satisfying: for each $x \in \mathbf{B}(\bar{x}, \nu)$, $y \in \mathbf{B}(x, \gamma) \backslash \mathbf{B}(\bar{x}, \nu)$ and $u \in \mathbf{F}(x)$, $(u-x)^T(y-\bar{x}) < 0$. Let x^0 by an arbitrary point in $\mathbf{R}^n$. If the above algorithm is executed using the starting point $x \in \{x^0\} \cup \mathbf{B}(\bar{x}, \nu + \gamma)$ and a special triangulation $\widetilde{K}_1$ with its mesh $\leqq \gamma$, then, the algorithm terminates in a finite number of steps with a linear approximate fixed point of $\mathbf{F}(x)$. Also, every linear approximate fixed point lies in $\mathbf{B}(\bar{x}, \nu + \gamma)$.*

We refer the reader to O. H. Merril's Ph. D. thesis [2.58] for a proof of this theorem. But it is very hard to verify whether these conditions hold in practical applications. In practical applications we apply the algorithm and let the path continue until some

prescribed upper bound on computer time is used up. If termination does not occur by then, one usually stops with the conclusion that the method has failed on that problem.

One strange feature of the sufficient conditions to guarantee finite termination of the above algorithm is the following. Let $f(x) = (f_1(x), \ldots, f_n(x))^T$ be a continuously differentiable function from $\mathbf{R}^n$ into $\mathbf{R}^n$, and suppose we are applying the algorithm discussed above, on the fixed point formulation for the problem of solving the system of equations "$f(x) = 0$". Solving the system "$f(x) = 0$" is equivalent to finding the Kakutani fixed point of either $\mathbf{F}_1(x) = \{f(x) + x\}$ or $\mathbf{F}_2(x) = \{-f(x) + x\}$. Mathematically, the problem of finding a fixed point of $\mathbf{F}_1(x)$ or $\mathbf{F}_2(x)$ are equivalent. However, if $\mathbf{F}_1(x)$ satisfies the sufficiency condition for finite termination, $\mathbf{F}_2(x)$ will not. Thus, if the algorithm is applied to find the fixed points of $\mathbf{F}_1(x)$, and $\mathbf{F}_2(x)$; the behavior of the algorithm on the two problems could be very different. On one of them the algorithm may have finite termination, and on the other it may never terminate. This point should be carefully noted in using this algorithm in practical applications.

Algorithm to generate an Approximate Fixed Point of $\mathbf{F}(x)$

Select a sequence of positive numbers $\delta_0 = 1, \delta_1, \delta_2, \ldots$ converging to zero. Let $x^0 = 0$. Set $t = 0$ and go to Step 1.

Step 1 : Define the piecewise linear approximation for $\mathbf{F}(x)$ relative to the special triangulation $\delta_t \widetilde{K}_1(x^t)$ choosing the point W from the interior of the translate of the n-dimensonal face of the initial simplex $\langle 0, I_{.1}, \ldots, I_{.n} \rangle$ on the bottom layer in this triangulation. Find a fixed point of this piecewise linear approximation using this special triangulation by the algorithm discussed above. Suppose the fixed point obtained is x^{t+1}. x^{t+1} is a linear approximate fixed point of $\mathbf{F}(x)$ relative to this special triangulation $\delta_t \widetilde{K}_1(x^t)$. If $x^{t+1} \in \mathbf{F}(x^{t+1})$, terminate, x^{t+1} is a fixed point of $\mathbf{F}(x)$. Otherwise go to Step 2.

Step 2 : Replace t by $t + 1$ and do Step 1.

So this method generates the sequence $\{x^1, x^2, x^3, \ldots\}$ of linear approximate fixed points for $\mathbf{F}(x)$. If at any stage $x^t \in \mathbf{F}(x^t)$, it is a fixed point of $\mathbf{F}(x)$ and we terminate. Otherwise, any limit point of the sequence $\{x^t : t = 1, 2, \ldots\}$ can be shown to be a fixed point of $\mathbf{F}(x)$. In practice, if finite termination does not occur, we continue until δ_t becomes sufficiently small and take the final x^t as an approximate fixed point of $\mathbf{F}(x)$.

To find Fixed Points of USC Maps Defined on a Compact Convex Subset $\mathbf{\Gamma} \subset \mathbf{R}^n$

Without any loss of generality we can assume that $\mathbf{\Gamma}$ has a nonempty interior (if the interior of $\mathbf{\Gamma}$ in $\mathbf{R}^n$ is $\emptyset$, the problem is not altered by replacing $\mathbf{R}^n$ by the affine hull of $\mathbf{\Gamma}$, in which $\mathbf{\Gamma}$ has a nonempty interior). Let $\mathbf{F}(x)$ be the given USC map. So $\mathbf{F}(x)$

is defined for all $x \in \Gamma$, and for all such x, $\mathbf{F}(x)$ is a compact convex subset of Γ. Since this map is only defined on Γ, and not on the whole of $\mathbf{R}^n$, the algorithm discussed above does not apply to this problem directly. However, as pointed out by B. C. Eaves [2.44], we can extend the definition of $\mathbf{F}(x)$ to the whole of $\mathbf{R}^n$ as below. Let c be any point from the interior of Γ.

$$\mathbf{F}^1(x) = \begin{cases} \{c\}, & \text{if } x \notin \Gamma \\ \text{convex hull of } \{c, \mathbf{F}(x)\}, & \text{if } x \in \text{boundary of } \Gamma \\ \mathbf{F}(x), & \text{if } x \in \text{interior of } \Gamma. \end{cases}$$

It can be verified that $\mathbf{F}^1(x)$ is now a USC map defined on $\mathbf{R}^n$, and that every fixed point of $\mathbf{F}^1(x)$ is in Γ and is also a fixed point of $\mathbf{F}(x)$ and vice versa. Since $\mathbf{F}^1(x)$ is defined over all of $\mathbf{R}^n$, the method discussed above can be applied to find a fixed point of it.

Homotopy Interpretation

In the algorithm discussed above for computing a fixed point of the piecewise linear approximation, there are two layers, the bottom layer and the top layer. We have the same triangulation of $\mathbf{R}^n$ in both the bottom and top layers. The labels for the vertices on the bottom layer are artificial labels corresponding to a very simple map for which we know the fixed point. The labels for the vertices on the top layer are natural labels corresponding to the piecewise linear map whose fixed point we want to find. The algorithm starts at the known fixed point of the artificial map of the bottom layer and walks its way through the triangulation until it reaches a fixed point of the piecewise linear map on the top layer. This makes it possible to interpret the above algorithm as a homotopy algorithm. Other homotopy algorithms for computing fixed points with continuous refinement of the grid size have been developed by B. C. Eaves [2.44] and B. C. Eaves and R. Saigal [2.47] and several others [2.40 to 2.80].

Comments 2.2 H. Scarf [2.68] first pointed out that the basic properties of the path followed by the complementary pivot algorithm in the LCP can be used to compute approximate Brouwer's fixed points using partitions of the space into sets called primitive sets, and T. Hansen and H. Scarf [2.69] extended this into a method for approximating Kakutani fixed points. The earliest algorithms for computing approximate fixed points using triangulations are those by B. C. Eaves [2.44], H. W. Kuhn [2.54]. These early algorithms suffered from computational inefficiency because they start from outside the region of interest. The first method to circumvent this difficulty is due to O. H. Merrill [2.57, 2.58] discussed above. The applications of fixed point methods in nonlinear programming discussed in Sections 2.7.3, 2.7.4, 2.7.5, 2.7.6 and 2.7.7 are due to O. H. Merrill [2.58]. Besides the triangulation K_1 discussed above, Merrill's algorithm can be implemented using other triangulations, see M. J. Todds book [2.72] and the papers [2.40 to 2.80].

2.8 COMPUTATIONAL COMPLEXITY OF THE COMPLEMENTARY PIVOT ALGORITHM

The **computational complexity** of an algorithm measures the growth of the computational effort involved in executing the algorithm as a function of the size of the problem. In the complementary pivot algorithm, we will assess the computational effort by the number of pivot steps carried out before the algorithm terminates. There are three commonly used measures for studying the computational complexity of an algorithm. These are discussed below.

Worst Case Computational Complexity

This measure is a tight mathematical upper bound on the number of pivot steps required before termination, as a function of the size of the problem. In studying the worst case computational complexity we will assume that the data is integer, or more generally, rational, that is, each m_{ij}, q_i in the matrices q, M is a ratio of two integers. In this case by multiplying all the data by a suitable positive integer, we can transform the problem into an LCP in which all the data is integer. Hence without any loss of generality we assume that all the data is integer, and define the size of the problem to be the total number of bits of storage needed to store all the data in the problem in binary form. See Chapter 6 where a mathematical definition of this size is given. The worst case computational complexity of an algorithm provides a guaranteed upper limit on the computational effort needed to solve any instance of the problem by the algorithm, as a function of the size of the instance. The algorithm is said to be **polynomially bounded** if this worst case computational complexity is bounded above by a polynomial of fixed degree in the size of the problem, that is, if there exist constants α, r independent of the size, such that the computational effort needed is always $\leq \alpha s^{\mathbf{r}}$ when the algorithm is applied on problems of size s. Even though the worst case computational complexity is measured in terms of the number of pivot steps, each pivot step needs $\mathcal{O}(n^{\mathbf{2}})$ basic arithmetical operations (addition, multiplication, division, comparison) on data each of which has at most s digits, where s is the size and n the order of the instance; so if the algorithm is polynomially bounded in terms of the number of pivot steps, it is polynomially bounded in terms of the basic arithmetical operations. In Chapter 6 we conclusively establish that the complementary pivot algorithm is not a polynomially bounded algorithm in this worst case sense. Using our examples discussed in Chapter 6, in [2.74] M. J. Todd constructed examples of square nonsingular systems of linear equations "$Ax - b = 0$", with integer data, for solving which the computational effort required by Merrill's algorithm of Section 2.7.8, grows exponentially with the size of the problem.

An algorithm may have a worst case computational complexity which is an exponentially growing function of the size of the problem, just because it performs very poorly on problem instances with a very rare pathological structure. Such an algorithm

might be extremely efficient on instances of the problem not having the rare pathological structure, which may never show up in practical applications. For this reason, the worst case measure is usually very poor in judging the computational efficiency of an algorithm, or its practical utility.

The Probabilistic Average Computational Complexity

Here we assume that the data in the problem is randomly generated according to some assumed probability distribution. The average computational complexity of the algorithm under this model is then defined to be the statistical expectation of the number of steps needed by the algorithm before termination, on problem instances with this data. Since the expectation is a multiple integral, this average analysis requires techniques for bounding the values of multiple integrals. If the probability distributions are continuous distributions, the data generated will in general be real numbers (not rational), and so in this case we define the size of the LCP to be its order n. We assume that each pivot step in the algorithm is carried out on the real data using exact arithmetic, but assess the computational complexity by the average number of pivot steps carried out by the algorithm before termination.

M. J. Todd performed the average analysis in [2.36] under the folowing assumptions on the distribution of the data (q, M).

i) With probability one, every square submatrix of M whose sets of row indices and column indices differ in at most one element, is nonsingular.

ii) q is nondegenerate in the LCP (q, M).

iii) The distributions of (q, M) are sign-invariant; that is, (q, M) and (Sq, SMS) have identical distributions for all sign matrices S (i. e., diagonal matrices with diagonal entries of $+1$ or -1).

Under these assumptions he showed that the expected number of pivot steps taken by the lexicographic Lemke algorithm (see Section 2.3.4) before termination when applied on the LCP (q, M) is at most $\frac{n(n+1)}{4}$.

M. J. Todd [2.36] also analysed the average computational complexity of the lexicographic Lemke algorithm applied on the LCP corresponding to the LP

$$\begin{aligned} &\text{minimize} && cx \\ &\text{subject to} && Ax \geqq b \\ & && x \geqq 0 \end{aligned}$$

under the following assumptions. A is a matrix of order $m \times N$. The probability distribution generating the data (A, b, c) and hence the data (q, M) in the corresponding LCP satisfies the following assumptions :

i) with probability one, the LP and its dual are nondegenerate (every solution of $Ax - u = b$ has at least m nonzero variables, and every solution of $yA + v = c$ has at least N nonzero variables), and every square submatrix of A is nonsingular.

ii) the distributions of (A, b, c) and of (S_1AS_2, S_1b, S_2c) are identical for all sign matrices S_1, S_2 of appropriate dimension). This is the sign invariance requirement.

Under these assumptions he showed that the expected number of pivot steps taken by the lexicographic Lemke algorithm when applied on the LCP corresponding to this LP is at most, minimum $\left\{\frac{m^2+5m+11}{2}, \frac{2N^2+5N+5}{2}\right\}$. See also [2.31] for similar results under slightly different probabilistic models.

In a recent paper, [8.20] R. Saigal showed that the expected number of pivot steps taken by the lexicographic Lemke algorithm when applied on the LCP corresponding to the above LP is actually bounded above by m and asymptotically approaches $\frac{m}{2}-1$, where m is the number of rows in A.

Unfortunately, these nice quadratic or linear bound expected complexity results seem very dependent on the exact manner in which the algorithm is implemented, and on the problabilistic model of the data. For example, it has not been possible so far to obtain comparable results for the complementary pivot algorithm of Section 2.2 which uses the column vector e of all 1's as the original column vector of the artificial variable z_0.

Empirical Average Computation Complexity

This measure of computational complexity is used more in the spirit of simulation. Here, a computational experiment is usually performed by applying the algorithm on a large number of problem instances of various sizes, and summary statistics are then prepared on how the algorithm performed on them. The data is usually generated according to some distribution (typically we may assume that each data element is a uniformly distributed random variable from an interval such as -100 to $+100$, etc.). In the LCP, we may also want to test how the complementary pivot algorithm performs under varying degrees of sparsity of q and M. For this, a certain percentage of randomly chosen entries in q and M can be fixed as zero, and the remaining obtained randomly as described above. It may also be possible to generate M so that it has special properties. As an example, if we want to experiment on LCPs associated with PSD symmetric matrices, we can generate a random square matrix A as above and take M to be A^TA. Such computational experiments can be very useful in practice. The experiments conducted on the complementary pivot algorithm, suggest that the empirical average number of pivot steps before termination grows linearly with n, the order of the LCP.

We know that Merrill's simplicial method for computing the fixed point of a piecewise linear map discussed in Section 2.7.8 may not terminate on some problems. Computational experiments indicate that on problems on which it did terminate, the average number of simplices that the algorithm walked through before termination, is $\mathcal{O}(n^2)$, as a function of the dimension of the problem. See [2.62 to 2.67].

2.9 THE GENERAL QUADRATIC PROGRAMMING PROBLEM

From the results in Section 2.3 we know that the complementary pivot method processes convex quadratic programs with a finite computational effort. Here we discuss the general, possibly nonconvex, quadratic programming problem. This is a problem in which a general quadratic objective function is to be minimized subject to linear constraints.

The Reduction Process

If there is an equality constraint on the variables, using it, obtain an expression for one of the variables as an affine function of the others, and eliminate this variable and this constraint from the optimization portion of the problem. A step like this is called a **reduction step**, it reduces the number of variables in the optimization problem by one, and the number of constraints by one. In the resulting problem, if there is another equality constraint, do a reduction step using it, and continue in the same manner. When this work is completed, only inequality constraints remain, and the system of constraints assumes the form $FX \geqq f$, which includes any sign restrictions and lower or upper bound constraints on the variables. We assume that this system is feasible. An inequality constraint in this system is said to be a **binding inequality constraint** if it holds as an equation at all feasible solutions. A binding inequality constraint can therefore be treated as an equality constraint without affecting the set of feasible solutions. Binding inequality constraints can be identified using a linear programming formulation. Introduce the vector of slack variables v and transform the system of constraints into $FX - v = f$, $v \geqq 0$. The i^{th} constraint in the system, $F_i.X \geqq f_i$, is a binding constraint iff the maximum value of v_i subject to $FX - v = f$, $v \geqq 0$, is zero. Using this procedure identify all the binding constraints, change each of them into an equality constraint in the system. Carry out further reduction steps using these equality constraints. At the end, the optimization portion of the problem reduces to one of the following form

$$\begin{array}{ll} \text{Minimize} & \theta(x) = cx + \frac{1}{2}x^T Dx \\ \text{Subject to} & Ax \geqq b \end{array} \tag{2.64}$$

satisfying the property that $Ax > b$ is feasible. Let A be of order $m \times n$. Without any loss of generality we assume that D is symmetric (because $x^T Dx = x^T \frac{D+D^T}{2} x$ and $\frac{D+D^T}{2}$ is a symmetric matrix). Let $\mathbf{K} = \{x : Ax \geqq b\}$. By our assumptions here $\mathbf{K} \neq \emptyset$ and in fact $\mathbf{K}$ has a nonempty interior. Every interior point of $\mathbf{K}$ satisfies $Ax > b$ and vice versa. We also assume that $\mathbf{K}$ is bounded. The solution of the problem when $\mathbf{K}$ is unbounded can be accomplished by imposing additional constraints $-\alpha \leqq x_j \leqq \alpha$ for each j, where α is a large positive valued parameter. The parameter α is not given

any specific value, but treated as being larger than any number with which it may be compared. The set of feasible solution of the augmented problem is bounded, and so the augmented problem can be solved by the method discussed below. If the optimum solution of the augmented problem is independent of α when α is positive and large, it is the optimum solution of the original problem (2.64). On the other hand if the optimum solution of the augmented problem depends on α however large α may be, and the optimum objective value diverges to $-\infty$ as α tends to $+\infty$, the objective function is unbounded below in the original problem. In the sequel we assume that $\mathbf{K}$ is bounded. Under these assumptions, (2.64) will have an optimum solution. If D is not PSD, we have the following theorem.

Theorem 2.16 *If D is not PSD, the optimum solution of (2.64) cannot be an interior point of $\mathbf{K}$.*

Proof. Proof is by contradiction. Suppose $\bar{x}$, an interior point of $\mathbf{K}$, is an optimum solution of (2.64). Since $\bar{x}$ is an interior point of $\mathbf{K}$, we have $A\bar{x} > b$, and a necessary condition for it to be optimum for (2.64) (or even for it to be a local minimum for (2.64)) is that the gradient vector of $\theta(x)$ at $\bar{x}$, which is $\nabla\theta(\bar{x}) = c + \bar{x}^T D = 0$. Since D is not PSD, there exists a vector $y \neq 0$ satisfying $y^T Dy < 0$. Using $c + \bar{x}^T D = 0$, it can be verified that $\theta(\bar{x} + \lambda y) = \theta(\bar{x}) + \frac{\lambda^2}{2} y^T Dy$. Since $\bar{x}$ satisfies $A\bar{x} > b$, we can find $\lambda > 0$ and sufficiently small so that $\bar{x} + \lambda y$ is feasible to (2.64), and $\theta(\bar{x} + \lambda y) = \theta(\bar{x}) + \frac{\lambda^2}{2} y^T Dy < \theta(\bar{x})$, contradiction to the hypothesis that $\bar{x}$ is optimal to (2.64). So if D is not PSD, every optimum solution must be a boundary point of $\mathbf{K}$, that is, it must satisfy at least one of the constraints in (2.64) as an equation.

□

The Method

Express the problem in the form (2.64), using the reduction steps discussed above as needed, so that the system $Ax > b$ is feasible. Suppose A is of order $m \times n$. Then we will refer to the problem (2.64) as being of order (m, n), where n is the number of decision variables in the problem, and m the number of inequality constraints on these variables.

Check whether D is PSD. This can be carried out by the efficient algorithm discussed in Section 1.3.1 with a computational effort of $\mathcal{O}(n^3)$. If D is PSD, (2.64) is a convex quadratic program, the optimum solution for it can be computed using the complementary pivot algorithm discussed in earlier sections, with a finite amount of computational effort. If D is not PSD, generate m candidate problems as discussed below. This operation is called the **branching operation**.

For $i = 1$ to m, the i^{th} **candidate problem** is the following :

$$\begin{array}{lll} \text{Minimize} & cx + \frac{1}{2}x^T Dx & \\ \text{Subject to} & A_p.x \geqq b_p, \quad p = 1 \text{ to } m, \ p \neq i & \\ & A_i.x = b_i \ . & \end{array} \tag{2.65}$$

If D is not PSD, by Theorem 2.16, every optimum solution for (2.64) must be an optimum solution of at least one of the m candidate problems.

Each of the candidate problems is now processed independently. The set of feasible solutions of each candidate problem is a subset (a face) of $\mathbf{K}$, the set of feasible solutions of the original problem (2.64). Using the equality constraint, a reduction step can be carried out in the candidate problem (2.65). In the resulting reduced problem identify any binding inequality constraints by a linear programming formulation discussed earlier. Treat binding constraints as equality constraints and carry out further reduction steps. The final reduced problem is one of the same form (2.64), but of order $\leqq (m-1, n-1)$. Test whether it is a convex quadratic programming problem (this could happen even if the original problem (2.64) is not a convex quadratic program) and if it is so, find the optimum solution for it using the complementary pivot algorithm and store its solution in a **solution list**. If it is not a convex quadratic program carry out the branching operation on it and generate additional candidate problems from it, and process each of them independently in the same way.

The total number of candidate problems to be processed is $\leqq 2^{\mathbf{m}}$. When there are no more candidate problems left to be procesed, find out the best solution (i. e., the one with the smallest objective value) among those in the solution list at that stage. That solution is an optimum solution of the original problem.

This provides a finite method for solving the general quadratic programming problem. It may be of practical use only if m and n are small numbers, or if the candidate problems turn out to be convex quadratic programs fairly early in the branching process. On some problems the method may require a lot of computation. For example, if D in the original problem (2.64) is negative definite, every candidate problem with one or more inequality constraints will be nonconvex, and so the method will only terminate when all the extreme points of $\mathbf{K}$ are enumerated in the solution list. In such cases, this method, eventhough finite, is impractical, and one has to resort to heuristics or some approximate solution methods.

2.9.1 Testing Copositiveness

Let M be a given square matrix of order n. Suppose it is required to check whether M is copositive. From the definition, it is clear that M is copositive iff the optimum objective value in the following quadratic program is zero.

$$\begin{array}{ll} \text{Minimize} & x^T M x \\ \text{Subject to} & x \geqq 0 \\ & e^T x \leqq 1 \ . \end{array} \tag{2.66}$$

where e is the column vector of all 1's in $\mathbf{R}^n$. We can check whether M is PSD with a computational effort of $\mathcal{O}(n^{\mathbf{3}})$ by the efficient pivotal methods discussed in Section 1.3.1. If M is PSD, it is also copositive. If M is not PSD, to check whether

it is copositive, we can solve the quadratic program (2.66) by the method discussed above. If the optimum objective value in it is zero, M is copositive, not otherwise. This provides a finite method for testing copositiveness. However, this method is not practilly useful when n is large. Other methods for testing copositiveness are discussed in [3.29, 3.59]. See also Section 2.9.3.

Exercise

2.4 Using the results from Section 8.7, prove that the general quadratic programming problem (2.64) with integer data is an $\mathcal{NP}$-hard problem.

Comments 2.2. Theorem 2.16 is from R. K. Mueller [2.23]. The method for the general quadratic programming problem discussed here is from [2.24] of K. G. Murty.

2.9.2 Computing a KKT point for a General Quadratic Programming Problem

Consider the QP (quadratic program)

$$\begin{array}{lll} \text{minimize} & Q(x) = cx + \frac{1}{2}x^T Dx & \\ \text{subject to} & Ax \geqq b & \\ & x \geqq 0 & \end{array} \tag{2.67}$$

where D is a symmetric matrix of order n, and A, b, c are given matrices of orders $m \times n$, $m \times 1$, and $1 \times n$ respectively. We let $\mathbf{K}$ denote the set of feasible solutions of this problem. If D is PSD, this is a convex quadratic program, and if $\mathbf{K} \neq \emptyset$, the application of the complementary pivot algorithm discussed in Sections 2.2, 2.3 on the LCP corresponding to this QP will either terminate with the global minimum for this problem, or provide a feasible half-line along which $Q(x)$ diverges to $-\infty$.

Here, we do not assume that D is PSD, so (2.67) is the general QP. In this case there can be local minima which are not global minima (see Section 10.2 for definitions of a global minimum, local minimum), the problem may have KKT points which are not even local minima (for example, for (2.66) verify that $x = 0$ is a KKT point, and that this is not even a local minimum for that problem if D is not copositive). The method discussed at the beginning of Section 2.9 is a total enumeration method (enumerating over all the faces of $\mathbf{K}$) applicable when $\mathbf{K}$ is bounded. In this section we do not make any boundedness assumption on $\mathbf{K}$. We prove that if $Q(x)$ is unbounded below on $\mathbf{K}$, there exists a half-line in $\mathbf{K}$ along which $Q(x)$ diverges to $-\infty$. We also prove that if $Q(x)$ is bounded below on $\mathbf{K}$, then (2.67) has a finite global minimum point. This result was first proved by M. Frank and P. Wolfe [10.14] but our proofs are based on the results of B. C. Eaves [2.9]. We also show that the complementary

pivot method applied on an LCP associated with (2.67) will terminate with one of three possible ways

(i) establish that $\mathbf{K} = \emptyset$, or
(ii) find a feasible half-line in $\mathbf{K}$ along which $Q(x)$ diverges to $-\infty$, or
(iii) find a KKT point for (2.67).

From the results in Chapter 1, we know that $\bar{x} \in \mathbf{K}$ is a KKT point for (2.67) iff there exist vectors $\bar{y}, \bar{v} \in \mathbf{R}^m$ and $\bar{u} \in \mathbf{R}^n$ which together satisfy

$$\begin{pmatrix} \bar{u} \\ \bar{v} \end{pmatrix} - \begin{pmatrix} D & -A^T \\ A & 0 \end{pmatrix} \begin{pmatrix} \bar{x} \\ \bar{y} \end{pmatrix} = \begin{pmatrix} c^T \\ -b \end{pmatrix}$$

$$\begin{pmatrix} \bar{u} \\ \bar{v} \end{pmatrix} \geqq 0, \quad \begin{pmatrix} \bar{x} \\ \bar{y} \end{pmatrix} \geqq 0, \quad \begin{pmatrix} \bar{u} \\ \bar{v} \end{pmatrix}^T \begin{pmatrix} \bar{x} \\ \bar{y} \end{pmatrix} = 0 \tag{2.68}$$

which is an LCP. We will call $(\bar{x}, \bar{y}, \bar{u}, \bar{v})$ a KKT solution corresponding to the KKT point $\bar{x}$. For the sake of simplicity, we denote

$$\begin{pmatrix} u \\ v \end{pmatrix} \text{ by } w, \text{ and } \begin{pmatrix} x \\ y \end{pmatrix} \text{ by } z$$

$$\begin{pmatrix} D & -A^T \\ A & 0 \end{pmatrix} \text{ by } M, \text{ and } \begin{pmatrix} c^T \\ -b \end{pmatrix} \text{ by } q$$

$$n + m \text{ by } N \ .$$

So, if $(\bar{w}, \bar{z})$ is complementary solution of the LCP (2.68), then $(\bar{z}_1, \ldots, \bar{z}_n) = \bar{x}$ is a KKT point for (2.67).

A KKT point $\bar{x}$ for (2.67) is said to be a **reduced** KKT **point** for (2.67) if the set of column vectors $\left\{ M_{.j} = \begin{pmatrix} D_{.j} \\ A_{.j} \end{pmatrix} \ : \ j \text{ such that } \bar{x}_j > 0 \right\}$ is linearly independent.

Lemma 2.12 *Let $\bar{x}$ be a KKT point for (2.67). From $\bar{x}$, we can derive either a reduced KKT point $\tilde{x}$ such that $Q(\tilde{x}) \leqq Q(\bar{x})$, or a feasible half-line in $\mathbf{K}$ along which $Q(x)$ diverges to $-\infty$.*

Proof. Let $\big(\bar{w} = (\bar{u}, \bar{v}), \bar{z} = (\bar{x}, \bar{y})\big)$ be a KKT solution associated with $\bar{x}$. Let $\mathbf{J}_1 = \{j : \bar{w}_j = 0\}$, $\mathbf{J}_2 = \{j : \bar{z}_j = 0\}$. By complementarity $\mathbf{J}_1 \cup \mathbf{J}_2 = \{1, \ldots, N\}$. From the fact that $(\bar{w}, \bar{z})$ is a KKT solution (i.e., it satisfies (2.68)) it can be verified that $Q(\bar{x}) = \frac{1}{2}(c\bar{x} + \bar{y}^T b) = \frac{1}{2}(c, b^T)\bar{z}$. Consider the following LP

$$\begin{array}{ll} \text{minimize} & \frac{1}{2}(c, b^T)\bar{z} \\ \text{subject to} & w - Mz = q \\ & w_j = 0 \text{ for } j \in \mathbf{J}_1 \\ & z_j = 0 \text{ for } j \in \mathbf{J}_2 \\ & w_j \geqq 0 \text{ for } j \notin \mathbf{J}_1 \\ & z_j \geqq 0 \text{ for } j \notin \mathbf{J}_2 \end{array} \tag{2.69}$$

If (w, z) is any feasible solution to this LP, from the constraints in (2.69) it is clear that the corresponding $x = (z_1, \ldots, z_n)$ is in $\mathbf{K}$, and that $w^T z = 0$ (complementarity), by this complementarity we have $Q(x) = \frac{1}{2}(c, b^T)z$.

There are only two possibilities for the LP (2.69). Either the objective function is unbounded below in it, in which case there exists a feasible half-line, say $\{(w^1, z^1) + \lambda(w^h, z^h) : \lambda \geqq 0\}$ along which the objective value diverges to $-\infty$ (this implies that the corresponding half-line $\{x^1 + \lambda x^h : \lambda \geqq 0\}$ is in $\mathbf{K}$ and $Q(x)$ diverges to $-\infty$ on it), or that it has an optimum solution, in which case it has an optimum BFS. If $(\tilde{w}, \tilde{z})$ is an optimum BFS of (2.69), the corresponding $\tilde{x}$ is a reduced KKT point for (2.67) and $Q(\tilde{x}) = \frac{1}{2}(c, b^T)\tilde{z} \leqq \frac{1}{2}(c, b^T)\bar{z} = Q(\bar{x})$.

□

Lemma 2.13 *If the QP has a global optimum solution, it has a global optimum solution $\bar{x}$ satisfying the property that the set of vectors $\left\{ \begin{pmatrix} D_{.j} \\ A_{.j} \end{pmatrix} : j \text{ such that } \bar{x}_j > 0 \right\}$ is linearly independent.*

Proof. Follows from Lemma 2.12.

□

Lemma 2.14 *For given D, A; there exists a finite set of matrices $L_1, \ldots, L_l$, each of order $n \times N$, such that for any c, b if x is a reduced KKT point of (2.67), then $x = L_t \begin{pmatrix} c^T \\ -b \end{pmatrix}$ for some t.*

Proof. Let x be a reduced KKT point for (2.67). Let $\big(w = (u, v), z = (x, y)\big)$ be the corresponding KKT solution. Then (w, z) is a BFS of an LP of the form (2.69). Since it is a BFS, there exists a basic vector and associated basis B for (2.69) such that this (w, z) is defined by

$$\begin{aligned} &\text{nonbasic variables} = 0 \\ &\text{basic vector} = B^{-1}q \end{aligned}$$

The matrix L_t can have its j^{th} row to be 0 if x_j is a nonbasic variable, or the r^{th} row of B^{-1}if x_j is the r^{th} basic variable in this basic vector. By complementarity, there are only $2^{\mathbf{N}}$ systems of the form (2.69), and each system has a finite number of basic vectors, so the collection of matrices of the form L_t constructed as above is finite and depends only on D, A. So, for any q, any reduced KKT point must be of the form $L_t q$ for some L_t in this finite collection.

□

Theorem 2.17 *Assume that $\mathbf{K} \neq \emptyset$. Either the QP (2.67) has a global minimum, or there exists a feasible half-line in $\mathbf{K}$ along which $Q(x)$ diverges to $-\infty$.*

Proof. Let $\{\alpha_p : p = 1, 2, \ldots\}$ be an increasing sequence of positive numbers diverging

to $+\infty$, such that $\mathbf{K} \cap \{x : ex \leqq \alpha_1\} \neq \emptyset$. Consider the QP

$$\begin{array}{ll} \text{minimize} & cx + \frac{1}{2}x^T Dx \\ \text{subject to} & Ax \geqq b \\ & x \geqq 0 \\ & ex \leqq \alpha_p \end{array} \tag{2.70}$$

For every p in this sequence, (2.70) has a non-empty bounded solution set, and hence has a global nimimum. By Lemma 2.12, it has a global minimum which is a reduced KKT point for (2.70). Applying Lemma 2.14 to the QP (2.70), we know that there exists a finite collection of matrices $\{L_1, \ldots, L_l\}$ independent of the data in the right hand side constants vector in (2.70), such that every reduced KKT point for (2.70) is of the form

$$L_t \begin{pmatrix} c^T \\ -b \\ \alpha_p \end{pmatrix} = L_t \begin{pmatrix} c^T \\ -b \\ 0 \end{pmatrix} + \alpha_p L_t \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \tag{2.71}$$

for some t. So, for each $p = 1, 2, \ldots$, there exists a t between 1 to l such that the global minimum of (2.70) for that p is of the form given in (2.71). Since there are only a finite number l, of these t's, there must exist a t, say t_1, which gives the global minimum for an infinite number of p's. Let the subsequence corresponding to these p's in increasing order be $P = \{p_1, p_2, \ldots\}$. Let

$$\tilde{x} = L_{t_1} \begin{pmatrix} c^T \\ -b \\ 0 \end{pmatrix} \quad , \quad \bar{y} = L_{t_1} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Then the global minimum for (2.70) is $x(p_r) = \tilde{x} + \alpha_{p_r}\bar{y}$ when $p = p_r$, for $r = 1, 2, \ldots$. So, the optimum objective value in this problem is $Q\big(x(p_r)\big) = Q\big(\tilde{x} + \alpha_{p_r}\bar{y}\big)$, and this is of the form $a_0 + a_1\alpha_{p_r} + a_2\alpha_{p_r}^2$. The quantity α_{p_r} is monotonic increasing with r, so the set of feasible solutions of (2.70) for $p = p_r$ becomes larger as r increases, so $Q\big(x(p_r)\big)$ is monotonic decreasing with r. These facts imply that either $a_2 < 0$ or $a_2 = 0$ and $a_1 \leqq 0$. If $a_2 < 0$ or $a_2 = 0$ and $a_1 < 0$, $Q\big(x(p_r)\big)$ diverges to $-\infty$ as r tends to $+\infty$, in this case $\{\tilde{x} + \lambda\bar{y} : \lambda \geqq \alpha_{p_1}\}$ is a half-line in $\mathbf{K}$ along which $Q(x)$ diverges to $-\infty$. On the other hand, if $a_2 = a_1 = 0$, $Q(x)$ is bounded below by a_0 on $\mathbf{K}$, and in this case $\tilde{x} + \alpha_{p_r}\bar{y}$ is a global minimum for (2.67) for any r.

□

The Algorithm

To compute a KKT point for (2.67), apply the complementary pivot method on the LCP (γ, F) of order $n + m + 1$, where

$$\gamma = \begin{pmatrix} c^T \\ -b \\ q_{n+m+1} \end{pmatrix} \quad , \qquad F = \begin{pmatrix} D & -A^T & e \\ A & 0 & 0 \\ -e^T & 0 & 0 \end{pmatrix}$$

where q_{n+m+1} is treated as a large positive valued parameter without giving any specific value for it (i.e., q_{n+m+1} is treated as being larger than any numer with which it is compared), with the original column vector of the artificial variable z_0 taken to be $(-1,-1,\ldots,-1,0) \in \mathbf{R}^{n+m+1}$. By Lemma 2.9, it can be verified that the matrix M defined above is an L_2-matrix. If the complementary pivot method terminates in a secondary ray, by Theorem 2.5, we conclude that

$$\begin{aligned} -Ax &\leqq -b \\ ex &\leqq q_{n+m+1} \\ x &\geqq 0 \end{aligned}$$

is infeasible for q_{n+m+1} arbitrarily large, that is

$$\begin{aligned} Ax &\geqq b \\ x &\geqq 0 \end{aligned}$$

is infeasible. So (2.67) is infeasible, if ray termination occurs in the complementary pivot algorithm when applied on the LCP (γ, F).

Suppose the complementary pivot method terminates with a complementary solution $\left(\bar{w} = (\bar{w}_j),\ \bar{z} = (\bar{z}_j)\right)$ where $\bar{w}, \bar{z} \in \mathbf{R}^{n+m+1}$. If $\bar{w}_{n+m+1} > 0$, $\bar{z}_{n+m+1} = 0$, it can be verified that $\left((\bar{w}_1,\ldots,\bar{w}_{n+m}),(\bar{z}_1,\ldots,\bar{z}_{n+m})\right)$ is a complementary solution for the LCP $\left(\begin{pmatrix} c^T \\ -b \end{pmatrix},\ \begin{pmatrix} D & -A^T \\ A & 0 \end{pmatrix} \right)$, that is, it is a KKT solution for (2.67) and $\bar{x} = (\bar{z}_1,\ldots,\bar{z}_n)^T$ is a KKT point for (2.67).

On the other hand, if $\bar{w}_{n+m+1} = 0$ and $\bar{z}_{n+m+1} > 0$ in the terminal complementary BFS, the basic variables are affine functions of the large positive parameter q_{n+m+1}. Let $\bar{x} = (\bar{z}_1,\ldots,\bar{z}_n)^T$, $\bar{y} = (\bar{z}_{n+1},\ldots,\bar{z}_{n+m})$. It can be verified that $Q(\bar{x}) = \frac{1}{2}(c\bar{x} + b^T y) - \frac{1}{2}q_{n+m+1}\bar{z}_{n+m+1}$ and as q_{n+m+1} tends to $+\infty$, this diverges to $-\infty$. Hence in this case, $Q(x)$ is unbounded below on $\mathbf{K}$, and a feasible half-line along which $Q(x)$ diverges to $-\infty$ can be obtained by letting the parameter q_{n+m+1} tend to $+\infty$ in the solution $\bar{x}$.

When D is not PSD, it is possible for (2.67) to have some KKT points, even when $Q(x)$ is unbounded below on $\mathbf{K}$. Thus in this case the fact that this algorithm has terminated with a KKT point of (2.67) is no guarantee that $Q(x)$ is bounded below on $\mathbf{K}$.

2.9.3 Computing a Global Minimum, or Even a Local Minimum in Nonconvex Programming Problems May be Hard

Consider the smooth nonlinear program (NLP)

$$\begin{array}{ll} \text{minimize} & \theta(x) \\ \text{subject to} & g_i(x) \geqq 0, \quad i = 1 \text{ to } m \end{array} \tag{2.72}$$

where each of the functions is a real valued function defined on $\mathbf{R}^n$ with high degrees of differentiability. (2.72) is convex NLP if $\theta(x)$ is convex and $g_i(x)$ are concave for all i, nonconvex NLP, otherwise.

A global minimum for (2.72) is a feasible solution $\bar{x}$ for it satisfying $\theta(x) \geqq \theta(\bar{x})$ for all feasible solutions x of the problem. See Section 10.2. For a convex NLP, under some constraint qualifications (see Appendix 4) necessary and sufficient optimality conditions are known. Given a feasible solution satisfying the constraint qualification, using these optimality conditions, it is possible to check efficiently whether that point is a (global) optimum slution of the problem or not.

For a smooth nonconvex nonlinear program, the problem of computing a global minimum, or checking whether a given feasible solution is a global minimum, are hard problems in general. To establish these facts mathematically, consider the subset sum problem, a hard problem in discrete optimization, which is known to be $\mathcal{NP}$-complete (see reference [8.12] for a complete discussion of $\mathcal{NP}$-completeness): given postive integers $d_0, d_1, \ldots, d_n$; is there a solution to

$$\sum_{j=1}^{n} d_j y_j = d_0$$

$$y_j = 0 \text{ or } 1 \text{ for all } j$$

Now consider the quadratic programming problem (QP)

$$\begin{array}{ll} \text{minimize} & \left(\sum_{j=1}^{n} d_j y_j - d_0 \right)^2 + \sum_{j=1}^{n} y_j (1 - y_j) \\ \text{subject to} & 0 \leqq y_j \leqq 1, \quad j = 1 \text{ to } n \ . \end{array}$$

Because of the second term in the objective function, QP is a nonconvex quadratic programming problem. Clearly, the subset-sum problem given above has a feasible solution iff the global minimum objective value in QP is zero. Since the problem of checking whether the subset-sum problem is $\mathcal{NP}$-complete, computing the global minimum for QP, a very special and simple case of a smooth nonconvex NLP, is an $\mathcal{NP}$-hard problem (see reference [8.12] for a complete discussion of $\mathcal{NP}$-hardness). This shows that in general, the problem of computing a global minimum in a smooth nonconvex NLP may be a hard problem. See also Section 10.3 where some of the outstanding difficult problems in mathematics have been formulated as those of finding global minima in smooth nonconvex NLPs (for example, there we show that the well known **Fermat's last Theorem** in number theory, unresolved since 1637 AD, can be posed as the problem of checking whether the global minimum objective value in a smooth nonconvex NLP, (10.1), is zero or greater than zero).

Since the problem of computing a global minimum in a nonconvex NLP is a hard problem, we will now study the question whether it is at least possible to compute a local minimum for such a problem by an efficient algorithm.

For nonconvex NLPs, under constraint qualifications, some necessary conditions for a local minimum are known (see Section 10.2 for the definitions of a local minimum, and Appendix 4 for a discussion of necessary conditions for a local minimum) and there are some sufficient conditions for a point to be a local minimum. But there are no simple conditions known, which are both necessary and sufficient for a given point to be a local minimum. The **complexity** of checking whether a given feasible solution is a local minimum in a nonconvex NLP, is not usually addressed in the literature. Many textbooks in NLP, when they discuss algorithms, leave the reader with the impression that these algorithms converge to a global minimum in convex NLPs, and to a local minimum in nonconvex NLPs. The documentations distributed for many professional NLP software packages also create the same impression. This impression could be quite erroneous, in the general case. In this section we study this problem by examining the computational complexity of determining whether a given feasible solution is not a local minimum, and that of determining whether the objective function is not bounded below on the set of feasible solutions, in smooth continuous variable, nonconvex NLPs. For this purpose, we use the very special instance of an nonconvex quadratic programming problem studied in K. G. Murty and S. N. Kabadi [10.32] with integer data, which may be considered as the simplest nonconvex NLP. It turns out that the questions of determining whether a given feasible solution is not a local minimum in this problem, and to check whether the objective function is not bounded below in this problem, can both be studied using the discrete techniques of computational complexity theory, and in fact these questions are $\mathcal{NP}$-complete problems (see reference [8.12] for definition of $\mathcal{NP}$-completeness). This clearly shows that in general, it is a hard problem to check whether a given feasible solution in a nonconvex NLP is even a local minimum, or to check whether the objective function is bouned below. This indicates the following: when a nonlinear programming algorithm is applied on a nonconvex NLP, unless it is proved that it converges to a point satisfying some known sufficient condition for a local minimum, claims that it leads to a local minimum are hard to verify in the worst case. Also, in continuous variable smooth nonconvex minimization, even the down-to-earth goal of guaranteeing that a local minimum will be obtained by the algorithm (as opposed to the lofty goal of finding the global minimum) may be hard to attain.

We review the known optimality conditions for a given feasible solution $\bar{x}$ to (2.72) to be a local minimum. Let $\mathbf{J} = \{i : g_i(\bar{x}) = 0\}$. Optimality conditions are derived under the assumption that some constraint qualifications (CQ, see Appendix 4) are satisfied at $\bar{x}$, which we assume.

First Order Necessary Conditions for $\bar{x}$ to be a Local Minimum for (2.72)

There must exist a $\bar{\mu}_{\mathbf{J}} = (\bar{\mu}_i : i \in \mathbf{J})$ such that

$$\begin{aligned} &\nabla\theta(\bar{x}) - \sum_{i\in\mathbf{J}} \bar{\mu}_i \nabla g_i(\bar{x}) = 0 \\ &\bar{\mu}_i \geqq 0, \quad \text{for all } i \in \mathbf{J} \ . \end{aligned} \tag{2.73}$$

Given the feasible solution $\bar{x}$, it is possible to check whether these conditions hold, efficiently, using Phase I of the simplex method for linear programming.

Second Order Necessary Conditions for $\bar{x}$ to be a Local Minimum for (2.72)

These conditions include (2.73). Given $\bar{\mu}_{\mathbf{J}}$ satisfying (2.73) together with $\bar{x}$, let $L(x, \bar{\mu}_{\mathbf{J}}) = \theta(x) - \sum_{i\in\mathbf{J}} \bar{\mu}_i g_i(x)$. In addition to (2.73) these conditions require

$$y^T H y \geqq 0, \text{ for all } y \in \{y : \nabla g_i(\bar{x}) y = 0 \text{ for each } i \in \mathbf{J}\} \tag{2.74}$$

where H is the Hessian matrix of $L(x, \bar{\mu}_{\mathbf{J}})$ with respect to x at $x = \bar{x}$. Condition (2.74) requires the solution of a quadratic program involving only equality constraints, which can be solved efficiently. It is equivalent to checking the positive semidefiniteness of a matrix which can be carried out efficiently using Gaussian pivot steps (see Section 1.3.1).

Sufficient Conditions for $\bar{x}$ to be a Local Minimum for (2.72)

Given the feasible solution $\bar{x}$, and $\bar{\mu}_{\mathbf{J}}$ which together satisfy (2.73), the most general known sufficient optimality condition states that if

$$y^T H y > 0 \text{ for all } y \in \mathbf{T}_1 \tag{2.75}$$

where $\mathbf{T}_1 = \big\{y : y \neq 0 \text{ and } \nabla g_i(\bar{x}) y = 0 \text{ for each } i \in \{i : i \in \mathbf{J} \text{ and } \bar{\mu}_i > 0\}$, and $\nabla g_i(\bar{x}) y \geqq 0$ for each $i \in \{i : i \in \mathbf{J} \text{ and } \bar{\mu}_i = 0\}\big\}$, then $\bar{x}$ is a local minimum for (2.72). Unfortunately, when H is not positive semidefinite, the problem of checking whether (2.75) holds, leads to a nonconvex QP, which, as we will see later, may be hard to solve.

Aside from the question of the difficulty of checking whether (2.75) holds, we can verify that the gap between conditions (2.74) and (2.75) is very wide, particulary when the set $\{i : i \in \mathbf{J} \text{ and } \bar{\mu}_i = 0\} \neq \emptyset$. In this case, condition (2.74) may hold, and even if we are able to check (2.75), if it is not satisfied, we are unable to determine whether $\bar{x}$ is a local minimum for (2.72) with present theory.

Now we will use a simple indefinite QP, related to the problem of checking whether the sufficient optimality condition (2.75) holds, to study the following questions :

i) Given a smooth nonconvex NLP and a feasible solution for it, can we check whether it is a local minimum or not efficiently ?

ii) At least in the simple case when the constraints are linear, can we check efficiently whether the objective function is bounded below or not on the set of feasible solutions ?

Let D be an integer square symmetric matrix of order n. The problem of checking whether D is not PSD involves the question

$$\text{"is there an } x \in \mathbf{R}^n \text{ satisfying } x^T Dx < 0 \text{ ?"} \tag{2.76}$$

This can be answered with an effort of at most n Gaussian pivot steps, by the techniques discussed in Section 1.3.1. This leads to an $\mathcal{O}(n^3)$ algorithm for this problem. At the termination of this algorithm, it is in fact possible to actually produce a vector x satisfying $x^T Dx < 0$, if the answer to (2.76) is in the affirmative.

All PSD matrices are copositive, but a matrix which is not PSD may be copositive. Testing whether the given matrix D is not copositive involves the question

$$\text{"is there an } x \geqq 0 \text{ satisfying } x^T Dx < 0 \text{ ?"} \tag{2.77}$$

If D is not PSD, no efficient algorithm for this question is known (the computational complexity of the enumerative method of Section 2.9.1 grows exponentially with n in the worst case). In fact we show later that this question is $\mathcal{NP}$-complete. To study this question, we are naturally lead to the NLP

$$\begin{array}{ll} \text{minimize} & Q(x) = x^T Dx \\ \text{subject to} & x \geqq 0 \end{array} \tag{2.78}$$

We will show that this problem is an $\mathcal{NP}$-hard problem.

We assume that D is not PSD. So $Q(x)$ is nonconvex and (2.78) is a nonconvex NLP. It can be considered the **simplest nonconvex NLP**. We consider the following decision problems.

Problem 1: Is $x = 0$ not a local minimum for (2.78) ?

Problem 2: Is $Q(x)$ not bounded below on the set of feasible solutions of (2.78) ?

Clearly, the answer to problem 2 is in the affirmative iff the answer to problem 1 is. We will show that both these problems are $\mathcal{NP}$-complete. To study problem 1, we can replace (2.78) by the NLP

$$\begin{array}{ll} \text{minimize} & Q(x) = x^T Dx \\ \text{subject to} & 0 \leqq x_j \leqq 1, \quad j = 1 \text{ to } n \end{array} \tag{2.79}$$

Lemma 2.15 *The decision problem "is there an $\bar{x}$ feasible to (2.79) which satisfies $Q(\bar{x}) < 0$", is in the class $\mathcal{NP}$ (see [8.12] for the definition of the class $\mathcal{NP}$ of decision problems).*

Proof. Given an x feasible to (2.79), to check whether $Q(x) < 0$, can be done by computing $Q(x)$ which takes $\mathcal{O}(n^2)$ time. Also, if the answer to the problem is in the affirmative, an optimum solution $\bar{x}$ of (2.79) satisfies $Q(\bar{x}) < 0$. There is a linear complementarity problem (LCP) corresponding to (2.79) and an optimum solution for (2.79) must correspond to a BFS for this LCP. Since there are only a finite number of BFSs for an LCP, and they are all rational vectors, a nondeterministic algorithm can find one of them satisfying $Q(x) < 0$, if it exists, in polynomial time. Hence, this problem is in the class $\mathcal{NP}$.

□

Lemma 2.16 *The optimum objective value in (2.79) is either 0 or* $\leqq -2^{-\mathbf{L}}$ *where* L *is the size of D, (i.e., the total number of binary digits in all the data in D).*

Proof. Since the set of feasible solutions of (2.79) is a compact set and $Q(x)$ is continuous, (2.79) has an optimum solution. The necessary optimality conditions for (2.79) lead to the following LCP

$$\begin{pmatrix} u \\ v \end{pmatrix} - \begin{pmatrix} D & I \\ -I & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ e \end{pmatrix} \tag{2.80}$$

$$\begin{pmatrix} u \\ v \end{pmatrix} \geqq 0 \ , \qquad \begin{pmatrix} x \\ y \end{pmatrix} \geqq 0 \ , \tag{2.81}$$

$$\begin{pmatrix} u \\ v \end{pmatrix}^T \begin{pmatrix} x \\ y \end{pmatrix} = 0 \tag{2.82}$$

It can be verified that whenever (u, v, x, y) satisfies (2.80), (2.81) and (2.82), $x^T Dx = -e^T y$, a linear function, where e is the column vector of all 1's in $\mathbf{R}^n$. There exists an optimum solution of (2.79) which is a BFS of (2.80), (2.81). By the results under the ellipsoid algorithm (see, for example Chapter 8 in this book, or Chapter 15 in [2.26]), in every BFS of (2.80), (2.81), each y_j is either 0 or $\geqq 2^{-\mathbf{L}}$. If the optimum objective value in (2.79) is not zero, it must be < 0, and this together with the above facts implies that an optimum solution x or (2.79) corresponds to a BFS (u, v, x, y) of (2.80), (2.81) in which $-e^T y < 0$. All these facts clearly imply that the optimum objective value in (2.79) is either 0 or $\leqq -2^{-\mathbf{L}}$.

□

We now make a list of several decision problems, some of which we have already seen, and some new ones which we need for establishing our results.

Problem 3: Is there an $x \geqq 0$ satisfying $Q(x) < 0$?

Problem 4: For any positive integer a_0, is there an $x \in \mathbf{R}^n$ satisfying $e^T x = a_0$, $x \geqq 0$ and $Q(x) < 0$?

Now consider a subset sum problem with data d_0; $d_1, \ldots, d_n$, which are all positive integers. Let γ be a positive integer $> 4 \left(d_0 \left(\sum_{j=1}^{n} d_j \right) \right)^{\mathbf{2}} n^{\mathbf{3}}$. Let l be the size of this

subset sum problem, that is, the total number of binary digits in all the data for the problem. Let ε be a positive rational number $< 2^{-\mathbf{nl^2}}$. The subset sum problem is :

Problem 5: Subset sum problem: Is there a $y = (y_j) \in \mathbf{R}^n$ satisfying $\sum_{j=1}^{n} d_j y_j = d_0$, $0 \leqq y_j \leqq 1$, $j = 1$ to n, and y integer vector ?

We now define several functions involving nonnegative variables $y = (y_1, \ldots, y_n)^T$ and $s = (s_1, \ldots, s_n)^T$, related to the subset sum problem.

$$f_1(y,s) = \left(\sum_{j=1}^{n} d_j y_j - d_0\right)^{\mathbf{2}} + \gamma \left(\sum_{j=1}^{n} (y_j + s_j - 1)^{\mathbf{2}}\right) + \sum_{j=1}^{n} y_j s_j$$

$$= \left(\sum_{j=1}^{n} d_j y_j\right)^{\mathbf{2}} + \sum_{j=1}^{n} y_j s_j + \gamma \sum_{j=1}^{n} (y_j + s_j)^{\mathbf{2}}$$

$$- 2d_0 \left(\sum_{j=1}^{n} d_j y_j\right) + 2\gamma \sum_{j=1}^{n} (y_j + s_j) + n\gamma + d_0^{\mathbf{2}}$$

$$f_2(y,s) = f_1(y,s) + 2d_0 \left(\sum_{j=1}^{n} d_j y_j (1 - y_j)\right)$$

$$= \left(\sum_{j=1}^{n} d_j y_j\right)^{\mathbf{2}} + \gamma \sum_{j=1}^{n} (y_j + s_j)^{\mathbf{2}} + \sum_{j=1}^{n} y_j s_j$$

$$- 2d_0 \left(\sum_{j=1}^{n} d_j y_j^{\mathbf{2}}\right) + 2\gamma \sum_{j=1}^{n} (y_j + s_j) + n\gamma + d_0^{\mathbf{2}}$$

$$f_3(y,s) = \left(\sum_{j=1}^{n} d_j y_j\right)^{\mathbf{2}} + \gamma \sum_{j=1}^{n} (y_j + s_j)^{\mathbf{2}} + \sum_{j=1}^{n} y_j s_j$$

$$- 2d_0 \left(\sum_{j=1}^{n} d_j y_j^{\mathbf{2}}\right) + d_0^{\mathbf{2}} - n\gamma$$

$$f_4(y,s) = \left(\sum_{j=1}^{n} d_j y_j\right)^{\mathbf{2}} + \gamma \sum_{j=1}^{n} (y_j + s_j)^{\mathbf{2}} + \sum_{j=1}^{n} y_j s_j$$

$$- 2d_0 \sum_{j=1}^{n} d_j y_j^{\mathbf{2}} + \left(\frac{d_0^{\mathbf{2}} - n\gamma}{n^{\mathbf{2}}}\right) \left(\sum_{j=1}^{n} (y_j + s_j)\right)^{\mathbf{2}}$$

$$f_5(y,s) = f_4(y,s) - \left(\frac{\varepsilon}{n^{\mathbf{2}}}\right) \left(\sum_{j=1}^{n} (y_j + s_j)\right)^{\mathbf{2}}$$

Let $\mathbf{P} = \{(y, s) : y \geqq 0, s \geqq 0, \sum_{j=1}^{n}(y_j + s_j) = n\}$. Consider the following additional decision problems

Problem 6: Is there a $(y, s) \in \mathbf{P}$ satisfying $f_1(y, s) \leqq 0$?

Problem 7: Is there a $(y, s) \in \mathbf{P}$ satisfying $f_2(y, s) \leqq 0$?

Problem 8: Is there a $(y, s) \in \mathbf{P}$ satisfying $f_4(y, s) \leqq 0$?

Problem 9: Is there a $(y, s) \in \mathbf{P}$ satisfying $f_5(y, s) < 0$?

Theorem 2.18 *Problem 4 is an $\mathcal{NP}$-hard problem (see [8.11] for the definitions of an $\mathcal{NP}$-hard problem).*

Proof. Since $f_1(y, s)$ is a sum of nonnegative terms whenever $(y, s) \in \mathbf{P}$, if $(\bar{y}, \bar{s}) \in \mathbf{P}$ satisfies $f_1(y, s) \leqq 0$, then we must have $f_1(\bar{y}, \bar{s}) = 0$, this clearly implies from the definition of $f_1(y, s)$, that the following conditions must hold.

$$\sum_{j=1}^{n} d_j \bar{y}_j = d_0, \quad \bar{y}_j \bar{s}_j = 0 \text{ and } \bar{y}_j + \bar{s}_j = 1, \text{ for all } j = 1 \text{ to } n \ .$$

These conditions clearly imply that $\bar{y}$ is a solution of the subset sum problem and that the answer to problem 5 is in the affirmative. Conversely if $\hat{y} = (\hat{y}_j)$ is a solution to the subset sum problem, define $\hat{s} = (\hat{s}_j)$ where $\hat{s}_j = 1 - \hat{y}_j$ for each $j = 1$ to n, and it can be verified that $f_1(\hat{y}, \hat{s}) = 0$. This verifies that problems 5 and 6 are equivalent.

Whenever $\bar{y}$ is a 0–1 vector, we have $\bar{y}_j = \bar{y}_j^{\mathbf{2}}$ for all j, and this implies that $f_1(\bar{y}, s) = f_2(\bar{y}, s)$ for any s. So, from the above arguments, we see that if $(\bar{y}, \bar{s}) \in \mathbf{P}$ satisfies $f_1(\bar{y}, \bar{s}) \leqq 0$, then $f_1(\bar{y}, \bar{s}) = f_2(\bar{y}, \bar{s}) = 0$. If $0 \leqq y_j \leqq 1$, we have $2d_0 d_j y_j(1-y_j) \geqq 0$. If $(y, s) \in \mathbf{P}$, and $y_j > 1$, then $\frac{\gamma}{2}(y_j + s_j - 1)^{\mathbf{2}} + 2d_0 d_j y_j(1 - y_j) \geqq 0$, since γ is large (from the definition of γ). Using this and the definitions of $f_1(y, s)$, $f_2(y, s)$, it can be verified that for $(y, s) \in \mathbf{P}$, if $f_2(y, s) \leqq 0$ then $f_1(y, s) \leqq 0$ too. These facts imply that problems 6 and 7 are equivalent.

Clearly, problems 7 and 8 are equivalent.

From the definition of ε (since it is sufficiently small) and using Lemma 2.16, one can verify that problems 8 and 9 are equivalent.

Problem 9 is a special case of problem 4. Since problem 5 is $\mathcal{NP}$-complete, from the above chain of arguments we conclude that problem 4 is $\mathcal{NP}$-hard.

□

Theorem 2.19 *Problem 4 is $\mathcal{NP}$-complete.*

Proof. The answer to problem 4 is in the affirmative iff the answer to the decision problem in the statement of Lemma 2.15 is in the affirmative. So, from Lemma 2.15 we conlcude that problem 4 is in $\mathcal{NP}$. From Theorem 2.18, this shows that problem 4 is $\mathcal{NP}$-complete.

□

Theorem 2.20 *Problem 3 is $\mathcal{NP}$-complete.*

Proof. Problems 3 and 4 are clearly equivalent, this result follows from Theorem 2.19. □

Theorem 2.21 *Both problems 1 and 2 are $\mathcal{NP}$-complete.*

Proof. Problems 1 and 2 are both equivalent to problem 3, so this result follows from Theorem 2.20.

□

Theorem 2.22 *Given an integer square matrix D, the decision problem "is D not copositive ?" is $\mathcal{NP}$-complete.*

Proof. The decision problem "is D not copositive ?" is equivalent to problem 1, hence this result follows from Theorem 2.21.

□

Can We Check Local Minimality Efficiently In Unconstrained Minimization Problems ?

Let $\theta(x)$ be a real valued smooth function defined on $\mathbf{R}^n$. Consider the unconstrained problem

$$\text{minimize} \quad \theta(x) \ . \tag{2.83}$$

A necessary condition for a given point $\bar{x} \in \mathbf{R}^n$ to be a local minimum for (2.83) is (see Appendix 4)

$$\nabla\theta(\bar{x}) = 0, \quad H(\theta(\bar{x})) \text{ is PSD} \tag{2.84}$$

where $H(\theta(\bar{x}))$ is the Hessian matrix (the matrix of second order partial derivatives) of $\theta(x)$ at $\bar{x}$. A sufficient condition for $\bar{x}$ to be a local minimum for (2.83) is

$$\nabla\theta(\bar{x}) = 0, \quad H(\theta(\bar{x})) \text{ is positive definite.} \tag{2.85}$$

Both conditions (2.84) and (2.85) can be checked very efficiently. If (2.84) is satisfied, but (2.85) is violated, there are no simple conditions known to check whether or not $\bar{x}$ is a local minimum for (2.83). Here, we investigate the complexity of checking whether or not a given point $\bar{x}$ is a local minimum for (2.83), and that of checking whether $\theta(x)$ is bounded below or not over $\mathbf{R}^n$.

As before, let $D = (d_{ij})$ be an integer square symmetric matrix of order n. Consider the unconstrained problem,

$$\text{minimize } h(u) = (u_1^{\mathbf{2}}, \ldots, u_n^{\mathbf{2}})D(u_1^{\mathbf{2}}, \ldots, u_n^{\mathbf{2}})^T \tag{2.86}$$

Clearly, (2.86) is an instance of the general unconstrained minimization problem (2.83). Consider the following decision problems.

Problem 10: Is $\bar{u} = 0$ not a local minimum for (2.86) ?

Problem 11: Is $h(u)$ not bounded below on $\mathbf{R}^n$?

We have, for $i, j = 1$ to n

$$\frac{\partial h(u)}{\partial u_j} = 4u_j\left((u_1^{\mathbf{2}}, \ldots, u_n^{\mathbf{2}})D_{.j}\right)$$
$$\frac{\partial^2 h(u)}{\partial u_i \partial u_j} = 8u_i u_j d_{ij}, \quad i \neq j$$
$$\frac{\partial^2 h(u)}{\partial u_j^2} = 4(u_1^{\mathbf{2}}, \ldots, u_n^{\mathbf{2}})D_{.j} + 8u_j^{\mathbf{2}} d_{jj}$$

where $D_{.j}$ is the j^{th} column vector of D. So, $\bar{u} = 0$ satisfies the necessary conditions for being a local minimum for (2.86), but not the sufficient condition given in (2.85).

Using the transformation $x_j = u_j^{\mathbf{2}}$, $j = 1$ to n, we see that (2.86) is equivalent to (2.78). So problem 1 and 10 are equivalent. Likewise, problems 2 and 11 are equivalent. By Theorem 2.21, we conclude that both problems 10 and 11 are $\mathcal{NP}$-hard. Thus, even in the unconstrained minimization problem, to check whether the objective function is not bounded below, and to check whether a given point is not a local minimum, may be hard problems in general. This also shows that the problem of checking whether a given smooth nonlinear function (even a polynomial) is or is not locally convex at a given point, may be a hard problem in general.

What Are Suitable Goals for Algorithms in Nonconvex NLP ?

Much of nonlinear programming literature stresses that the goal for algorithms in nonconvex NLPs should be to obtain a local minimum. Our results here show that in general, this may be hard to guarantee.

Many nonlinear programming algorithms are iterative in nature, that is, beginning with a initial point x^0, they obtain a sequence of points $\{x^r : r = 0, 1, \ldots\}$. For some of the algorithms, under certain conditions, it can be shown that the sequence converges to a KKT point for the original problem, (a KKT point is a feasible solution at which the first order necessary conditions for a local minimum, (2.73), hold). Unfortunately, there is no guarantee that a KKT point will be a local minimum, and our results point out that in general, checking whether or not it is a local minimum may be a hard problem.

Some algorithms have the property that the sequence of points obtained is actually a descent sequence, that is, either the objective function, or a measure of the infesibility of the current solution to the problem, or some merit function or criterion function which is a combination of both, strictly decreases along the sequence. Given x^r, these algorithms generate a $y^r \neq 0$ such that the direction $x^r + \lambda y^r$, $\lambda \geqq 0$, is a descent direction for the functions discussed above. The next point in the sequence x^{r+1} is usually taken to be the point which minimizes the objective or criterion function on the half-line $\{x^r + \lambda y^r : \lambda \geqq 0\}$, obtained by using a line minimization algorithm. On general nonconvex problems, these methods suffer from the same difficulties, they cannot theoretically guarantee that the point obtained at termination is even a local

minimum. However, it seems reasonable to expect that a solution obtained through a descent process is more likely to be a local minimum, than a solution obtained purely based on necessary optimality conditions. Thus a suitable goal for algorithms for non-convex NLPs seems to be a descent sequence converging to a KKT point. Algorithms, such as the sequential quadratic programming methods discussed in Section 1.3.6, and those discussed in Chapter 10, reach this goal.

2.10 Exercises

2.5 Let $\theta(x)$ be a convex function defined on $\mathbf{R}^n$, which is known to be unbounded below on $\mathbf{R}^n$. Does there exist a half-line along which $\theta(x)$ diverges to $-\infty$? Either prove that it does, or construct a counterexample. Does the answer change if $\theta(x)$ is known to be a differentiable convex function ?

2.6 Consider the problem

$$\begin{array}{ll} \text{Minimize} & \theta(x) \\ \text{Subject to} & Ax \geqq b \end{array}$$

where A is a matrix of order $m \times n$, and $\theta(x)$ is a convex function. Suppose it is known that $\theta(x)$ is unbounded below in this problem. Does there exist a feasible half-line along which $\theta(x)$ diverges to $-\infty$? Either prove that it does, or construct a counterexample. Does the answer change if $\theta(x)$ is a differentiabl convex function ?

2.7 If the data in the LCP (q, M) satisfies

i) $M + M^T \geqq 0$, and

ii) $q - M^T z \geqq 0, \quad z \geqq 0$ is feasible,

prove that the complementary pivot algorithm will terminate with a solution when applied on the LCP (q, M).
(Philip C. Jones [2.16])

2.8 Let $\mathbf{G}_j$ be the set $\{(w, z) : w - Mz = q,\ w \geqq 0,\ z \geqq 0,\ w_i z_i = 0 \text{ for all } i \neq j\}$, and let $\mathbf{G} = \bigcup(\mathbf{G}_j : j = 1 \text{ to } n)$. If M is PSD or a P-matrix, prove that $\mathbf{G}$ is a connected subset of $\mathbf{R}^n$. If (q, M) is the LCP corresponding to the following quadratic program, show that $\mathbf{G}$ is not connected.

$$\begin{array}{ll} \text{minimize} & cx + \frac{1}{2}x^T Dx \\ \text{subject to} & 0 \leqq x \leqq u \end{array}$$

where $D = \begin{pmatrix} -2 & -3 & -3 \\ -3 & -5 & -1 \\ -1 & -1 & 4 \end{pmatrix}$, $u = \begin{pmatrix} 10 \\ 10 \\ 10 \end{pmatrix}$, $c^T = \begin{pmatrix} 4 \\ 3 \\ 5 \end{pmatrix}$.
(W. P. Hallman and I. Kaneko [2.15])

2.9 Prove that the complementary pivot algorithm will process the LCP (q, M) if M is a Z-matrix.
(R. Saigal [2.32])

2.10 Let $\{A_{.1}, \ldots, A_{.n-1}\}$ be a linearly independent set of column vectors in $\mathbf{R}^n$. Let $\S = \{y^1, \ldots, y^r\}$ be another finite set of column vectors in $\mathbf{R}^n$, and let $b \in \mathbf{R}^n$ be another given column vector. It is required to choose $A_{.n} \in \S$ so that the minimum distance from b to $\text{Pos}\{A_{.1}, \ldots, A_{.n}\}$ is a small as possible. Develop an efficient algorithm for doing it.

2.11 Let

$$\widehat{M} = \begin{pmatrix} -1 & 2 \\ 2 & -1 \end{pmatrix}, \qquad \hat{q} = \begin{pmatrix} -1 \\ -2 \end{pmatrix} .$$

Show that the LCP $(\hat{q}, \widehat{M})$ has a solution. However, show that all the variants of the complementary pivot algorithm discussed in this Chapter are unable to find a solution to this LCP $(\hat{q}, \widehat{M})$.

2.12 Let (P) be a linear programming problem, and (Q) the corresponding linear complementary problem as obtained in Section 1.2. It has been suggested that the sequence of solutions generated when the LCP, (Q), is solved by the complementary pivot method, is the same as the sequence of solutions generated when the LP, (P), is solved by the self-dual parametric algorithm (see Section 8.13 of [2.26]). Discuss, and examine the similarities between the self-dual parametric algorithm applied to (P) and the complementary pivot method applied on (Q).

2.13 Let

$$M = \begin{pmatrix} 2 & 2 & 1 & 2 \\ 3 & 3 & 2 & 3 \\ -2 & 1 & 5 & -2 \\ 1 & -2 & 1 & 2 \end{pmatrix}, \qquad q = \begin{pmatrix} -4 \\ -6 \\ 4 \\ 4 \end{pmatrix} .$$

i) Prove that M is strictly copositive.
ii) Show that the LCP (q, M) has an infinite number of complementary feasible solutions.

2.14 Given a square matrix M of order n, let $\mathbf{K}(M)$ denote the union of all the complementary cones in $\mathcal{C}(M)$. Prove that $\mathbf{K}(M)$ is convex iff $\mathbf{K}(M) = \{q : q + Mz \geqq 0, \text{ for some } z \geqq 0\}$.
(B. C. Eaves [2.8])

2.15 Let $a_1, \ldots, a_n, b$ be positive integers satisfying $b > \max\{a_1, \ldots, a_n\}$. Let

$$q(n+2) = (a_1, \ldots, a_n, -b, b)^T$$

$$M(n+2) = \begin{pmatrix} & 0 & 0 \\ -I_n & \vdots & \vdots \\ & 0 & 0 \\ e_n^T & -1 & 0 \\ -e_n^T & 0 & -1 \end{pmatrix}$$

where I_n is the identity matrix of order n, and e_n^T is the row vector in $\mathbf{R}^n$ all the entries in which are "1". Consider the LCP $(q(n+2), M(n+2))$ of order $n+2$. Are any of the algorithms discussed in this chapter able to process this LCP ? Why ? If not, develop an algorithm for solving this LCP using the special structure of the matrix M.

2.16 Consider the quadratic program

$$\begin{array}{ll} \text{minimize} & -x_1 - 2x_2 + \frac{1}{2}(2x_1^2 + 4x_1x_2 + 4x_2^2) \\ \text{subject to} & 3x_1 - 2x_2 - x_3 \qquad = 2 \\ & -x_1 + 2x_2 \qquad - x_4 = 6 \\ & x_j \geqq 0 \quad \text{for all } j \ . \end{array}$$

Formulate this program as an LCP of order 4 and write down this LCP clearly. Does a solution of this LCP lead to a solution of this quadratic program ? Why ?

It is required to solve this LCP using the variant of complementary pivot method in which the column vector of the artificial variable is $(1, 2, 2, 6)^T$. Obtain the canonical tableau corresponding to the initial almost complementary basic vector, and then carry out exactly one more pivot step in this algorithm.

2.17 Suppose $B \geqq 0$, and the linear programs

i) Maximize $c^T x$; subject to $Ax \leqq b$, $x \geqq 0$ and

ii) Minimize $b^T y$; subject to $(A+B)^T y \geqq c$, $y \geqq 0$

have finite optimum solutions. Show that the complementary pivot algorithm terminates with a complementary feasible solution for the LCP (q, M) with

$$q = \begin{pmatrix} -c \\ b \end{pmatrix} \ , \qquad M = \begin{pmatrix} 0 & (A+B)^T \\ -A & 0 \end{pmatrix} \ .$$

(G. B. Dantzig and A. S. Manne [2.6])

2.18 Let $\mathbf{\Gamma}$ be a nonenmpty closed convex subset of $\mathbf{R}^n$. For each $x \in \mathbf{R}^n$ let $P_\Gamma(x)$ denote the nearest point in $\mathbf{\Gamma}$ to x in terms of the usual Euclidean distance. Prove the following :

(i) $||P_\Gamma(x) - y||^2 \leqq ||x - y||^2$ for all $x \in \mathbf{R}^n, y \in \mathbf{\Gamma}$.

(ii) $||P_\Gamma(x) - P_\Gamma(y)||^2 \leqq ||x - y||^2$ for all $x, y \in \mathbf{R}^n$.

(Y. C. Cheng [3.6])

2.19 Let G and H be symmetric PSD matrices of order n and m respectively. Consider the following quadratic programs :

$$\begin{array}{ll} \text{maximize} & cx - \frac{1}{2}x^T Gx - \frac{1}{2}y^T Hy \\ \text{subject to} & Ax - Hy \leqq b \\ & x \quad\quad \geqq 0 \end{array}$$

and

$$\begin{array}{ll} \text{minimize} & b^T y + \frac{1}{2}x^T Gx + \frac{1}{2}y^T Hy \\ \text{subject to} & Gx + A^T y \geqq c^T \\ & y \geqq 0 \end{array}$$

Prove that if both the problems are feasible, then each has an optimal solution, and the optimum objective values are equal; moreover, the optimal solutions can be taken to be the same.
(R. W. Cottle [2.5] and W. S. Dorn [2.7])

2.20 Let M be a nondegenerate square matrix of order n. Let $d \in \mathbf{R}^n$, $d > 0$ be such that for every $\mathbf{J} \subset \{1, \ldots, n\}$, if $d_{\mathbf{J}} = (d_j : j \in \mathbf{J})$, $M_{\mathbf{JJ}} = (m_{ij} : i, j \in \mathbf{J})$, then $(M_{\mathbf{JJ}})^{-1} d_{\mathbf{J}} \geqq 0$. Then prove that if the LCP (q, M) is solved by the variant of the complementary pivot algorithm discussed in Section 2.3.3 with $-d$ as the original column vector for the artificial variable z^0, it will terminate with a solution of the LCP after at most $(n + 1)$ pivot steps.
(J. S. Pang and R. Chandrasekaran [8.18])

2.21 Consider the process of solving the LCP (q, M) by the complementary pivot algorithm. Prove that the value of the artificial variable z_0 decreases as the algorithm progresses, whenever M is either a PSD matrix or a P-matrix or a P_0-matrix, until termination occurs.
(R. W. Cottle [4.5] and B. C. Eaves [2.8])

2.22 Consider the process of solving the LCP (q, M) by the variant of the complementary pivot algorithm discussed in Section 2.3.3 with the column vector $d > 0$ as the initial column vector associated with the artificial variable z_0. Prove that in this process, there exists no secondary ray for all $d > 0 > q$ iff M is an $L_\star$-matrix. Using this prove that the variant of the complementary pivot algorithm discussed in Section 2.3.3 with the lexico minimum ratio rule for the dropping variable section in each step, will always terminate with a complementary solution for all q, no matter what $d > 0$ is used, iff M is an $L_\star$-matrix.
(B. C. Eaves [2.8])

2.23 Consider the convex quadratic programming problem

$$\begin{array}{lll} \text{minimize} & Q(x) = cx + \frac{1}{2}x^T Dx \\ \text{subject} & Ax \geqq b \\ & x \geqq 0 \end{array}$$

where D is a symmetric PSD matrix. If the problem has alternate optimum solution prove the following :

(i) the set of optimum solutions is a convex set,

(ii) $(y-x)^T D(y-x) = 0$ and actually $(y-x)^T D = 0$ for every pair of optimum solutions x and y, of the problem,

(iii) the gradient vector of $Q(x)$, $\nabla Q(x)$ is a constant on the set of optimum solutions,

(iv) the set of optimum solutions is the intersection of the constraint set with some linear manifold.

(M. Frank and P. Wolfe [10.14])

2.24 Let A^1, B^1, two given matrices of orders $m \times n$ each, be the loss matrices in a bimatrix game problem. Prove that the problem of computing a Nash equilibrium strategy pair of vectors for this bimatrix game, can be posed as the LCP (q, M), where

$$q = \begin{pmatrix} -e_m \\ e_n \end{pmatrix} \quad , \qquad M = \begin{pmatrix} 0 & A \\ B^T & 0 \end{pmatrix}$$

where $A > 0$ and $B < 0$. Prove (use Lemma 2.8) that the complementary pivot algorithm will terminate with a solution when applied on this LCP.
(B. C. Eaves [2.8])

2.25 Consider the LCP (q, M) of order n. Let $\mathbf{C}_1$ be the set of feasible solutions of the system

$$\begin{array}{rl} w - Mz = & q \\ w,\ z \geqq & 0 \\ w_j z_j = & 0,\ \ j = 2 \text{ to } n. \end{array}$$

If q is nondegenerate in the LCP (q, M) (i.e., if in every solution (w, z) of the system of linear equations "$w - Mz = q$", at least n variables are nonzero) prove that $\mathbf{C}_1$ is a disjoint union of edge paths. What happens to this result if q is degenerate ?

2.26 In Merrill's algorithm for computing a Kakutani fixed point discussed in Section 2.7.8, we defined the piecewise linear map in the top layer of the special triangulation of $\mathbf{R}^n \times [0,1]$ by defining for any vertex $V = \begin{pmatrix} v \\ 1 \end{pmatrix}$, $f(V) = \begin{pmatrix} f(v) \\ 1 \end{pmatrix}$ where $f(v)$ is an arbitrary point chosen from the set $\mathbf{F}(v)$. Examine the advantages that could be gained by defining $f(v)$ to be the nearest point (in terms of the usual Euclidean distance) in the set $\mathbf{F}(v)$ to v.

2.27 Let M, q be given matrices of orders $n \times n$ and $n \times 1$ respectively. If $y^T My + y^T q$ is bounded below on the set $\{y : y \geqq 0\}$, prove that the LCP (q, M) has a complementary solution, and that a complementary solution can be obtained by applying the complementary pivot algorithm on the LCP of order $(n+1)$ with data

$$q = \begin{pmatrix} q \\ q_{n+1} \end{pmatrix} \quad , \qquad M = \begin{pmatrix} M & e \\ -e^T & 0 \end{pmatrix}$$

where $q_{n+1} > 0$, with the initial column vector associated with the artificial variable z_0 to be $(-1, \ldots, -1, 0) \in \mathbf{R}^{n+1}$.
(B. C. Eaves [2.8])

2.28 Consider the general quadratic program (2.67). If $Q(x)$ is unbounded below on the set of feasible solutions $\mathbf{K}$ of this problem, prove that there exists a feasible half-line through an extreme point of $\mathbf{K}$ along which $Q(x)$ diverges to $-\infty$.
(B. C. Eaves [2.9])

2.29 Let M be a given square matrix of order n. Let $\{B_{.1}, \ldots, B_{.r}\}$ be a given set of column vectors in $\mathbf{R}^n$. It is required to check whether $x^T Mx$ is $\geqq 0$ for all $x \in \text{Pos}\{B_{.1}, \ldots, B_{.r}\}$. Transform this into the problem of checking the copositivity of a matrix.

Can the problem of checking whether $x^T Mx$ is $\geqq 0$ for all $x \in \{x : Ax \geqq 0\}$ where A is a given matrix of order $m \times n$, be also transformed into the problem of checking the copositivity of a matrix ? How ?

2.30 (Research Problem) Application to pure 0-1 Integer Programming

Consider the pure 0-1 integer programming problem

$$\begin{array}{ll} \text{minimize} & cx \\ \text{subject to} & Ax = b \\ & Dx \geqq d \\ & x_j = 0 \text{ or } 1 \text{ for all } j \end{array}$$

where $x \in \mathbf{R}^n$, and c, A, b, D, d are the data in the problem. In the interval $0 \leqq x_j \leqq 1$, the function $x_j(1 - x_j)$ is non-negative, and is zero iff x_j is either 0 or 1. Using this we can transfom the above discrete problem into a continuous variable optimization by a penalty transformation as given below

$$\begin{array}{ll} \text{minimize} & cx + \alpha(\sum_{j=1}^{n} x_j(1 - x_j)) \\ \text{subject to} & Ax = b \\ & Dx \geqq d \\ & 0 \leqq x_j \leqq 1, \ j = 1 \text{ to } n \end{array}$$

where α is a large positive penalty parameter. This is now a quadratic programming problem (unfortunately, it is a concave minimization problem and may have lots of local minima, in fact it can be verified that every integer feasible solution is a local minima for this problem). Check whether any of the algorithm for LCP discussed here are useful to approach the integer program through the LCP formulation of the above quadratic program.

2.31 Consider the system

$$w - Mz = q$$
$$w, z \geqq 0$$

where M is a given square matrix of order n. Let $\mathbf{C}_1$ be the set of feasible solutions of this problem satisfying the additional conditions

$$w_j z_j = 0, \quad j = 2 \text{ to } n \ .$$

Assuming that q is nondegenerate in this system (i.e., that in every solution (w, z) of the system of equations "$w - Mz = q$", at last n variables are non-zero), study whether $\mathbf{C}_1$ can contain an edge path terminating with extreme half-lines at both ends, when M is a copositive plus matrix.

2.32 (Research Problem) : Consider the general quadratic programming problem (2.67) of Section 2.9.2, and let $\mathbf{K}$ be its set of feasible solutions.

Develop necessary and sufficient conditions for $Q(x)$ to be unbounded below on $\mathbf{K}$. Develop an efficient procedure to check whether $Q(x)$ is unbounded below on $\mathbf{K}$.

In (2.67), the objective function is said to be strongly unbounded below, if it remains unbounded below whatever the vector c may be, as long as all the other data in the problem remains unchanged. Develop necessary and sufficient conditions for and an efficient procedure to check this strong unboundedness.

Extend the enumeration procedure for solving the general quadratic programming problem under the assumption of a bounded feasible set discussed in Section 2.9, to the case when $\mathbf{K}$ is unbounded.

The method discussed in Section 2.9 for solving this problem, may be viewed as a total enumeration method (enumerating over all the faces of $\mathbf{K}$). Develop an efficient method for computing a lower bound for $Q(x)$ on $\mathbf{K}$, and using it, develop a branch and bound method for solving this problem (this will be an efficient partial enumeration method). (See B. C. Eaves [2.9] for some useful information on this problem.)

2.33 Let M be a square matrix of order n which is $D + E$ where

D is symmetric and copositive plus

E is copositive.

Let $q \in \mathbf{R}_n$. If the system $Dx - E^T y \geqq -q$, $y \geqq 0$ is feasible, prove that the complementary pivot algorithm will terminate with a solution when applied on the LCP (q, M).
(P. C. Jones [2.17])

2.34 Let $(\bar{w}, \bar{z})$ be the solution of the LCP (q, M).

i) If M is PSD, prove that $\bar{z}^T q \leqq 0$.

ii) If the LCP (q, M) comes from an LP prove that $\bar{z}^T q = 0$.

2.35 Prove that if M is a copositive plus matrix of order n, and $q \in \mathbf{R}^n$ then the optimum objective value in the following quadratic program is zero, if the problem has a feasible solution.

$$\begin{array}{lll} \text{minimize} & Q(x) = x^T(Mx + q) & \\ \text{subject to} & Mx + q & \geqq 0 \\ & x & \geqq 0 \end{array}$$

2.36 In Section 2.9.2, we have seen that if a quadratic function $Q(x)$ is bounded below on a convex polyhedron, then $Q(x)$ has a finite global minimum point on that polyhedron. Does this result hold for a general polynomial function ?
(Hint: Examine the fourth degree polynomial function $f(x) = x_1^2 + (x_1 x_2 - 1)^2$ defined over $\mathbf{R}^2$).
(L. M. Kelly)

2.37 Apply the Complementary pivot method on the LCP with the following data.

a) $$q = \begin{pmatrix} -4 \\ -5 \\ -1 \end{pmatrix}, \quad M = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$

b) $$q = \begin{pmatrix} -1 \\ -2 \\ -3 \end{pmatrix}, \quad M = \begin{pmatrix} 1 & 2 & 0 \\ -2 & -1 & 0 \\ -1 & -3 & -1 \end{pmatrix}$$

c) $$q = \begin{pmatrix} -1 \\ -2 \\ -3 \end{pmatrix}, \quad M = \begin{pmatrix} -1 & 2 & -2 \\ 2 & -1 & 2 \\ -2 & 2 & -1 \end{pmatrix}.$$

Verify that (z_1, z_2, z_3) is a complementary feasible basic vector for (c).

Also, solve (a) by the variant of the complementary pivot method discussed in Section 2.4.

2.11 References

2.1 I. Adler, "The Expected Number of Pivots Needed to Solve Parametric Linear Programs and the Efficiency of the Self-dual Simplex Method", Department of Industrial Engineering and Operations Research, University of California, Berkeley, California (May 1983).

2.2 V. Aggarval, "On the Generation of all Equilibrium Points for Bimatrix Games Through the Lemke-Howson Algorithm", *Mathematical Programming*, 4 (1973) 233–234.

2.3 F. A. Al-Khayyal, "Linear, Quadratic and Bilinear Programming Approaches to the Linear Complementary Problem", Industrial and Systems Engineering Report No.J–83–5, Georgia Institute of Technology, Atlanta, Georgia (September 1983).

2.4 J. C. G. Boot, "Quadratic Programming", North-Holland, Amsterdam, The Netherlands (1964).

2.5 R. W. Cottle, "Symmetric Dual Quadratic Programs", *Quarterly of Applied Mathematics*, 21 (1963), 237–243.

2.6 G. B. Dantzig & A. Manne, "A Complementarity Algorithm for an Optimal Capital Path with Invariant Proportions", *Journal of Economic Theory*, 9 (1974) 312–323.

2.7 W. S. Dorn, "Duality in Quadratic Programming", *Quaterly of Applied Mathematics*, 18 (1960) 155–162.

2.8 B. C. Eaves, "The Linear Complementarity Problem", Management Science, 17, 9 (May 1971) 612–634.

2.9 B. C. Eaves, "On Quadratic Programming", Management Science, 17, 11 (July 1971) 698–711.

2.10 B. C. Eaves & H. Scarf, "The solution of Systems of Piecewise Linear Equations", *Mathematics of Operations Research*, 1 (1976) 1–27.

2.11 J. J. Evers, "More With Lemke Complementarity Algorithm", *Mathematical Programming*, 15 (1978) 214–219.

2.12 P. E. Gill, N. I. M. Gould, W. Murray, M. A. Saunders & M. H. Wright, "A Range-space Method for Quadratic Programming with Bounds and General Constraints", Technical Report SOL 82–15, Department of Operations Research, Stanford University, Stanford, California (October 1982).

2.13 P. E. Gill & W. Murray, "Numerically Stable Methods for Quadratic Programming", *Mathematical Programming*, 14 (1978) 349–372.

2.14 F. J. Gould & J. W. Tolle, "A Unified Approach to Complementarity in Optimization", *Discrete Mathematics*, 7 (1974) 225–271.

2.15 W. P. Hallman & I. Kaneko, "On the Connectedness of the Set of Almost Complementary Paths in a Linear Complementarity Problem", *Mathematical Programming*, 16 (1979) 384–385.

2.16 P. C. Jones, "Computing an Optimal Invariant Capital Stock", *SIAM Journal of Algebraic and Discrete Methods*, 3 (No.2)(1982) 145–150.

2.17 P. C. Jones, "Even more with the Lemke complementary algorithm", Technical Report, Department of IEMS, Northwestern University, Evanston, IL 60201, December 1984.

2.18 E. L. Keller, "Quadratic Optimization and Linear Complementarity", Ph.D. Dissertation, University of Michigan (1969).

2.19 E. L. Keller, "The General Quadratic Optimization Problem", *Mathematical Programming*, 5 (1973) 311–337.

2.20 M. M. Kostreva, "Cycling in Linear Complementarity Problems", *Mathematical Programming*, 16 (1979) 127–130.

2.21 C. E. Lemke, "Bimatrix Equilibrium Points and Mathematical Programming", *Management Science*, 11 (1965) 681–689.

2.22 C. E. Lemke, "Some Pivot Schemes for the Linear Complementarity Problem", *Mathematical Programming Study*, 7 (1978) 15–35.

2.23 R. K. Mueller, "A Method for Solving the Indefinite Quadratic Programming Problem", *Management Science (Theory)*, 16 (No.5)(1970) 333–339.

2.24 K. G. Murty, "On the Finite Nature of Quadratic Programming: A Note on R. K. Mueller's Paper", Technical Report 71–5, Department of Industrial and Operations Engineering, University of Michigan, Ann Arbor, Michigan (1971).

2.25 K. G. Murty, "Algorithm for Finding all the Feasible Complementary Bases for a Linear Complementarity Problem", Technical Report No. 72–2, Department IOE, University of Michigan, Ann Arbor, Michigan (February 1972).

2.26 K. G. Murty, *Linear Programming,* John Wiley & Sons, New York (1983).

2.27 H. Nishino, M. Kojima & T. Sekine, "An Extension of Lemke's Method to the Piecewise Linear Complementarity Problem", *SIAM Journal on Applied Mathematics*, 31 (1976) 600–613.

2.28 T. D. Parsons, "A Combinatorial Approach to Convex Quadratic Programming", Ph. D. Dissertation, Department of Mathematics, Princeton University (1966).

2.29 F. J. Pereira R., "On Characterizations of Copostive Matrices", Technical Report No. 72–8, OR House, Stanford University, Stanford, California (May 1972).

2.30 A. Ravindran, "A Comparison of the Primal Simplex and Complementary Pivot Methods for Linear Programming", *Naval Research Logistics Quarterly*, 20 (No.1) (1973) 95–100.

2.31 R. Saigal, "On Some Average Results for Linear Complementarity Problems", Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, Illinois (August 1983).

2.32 R. Saigal, "Lemke's Algorithm and a Special Linear Complementarity Problem", *Opsearch* 8 (1971) 201–208.

2.33 L. S. Shapley, "A Note on the Lemke-Howson Algorithm", *Mathematical Programming Study 1* (1974) 175–189.

2.34 M. J. Todd, "Extensions of Lemke's Algorithm for the Linear Complementarity Problem", *Journal of Optimization Theory and Applications*, 20 (1976) 397–416.

2.35 M. J. Todd, "Linear and Quadratic Programming in Oriented Matroids", Technical Report No.565, School of Operations Research and Industrial Engineering, Cornell University, Ithaca, New York (1983).

2.36 M. J. Todd, "Polynomial Expected Behavior of a Pivoting Algorithm for Linear Complementarity and Linear Programming Problems", Technical Report No.595, School of Operations Research and Industrial Engineering, Cornell University, Ithaca, New York (1983).

2.37 C. Van de Panne, "A Complementary Variant of Lemke's Method for the Linear Complementarity Problem", *Mathematical Programming*, 7 (1974) 283–310.

2.38 L. Van der Heyden, "A Variable Dimension Algorithm for the Linear Complementarity Problem", *Mathematical Programming*, 19 (1980) 328–346.

2.39 L. Van der Heyden, "On a Variable Dimension Algorithm for the Linear Complementarity Problem", Cowles Foundation Discussion Paper No.689 (January 1984).

References on Fixed Point Computing Methods

2.40 E. Allgower & K. Georg, "Simplicial and Continuation Methods for Approximating Fixed Points and Solutions to Systems of Equations", *SIAM Review*, 22 (1980) 28–85.

2.41 M. N. Broadie, "Computational Experience with the Octahedral Algorithm and Related Results", Department of OR, Stanford University, Stanford, California (March 1983).

2.42 S. N. Chow, J. Mallet-Paret & J. A. Torke, "Finding Zeros of Maps: Homotopy Methods that are Constructive with Probability One", *Math. Comput.*; 32, pp. 887–899, 1978.

2.43 B. C. Eaves, "Computing Kakutani Fixed Points", *SIAM Journal of Applied Mathematics*, 21, pp. 236–244, 1971.

2.44 B. C. Eaves, "Homotopies for Computation of Fixed Points", *Mathematical Programming* 3 (1972) 1–22.

2.45 B. C. Eaves, *A course in triangulations for solving equations with deformations,* Lecture Notes in Economics and Mathematical Systems 293, Springer Verlag, New York, 1984.

2.46 B. C. Eaves, F. J. Gould, H. O. Peitgen & M. J. Todd (editors), *Homotopy Methods and Global Convergence,* Plenum Press, New York (1983).

2.47 B. C. Eaves & R. Saigal, "Homotopies for Computation of Fixed Points on Unbounded Regions", *Mathematical Programming*, 3 (1972) 225–237.

2.48 C. B. Garcia & W. I. Zangwill, *Pathways to Solutions, Fixed Points and Equilibria,* Prentice-Hall, Englewood Cliffs, New Jersey (1981).

2.49 F. J. Gould & J. W. Tolle, *Complementary Pivoting on a Pseudo-manifold Structure with Applications in the Decision Sciences,* Heldermann Verlag, Berlin (1983).

2.50 S. Kakutani, "A Generalization of Brouwer's Fixed Point Theorem", *Duke Mathematical Journal*, 8 (1941) 457–459.

2.51 S. Karamardian & C. B. Garcia (editors), *Fixed Point Algorithms and Applications,* Academic Press, New York (1977).

2.52 R. B. Kellog, T. Y. Li & J. Yorke, "A Constructive Proof of the Brouwer Fixed Point Theorem and Computational Results", *SIAM Journal on Numerical Analysis*, 13, pp. 473–483, 1976.

2.53 M. Kojima & Y. Yamamoto, "A unified Approach to the Implementation of Several Restart Fixed Point Algorithms and a New Variable Dimension Algorithm", *Mathematical Programming* 28 (1984) 288–328.

2.54 H. W. Kuhn, "Approximate Search for Fixed Points", *Computing Methods in Optimization Problems*, 2, Academic Press, New York (1969).

2.55 H. J. Luethi, "A Simplicial Approximation of a Solution for the Nonlinear Complementarity Problem", *Mathematical Programming*, 9 (1975) 278–293.

2.56 O. L. Mangasarian, "Equivalence of the Complementarity Problem to a System of Nonlinear Equations", *SIAM Journal of Applied Mathematics*, 31 (No.1) (1976) 89–92.

2.57 O. H. Merrill, "A Summary of Techniques for Computing Fixed Points of Continuous Mappings", *Mathematical Topics in Economic Theory and Computation*, SIAM, Philadelphia, Pennsylvania (1972).

2.58 O. H. Merrill, "Application and Extensions of an Algorithm that Computes Fixed Points of Upper Semicontinuous Point to Set Mappings", Ph. D. Dissertation, Department of Industrial and Operations Engineering, The University of Michigan (1972). Available from University Microfilms, 300 N. Zeeb Road, Ann Arbor, Michigan, 48106, USA, under number 73–6876.

2.59 A. Netravali & R. Saigal, "Optimum Quantizer Design Using a Fixed Point Algorithm", *The Bell System Technical Journal*, 55 (No.9) (November 1976) 1423–1435.

2.60 J. M. Ortega & C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables,* Academic Press, New York (1970).

2.61 S. M. Robinson (Editor), *Analysis and Computation of Fixed Points,* Academic Press, (1980).

2.62 R. Saigal, D. Solow & L. Wolsey, "A Comparative Study of Two Algorithms that Compute Fixed Points in Unbounded Regions", VIII Internationsl Symposium of Mathematical Programming, Stanford University, Stanford, California (August 1973).

2.63 R. Saigal, "Fixed Point Coumputing Methods", in J. Belzer, A. G. Holzman and A. Kent (editors) *Encyclopedia of Computer Science and Technology*, Volume 8, Marcel Dekker, Inc., New York (1976) 213–235.

2.64 R. Saigal, "On the Convergence Rate of Algorithm for Solving Equations that are Based on Methods of Complementary Pivoting", *Mathematics of Operations Research*, 2 (1977) 108–124.

2.65 R. Saigal & M. J. Todd, "Efficient Acceleration Techniques for Fixed Point Algorithms", *SIAM Journal on Numerical Mathematics*, 15 (1978) 997–1007.

2.66 R. Saigal, "A Homotopy for Solving Large, Sparse and Structured Fixed Point Problems", *Mathematics of Operations Research*, 8 (No.4) (1983) 557–578.

2.67 R. Saigal, "Computational Complexity of a Piecewise Linear Homotopy Algorithm", *Mathematical Programming*, 28 (1984) 164–173.

2.68 H. E. Scarf, "The Approximation of Fixed Points of Continuous Mappings", *SIAM Journal of Applied Mathematics*, 15 (1967) 1328–1343.

2.69 H. E. Scarf (in collaboration with T. Hansen), *The Computation of Economic Equilibria,* Yale University Press, New Haven, Connecticut (1973).

2.70 D. Solow, "Homeomorphisms of Triangulations with Applications to Computing Fixed Points", *Mathematical Programming*, 20, 2, pp. 213–224, March 1981.

2.71 M. J. Todd, "On Triangulations for Computing Fixed Points", *Mathematical Programming*, 10 (1976a) 322–346.

2.72 M. J. Todd, *The Computation of Fixed Points and Applications,* Lecture Notes in Economics and Mathematical Systems 124, Springer Verlag, New York (1976b).

2.73 M. J. Todd, "Exploiting Structure in Piecewise-Linear Homotopy Algorithms for Solving Equations", *Mathematical Programming*, 18 (1980) 233–247.

2.74 M. J. Todd, "On the Computational Complexity of Piecewise Linear Homotopy Algorithms", *Mathematical Programming*, 24 (1982) 216–224.

2.75 M. J. Todd, "J': A New Triangulation of $\mathbf{R}^n$", *SIAM Journal on Algebraic and Discrete Methods*, 5 No.2 (June 1984) 244–254.

2.76 M. J. Todd, "Fat Triangulations, or solving certain nonconvex matrix optimization problems", *Mathematical Programming*, 31 (1985) 123–136.

2.77 G. Van der Laan & A. J. J. Talman, "A Class of Simplicial Restart Fixed Point Algorithms Without an Extra Dimension", *Mathematical Programming*, 20 (1981) 33–48.

2.78 L. T. Watson, "A Globally Convergent Algorithm for Computing Fixed Points of C^2 Maps", *Applied Mathematics and Computation*, 5, pp. 297–311, 1979.

2.79 R. J. Wilmuth, "The Computation of Fixed Points", Ph. D. Dissertation, Department of OR, Stanford University, Stanford, California (1973).

2.80 A. H. Wright, "The Octahedral Algorithm: A New Simplicial Fixed Point Algorithm", *Mathematical Programming*, 21 (1981) 47–69.

References on Available Computer Codes

2.81 M. N. Broadie, "Octasolve User's Guide", Department of OR, Stanford University, Stanford, Californis (1983).

2.82 R. L. Crane, K. E. Hillstrom & M. Minkoff, "Solution of the General Nonlinear Programming Problem with Subroutine VMCON", ANL80–64, Argonne National Laboratory, Illinois 60439 (July 1980).

2.83 P. E. Gill, W. Murray, M. A. Saunders & M. H. Wright, "User's Guide for SOL/QPSOL", Technical Report SOL 83–7, Department of OR, Stanford University, Stanford, California (1983).

2.84 P. E. Gill, W. Murray, M. A. Saunders & M. H. Wright, "User's Guide for SOL/NPSOL: A Fortran Package for Nonlinear Programming", Technical Report SOL 83–12, Department of OR, Stanford University, Stanford, California (July 1983).

2.85 B. A. Murtagh & M. A. Saunders, "MINOS: A Large Scale Nonlinear Programming System for Problems with Linear Constraints, User's Guide", Technical Report SOL 77–9, Department of OR, Stanford University, Stanford, California 94305 (February 1977). Also, see "MINOS 5.0 User's Guide", Technical Report SOL 83–20, (December 1983).

2.86 R. Saigal, "Efficient Algorithms for Computing Fixed Points When Mappings May be Separable, A Computer Program and User's Guide", Department of Industrial Engineering, Northwestern University, Evanston, Illinois (1979).

2.87 K. Schittkowski, "User's Guide for the Nonlinear Programming Code NLPQL", Institut fuer Informatik, Universitaet Stuttgart, Azenbergstr. 12, D-7000 Stuttgart 1, Federal Republic of Germany, 1983.

2.88 M. J. Todd, "PLALGO: A Fortran Implementation of a Piecewise-linear Homotopy Algorithm for Solving Systems of Nonlinear Equations", Technical Report No.452, School of OR and IE, Cornell University, Ithaca, New York (1981).

2.89 J. Tomlin, "Robust Implementation of Lemke's Method for the Linear Complementarity Problem", *Mathematical Programming Study* 7 (1978) 55–60.

2.90 J. Tomlin, "User's Guide to LCPL, A Program for Solving Linear Complementarity Problems by Lemke's Method", Technical Report SOL76–16, Department of Operations Research, Stanford University, Stanford, California (1976).

2.91 L. T. Watson, S. C. Billups & A. P. Morgan, "HOMPACK: A suite of codes for globally convergent homotopy algorithms", Technical Report 85–34, Industrial and Operations Engineering Department, The University of Michigan, Ann Arbor, Michigan 48109, USA, November 1985.

References on Convexity and Convex Programming

2.92 R. T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, New Jersey (1970).

2.93 R. T. Rockafellar, "The Theory of Subgradients and Its Applications on Problems of Optimization: Convex and Nonconvex Functions", Heldermann Verlag, Berlin (1981).

2.94 J. Stoer & C. Witzgall, "Convexity and Optimization in Finite Dimensions I", Springer Verlag, Berlin (1970).

Chapter 3

SEPARATION PROPERTIES, PRINCIPAL PIVOT TRANSFORMS, CLASSES OF MATRICES

In this chapter we present the basic mathematical results on the LCP. Many of these results are used in later chapters to develop algorithms to solve LCPs, and to study the computational complexity of these algorithms. Here, unless stated otherwise, I denotes the unit matrix of order n. M is a given square matrix of order n. In tabular form the LCP (q, M) is

w	z	q
I	$-M$	q

$$w \geqq 0, \qquad z \geqq 0, \qquad w^T z = 0 \tag{3.1}$$

Definition: Subcomplementary Sets of Column Vectors

A vector $(y_1, \ldots, y_{i-1}, y_{i+1}, \ldots, y_n)$ where $y_r \in \{w_r, z_r\}$ for $r = 1, \ldots, i-1, i+1, \ldots, n$ is known as a **subcomplementary vector of variables** for the LCP (3.1). The complementary pair (w_i, z_i) is known as the **left-out complementary pair of variables** in the subcomplementary vector $(y_1, \ldots, y_{i-1}, y_{i+1}, \ldots, y_n)$. Let $A_{.j}$ be the column vector associated with y_j in (3.1). The ordered set $(A_{.1}, \ldots, A_{.i-1}, A_{.i+1}, \ldots, A_{.n})$ is known as a **subcomplementary set of column vectors** for the LCP (3.1), and $(I_{.i}, -M_{.i})$ is the **left-out complementary pair of column vectors** in this subcomplementary set of column vectors.

Sometimes we have to refer to subcomplementary sets which are complementary sets with several elements missing. For this, we adopt the following notation. Let $\mathbf{J} \subset \{1, \ldots, n\}$, $\mathbf{J} \neq \emptyset$, $\mathbf{J}$ a proper subset. The vector $(y_j : j \in \mathbf{J})$ where $y_j \in \{w_j, z_j\}$

for all $j \in \mathbf{J}$ is said to be a subcomplementary vector of variables for (3.1) associated with the subset $\mathbf{J}$. Let t_j be the complement of y_j and let $A_{.j}$ be the column vector associated with y_j in (3.1), and let $B_{.j}$ be the complement of $A_{.j}$, for $j \in \mathbf{J}$. Then $\{A_{.j} : j \in \mathbf{J}\}$ is said to be a subcomplementary set of column vectors associated with $\mathbf{J}$, and $\{B_{.j} : j \in \mathbf{J}\}$ is its complement. The subcomplementary vector $(t_j : j \in \mathbf{J})$ is the complement of the subcomplementary vector $(y_j : j \in \mathbf{J})$.

3.1 LCPs ASSOCIATED WITH PRINCIPALLY NONDEGENERATE MATRICES

If $y = (y_1, \ldots, y_n)$ is a complementary vector of variables for (3.1), define

$$\begin{aligned} \mathbf{Z}(y) &= \{j : j \text{ such that } y_j = z_j\} \\ \mathbf{W}(y) &= \{j : j \text{ such that } y_j = w_j\} . \end{aligned} \tag{3.2}$$

Theorem 3.1 *If y is a complementary vector of variables for (3.1), it is a complementary basic vector iff the principal subdeterminant of M corresponding to the subset $\mathbf{Z}(y)$ is nonzero.*

Proof. Let the cardinality of $\mathbf{Z}(y)$ be r. Let A be the complementary matrix associated with y. For $j \in \mathbf{W}(y)$, $A_{.j} = I_{.j}$ and for $j \in \mathbf{Z}(y)$, $A_{.j} = -M_{.j}$. If $r = 0$, $A = I$ and its determinant is 1. If $r > 0$, by expanding the determinant of A in terms of its elements in the jth column for each $j \in \mathbf{W}(y)$ in some order, we see that the determinant of A is $(-1)^r$ (principal subdeterminant of M corresponding to the subset $\mathbf{Z}(y)$). Since y is a complementary basic vector iff the determinant of A is nonzero, the result follows. □

As an example, let $n = 4$, and consider the LCP (q, M). Let $y = (w_1, z_2, w_3, z_4)$ be a complementary vector of variables for this problem. The corresponding complementary matrix is

$$\begin{pmatrix} 1 & -m_{12} & 0 & -m_{14} \\ 0 & -m_{22} & 0 & -m_{24} \\ 0 & -m_{32} & 1 & -m_{34} \\ 0 & -m_{42} & 0 & -m_{44} \end{pmatrix}$$

and its determinant is determinant $\begin{pmatrix} -m_{22} & -m_{24} \\ -m_{42} & -m_{44} \end{pmatrix}$, which is non-zero iff the principal subdeterminant of M corresponding to the subset $\mathbf{Z}(y) = \{2, 4\}$ is non-zero. Thus, in this problem, y is a complementary basic vector iff the principal subdeterminant of M corresponding to the subset $\mathbf{Z}(y)$ is non-zero.

Corollary 3.1 *Every complementary vector of variables is a basic vector for (3.1) iff M is a nondegenerate matrix. This follows from Theorem 3.1 and the definition of nondegeneracy of a matrix.*

Corollary 3.2 *The complementary cone associated with the complementary vector of variables y for (3.1) has a nonempty interior iff the principal subdeterminant of M corresponding to the subset $\mathbf{Z}(y)$ is nonzero.*

Proof. If A is the corresponding complementary matrix, the complementary cone is $\text{Pos}(A)$, and it has nonempty interior iff the determinant of A is nonzero. So the result follows from Theorem 3.1.

□

Corollary 3.3 *Every complementary cone in the class $\mathcal{C}(M)$ has a nonempty interior iff M is a nondegenerate matrix. This follows from Corollary 3.2.*

Theorem 3.2 *The LCP (q, M) has a finite number of solutions for each $q \in \mathbf{R}^n$ iff M is a nondegenerate matrix.*

Proof. Let $(\hat{w}, \hat{z})$ be a solution of the LCP (q, M). Let $A_{.j} = -M_{.j}$ if $\hat{z}_j > 0$, $I_{.j}$ otherwise; and $\alpha_j = \hat{z}_j$ if $\hat{z}_j > 0$, $\hat{w}_j$ otherwise. Then $(A_{.1}, \ldots, A_{.n})$ is a complementary set of column vectors and $q = \sum_{j=1}^{n} \alpha_j A_{.j}$. In this manner each solution of the LCP (q, M) provides an expression of q as a nonnegative linear combination of a complementary set of column vectors. There are only $2^{\mathbf{n}}$ complementary sets of column vectors. If $q \in \mathbf{R}^n$ is such that the LCP (q, M) has an infinite number of distinct solutions, there must exist a complementary set of column vectors, say $(A_{.1}, \ldots, A_{.n})$, such that q can be expressed as a nonnegative linear combination of it in an infinite number of ways. So there exist at least two vectors $\alpha^t = (\alpha_1^t, \ldots, \alpha_n^t)^T \geqq 0$, $t = 1, 2$ such that $\alpha^1 \neq \alpha^2$ and $q = A\alpha^1 = A\alpha^2$. So $A(\alpha^1 - \alpha^2) = 0$, and since $\alpha^1 \neq \alpha^2$, $\{A_{.1}, \ldots, A_{.n}\}$ is linearly dependent. By Theorem 3.1, this implies that M is degenerate.

Conversely suppose M is degenerate. So, by Theorem 3.1, there exists a complementary set of column vectors, say $\{A_{.1}, \ldots, A_{.n}\}$ which is linearly dependent. So there exists a $\beta = (\beta_1, \ldots, \beta_n) \neq 0$ such that $\sum_{j=1}^{n} \beta_j A_{.j} = 0$. Let $\delta = \text{Maximum } \{|\beta_j| : j = 1 \text{ to } n\}$. Since $\beta \neq 0$, $\delta > 0$. Define $\overline{q} = \delta \sum_{j=1}^{n} A_{.j}$. Let $(y_1, \ldots, y_n)$ be the complementary vector associated with $(A_{.1}, \ldots, A_{.n})$. Define a solution $(w(\lambda), z(\lambda))$ by

$$\begin{array}{ll} \text{Complement of} & y_j = 0,\ j = 1 \text{ to } n \\ & y_j = \delta + \lambda\beta_j,\ j = 1 \text{ to } n\,. \end{array} \tag{3.3}$$

Then $(w(\lambda), z(\lambda))$ is a solution of the LCP $(\overline{q}, M)$ for each $0 \leqq \lambda \leqq 1$, and since $\beta \neq 0$, each of these solutions is distinct. So if M is degenerate, there exist a $q \in \mathbf{R}^n$ such that the LCP (q, M) has an infinite number of distinct solutions.

□

Example 3.1

Consider the following LCP

w_1	w_2	z_1	z_2	
1	0	-1	-1	-2
0	1	-1	-1	-2
$w_1, w_2, z_1, z_2 \geqq 0$, $w_1 z_1 = w_2 z_2 = 0$				

We have

$$\begin{aligned} q = (-2,-2)^T &= (-M_{.1}) + (-M_{.2}) \\ 0 &= (-M_{.1}) - (-M_{.2}) \,. \end{aligned}$$

These facts imply that $(w_1, w_2; z_1, z_2) = (0, 0; 1+\theta, 1-\theta)^T$ is a complementary solution to this LCP for all $0 \leqq \theta \leqq 1$.

The set of q for which the number of complementary solutions for the LCP (q, M) is infinite, is always a subset of the union of all degenerate complementary cones. Also if the LCP (q, M) has an infinite number of complementary solutions, q must be degenerate in it (that is, q can be expressed as a linear combination of $(m-1)$ or less column vectors of $(I : -M)$).

Result 3.1 If q is nondegenerate in the LCP (q, M) of order n (that is, if in every solution to the system of equations $w - Mz = q$, at least n of the variables in the system are non-zero), every complementary solution of the LCP (q, M) must be a complementary BFS, and so the number of complementary solutions to the LCP (q, M) is finite and $\leqq 2^{\mathbf{n}}$.

Proof. In every complementary solution of the LCP (q, M) at most n variables can be positive by the complementarity constraint, and hence exactly n variables have to be positive by the nondegeneracy of q, that is one variable from every complementary pair of variables must be strictly positive. Consider a complementary solution $(\overline{w}, \overline{z})$ in which the positive variable from the complementary pair $\{w_j, z_j\}$ is y_j say, for $j = 1$ to n and suppose y_j has value $\overline{y}_j > 0$ in the solution. Let $A_{.j} = I_{.j}$ if $y_j = w_j$, or $-M_{.j}$ otherwise. So

$$q = \sum_{j=1}^{n} \overline{y}_j A_{.j} \,.$$

If $\{A_{.1}, \ldots, A_{.n}\}$ is linearly dependent, let the linear dependence relation be

$$0 = \sum_{j=1}^{n} \alpha_j A_{.j}$$

where $\alpha = (\alpha_1, \ldots, \alpha_n)^T \neq 0$. Suppose $\alpha_1 \neq 0$. Let $\lambda = -(\overline{y}_1/\alpha_1)$, then $\overline{y}_1 + \lambda\alpha_1 = 0$. From the above two equations, we have

$$q = \sum_{j=1}^{n} (\overline{y}_j + \lambda\alpha_j) A_{.j} = \sum_{j=2}^{n} (\overline{y}_j + \lambda\alpha_j) A_{.j}$$

that is, q is expressed as a linear combination of $\{A_{.2}, \ldots, A_{.n}\}$ which is a subset of $n-1$ columns of $(I \vdots -M)$, contradicting the nondegeneracy of q. So $\{A_{.1}, \ldots, A_{.n}\}$ must be linearly independent, that is $A = (A_{.1} \vdots \cdots \vdots A_{.n})$ is a complementary basis, and hence the representation of q as a linear combination of the columns of A is unique, and $(\overline{w}, \overline{z})$ is a complementary BFS. Thus under the nondegeneracy assumption of q, every complementary solution for the LCP (q, M) must be a complementary BFS. Since the total number of complementary bases is $\leqq 2^{\mathbf{n}}$, this implies that there are at most 2^n complementary solutions in this case.

□

3.2 PRINCIPAL PIVOT TRANSFORMS

Let $y = (y_j)$ be a complementary basic vector associated with the complementary basis A for (3.1). Let t_j be the complement of y_j for $j = 1$ to n (i. e., $t_j = w_j$ if $y_j = z_j$, $t_j = z_j$ if $y_j = w_j$). Let $B_{.j}$ be the complement of $A_{.j}$ for $j = 1$ to n, and $B = (B_{.1}, \ldots, B_{.n})$. Obtain the canonical tableau of (3.1) with respect to the basic vector y, and after rearranging the variables suppose it is

basic vector	$y_1 \ldots y_n$	$t_1 \ldots t_n$	
y	I	$-D$	$\overline{q}$

(3.4)

Then the matrix D is known as the **principal pivot transform** (PPT in abbreviation) of M associated with the complementary basic vector y or the corresponding complementary basis A of (3.1). Clearly $D = -A^{-1}B$. Also (3.4) can be viewed as the system of equations of an LCP in which the complementary pairs are (y_j, t_j), $j = 1$ to n. Remembering that the variables in (3.4) are just the variables in (3.1) arranged in a different order, we can verify that the canonical tableau of (3.4) with respect to its basic vector $(w_1, \ldots, w_n)$ is (3.1). This clearly implies that M is a PPT of D. Hence the property of being a PPT is a mutual symmetric relationship among square matrices of the same order.

Example 3.2

Consider the LCP (q, M) where

$$M = \begin{pmatrix} -1 & -2 & 0 & -1 \\ -1 & 1 & -1 & -2 \\ 0 & -1 & 1 & -1 \\ 0 & -2 & 0 & 2 \end{pmatrix}.$$

The LCP (q, M) is

w_1	w_2	w_3	w_4	z_1	z_2	z_3	z_4	
1	0	0	0	1	2	0	1	q_1
0	1	0	0	1	−1	1	2	q_2
0	0	1	0	0	1	−1	1	q_3
0	0	0	1	0	2	0	−2	q_4
				$w_j, z_j \geqq 0$, $w_j z_j = 0$ for all j.				

(z_1, w_2, z_3, w_4) is a complementary basic vector for this problem. The canonical tableau with respect to it is

z_1	w_2	z_3	w_4	w_1	z_2	w_3	z_4	
1	0	0	0	1	2	0	1	q'_1
0	1	0	0	−1	−2	1	2	q'_2
0	0	1	0	0	−1	−1	−1	q'_3
0	0	0	1	0	2	0	−2	q'_4

Thus the matrix

$$D = \begin{pmatrix} -1 & -2 & 0 & -1 \\ 1 & 2 & -1 & -2 \\ 0 & 1 & 1 & 1 \\ 0 & -2 & 0 & 2 \end{pmatrix}$$

is a PPT of M and vice versa.

Each complementary basic vector for (3.1) leads to a PPT of M. We thus get a class of matrices containing M, such that each matrix in the class is a PPT of each other matrix in the class. Some of the matrices in the class may be equal to the others as matrices (for example, it can be verified that every PPT of I is equal to I). This class of matrices is known as the **principal pivot transform class of** M.

Single and Double Principal Pivot Steps

If $y = (y_1, \ldots, y_n)$ is a complementary basic vector for (3.1), then y_r can be replaced in this basic vector by its complement, to yield another complementary basic vector for (3.1), iff the rth diagonal element in the PPT of M corresponding to y is nonzero. If this condition is satisfied, the pivot operation of replacing y_r by its complement, is known as a **single principal pivot step in the rth position in the complementary basic vector y.**

Suppose for $r \neq s$, the rth and sth diagonal elements in $M' = (m'_{ij})$, the PPT of M corresponding to the complementary basic vector y, are both zero. Then it is not possible to make a single principal pivot step either in the rth position, or in the sth position, in the complementary basic vector y. However, suppose $m'_{rs} \neq 0$ and $m'_{sr} \neq 0$. In this case we can perform two consecutive pivot steps, in the first one replacing y_r in the basic vector by the complement of y_s, and in the second one replacing y_s in the resulting basic vector by the complement of y_r. In the canonical tableau obtained at the end of these two pivot steps, the column vector associated with the complement of y_s is $I_{.r}$ and the column vector associated with the complement of y_r is $I_{.s}$. So, now interchange rows r and s in the canonical tableau. After this interchange it can be verified that in the new canonical tableau the column vector associated with the basic variable from the jth complementary pair, in the new complementary basic vector, is $I_{.j}$, for all j (including $j = r$ and s). This operation (one pivot step in position (r, s) replacing y_r in the basic vector by the complement of y_s, followed by another pivot step in position (s, r) replacing y_s in the resulting basic vector by the complement of y_r, followed by an interchange of rows r and s in the resulting canonical tableau) is called a **double principal pivot step in positions r and s in the complementary basic vector y.** Clearly, this double principal pivot step in positions r and s can only be carried out if the order two determinant $\begin{pmatrix} m'_{rr} & m'_{rs} \\ m'_{sr} & m'_{ss} \end{pmatrix} \neq 0$. If this order two determinant is nonzero, and one of its diagonal entries, say m'_{rr}, is nonzero; carrying out the double principal pivot in positions r and s in the complementary basic vector y, can be verified to have exactly the same effect as carrying out two single principal pivot steps, first in position r in y, and then in position s in the complementary basic vector resulting from the first. In general, in the algorithms discussed in the following chapters, a double principal pivot in positions r and s will only be performed if the diagonal entry in the PPT of M in at least one of the two positions r and s is zero (i. e., either $m'_{rr} = 0$ or $m'_{ss} = 0$ or both).

Example 3.3

Consider the following LCP

basic variable	w_1	w_2	w_3	w_4	z_1	z_2	z_3	z_4	
w_1	1	0	0	0	-1	1	-1	-1	q_1
w_2	0	1	0	0	-1	0	0	1	q_2
w_3	0	0	1	0	0	0	-1	-1	q_3
w_4	0	0	0	1	1	-1	1	0	q_4
	$w_j, z_j \geqq 0$, and $w_j z_j = 0$ for all j								

In this problem, in the complementary basic vector w, single principal pivot steps are only possible in positions 1 and 3. Carrying out a single principal pivot in the complementary basic vector w in position 1 leads to the following

basic variable	z_1	w_2	w_3	w_4	w_1	z_2	z_3	z_4	
z_1	1	0	0	0	-1	-1	1	1	q'_1
w_2	0	1	0	0	-1	-1	1	2	q'_2
w_3	0	0	1	0	0	0	-1	-1	q'_3
w_4	0	0	0	1	1	0	0	-1	q'_4

In the above canonical tableau, we have also rearranged the column vectors so that the basic variables, and the nonbasic variables, appear together and in their proper order.

We can make a double principal pivot step in the complementary basic vector w, in positions 2, 4 in this problem, because the determinant of the 2×2 matrix $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ is non-zero. Carrying out this double principal pivot step requires replacing the basic variable w_2 in the basic vector (w_1, w_2, w_3, w_4) by z_4, then replacing the basic variable w_4 in the resulting basic vector (w_1, z_4, w_3, w_4) by z_2, and finally interchanging rows 2 and 4 in the resulting canonical tableau. This is carried out below.

basic variable	w_1	w_2	w_3	w_4	z_1	z_2	z_3	z_4	
w_1	1	1	0	0	−2	1	−1	0	q'_1
z_4	0	1	0	0	−1	0	0	1	q'_2
w_3	0	1	1	0	−1	0	−1	0	q'_3
w_4	0	0	0	1	1	−1	1	0	q'_4
w_1	1	1	0	1	−1	0	0	0	q''_1
z_4	0	1	0	0	−1	0	0	1	q''_2
w_3	0	1	1	0	−1	0	−1	0	q''_3
z_2	0	0	0	−1	−1	1	−1	0	q''_4
w_1	1	1	0	1	−1	0	0	0	q''_1
z_2	0	0	0	−1	−1	1	−1	0	q''_4
w_3	0	1	1	0	−1	0	−1	0	q''_3
z_4	0	1	0	0	−1	0	0	1	q''_2

Block Principal Pivoting

Consider the LCP (q, M), (3.1). Let $\mathbf{J} \subset \{1, \ldots, n\}$ be such that $M_{\mathbf{JJ}}$, the principal submatrix of M corresponding to the subset $\mathbf{J}$, is nonsingular. Define the complementary vector $y = (y_j)$ by

$$y_j = \begin{cases} w_j, & \text{for } j \notin \mathbf{J} \\ z_j, & \text{for } j \in \mathbf{J} \end{cases}$$

and let A be the complementary matrix corresponding to y. Since $M_{\mathbf{JJ}}$ is nonsingular, A is a basis. Let t_j be the complement of y_j for each $j = 1$ to n, and let $t = (t_j)$. Multiplying (3.1) on the left by A^{-1} and rearranging the variables leads to the LCP

y	t	
I	$-D$	q'
$y, t \geq 0,$	$y^T t = 0$	

where

$$\begin{aligned} D_{\mathbf{JJ}} &= (M_{\mathbf{JJ}})^{-1}, D_{\mathbf{J\overline{J}}} = -(M_{\mathbf{JJ}})^{-1} M_{\mathbf{J\overline{J}}} \\ D_{\mathbf{\overline{J}J}} &= M_{\mathbf{\overline{J}J}} (M_{\mathbf{JJ}})^{-1}, D_{\mathbf{\overline{JJ}}} = M_{\mathbf{\overline{JJ}}} - M_{\mathbf{\overline{J}J}} (M_{\mathbf{JJ}})^{-1} M_{\mathbf{J\overline{J}}} \\ q'_{\mathbf{J}} &= -(M_{\mathbf{JJ}})^{-1} q_{\mathbf{J}}, q'_{\mathbf{\overline{J}}} = q_{\mathbf{\overline{J}}} - M_{\mathbf{\overline{J}J}} (M_{\mathbf{JJ}})^{-1} q_{\mathbf{J}} . \end{aligned}$$

Here $\overline{\mathbf{J}} = \{1, \ldots, n\} \setminus \mathbf{J}$, and $M_{\mathbf{J}\overline{\mathbf{J}}}$ is the submatrix $(m_{ij} : i \in \mathbf{J}, j \in \overline{\mathbf{J}})$, etc.; and $q_{\mathbf{J}} = (q_j : j \in \mathbf{J})$, etc. D is of course the PPT of M corresponding to the complementary basic vector y. The above LCP (q', D) is said to have been obtained from the LCP (q, M) in (3.1) by a **block principal pivot step in positions J** (or by **block principal pivoting** on $-M_{\mathbf{JJ}}$) in (3.1).

Corollary 3.4 *If M is a nondegenerate matrix, a single principal pivot step in any position is always possible in every complementary basic vector.*

Proof. Follows from Corollary 3.1 and the argument used in Theorem 3.1. □

Corollary 3.5 *A square matrix M of order n is nondegenerate (that is, principally nondegenerate to be specific) iff every diagonal entry in every PPT of M is non-zero.*

Proof. Follows from Corollary 3.1. □

Theorem 3.3 *If M is a* PD *or a P-matrix, or a nondegenerate matrix in general; starting with a complementary basic vector $y^1 = (y_1^1, \ldots, y_n^1)$, any other complementary basic vector $y^2 = (y_1^2, y_2^2, \ldots, y_n^2)$ for (3.1), can be obtained by performing a sequence of single principal pivot steps.*

Proof. In these cases, by Corollary 3.1 every complementary vector of variables is a complementary basic vector. Hence if y^1 and y^2 have $n - r$ common variables; each of the variables in y^1 which is not in y^2, can be replaced by its complement, to lead to y^2 after r single principal pivot steps. □

Theorem 3.4 *All PPTs of a nondegenerate matrix are nondegenerate.*

Proof. Let M be nondegenerate. Let y, $\hat{y}$ be distinct complementary vectors of variables associated with the complementary matrices A, $\hat{A}$ respectively in (3.1). Since M is nondegenerate, A is a complementary basis. Let (3.4) be the canonical tableau of (3.1) with respect to y. So D is the PPT of M corresponding to y. We will now prove that D is nondegenerate. Look at (3.4). The complementary matrix corresponding to the complementary vector of variables $\hat{y}$ in (3.4) is $A^{-1}\hat{A}$, and this matrix is nonsingular since both A and $\hat{A}$ are. Hence $\hat{y}$ is a complementary basic vector for (3.4). Since $\hat{y}$ is an arbitrary complementary vector of variables, this implies that all complementary vectors of variables in (3.4) are basic vectors.
Hence by Corollary 3.1, D is nondegenerate. □

Theorem 3.5 *All PPTs of a P-matrix are P-matrices.*

Proof. Let $M = (m_{ij})$ be a P-matrix of order n. Consider a single principal pivot step on (3.1) in any position, say position 1. The pivot matrix corresponding to this pivot step is P, which is the same as the unit matrix of order n, with the exception that its first column vector is $(-1/m_{11}, -m_{12}/m_{11}, \ldots, -m_{1n}/m_{11})^T$. Let M' be the

PPT of M obtained after this pivot step. Let $\mathbf{J} = \{j_1, \ldots, j_r\} \subset \{1, \ldots, n\}$, $\mathbf{J} \neq \emptyset$, and let Δ be the principal subdeterminant of M' corresponding to the subset $\mathbf{J}$. We will now prove that $\Delta > 0$. We consider two cases separately.

Case 1: $1 \notin \mathbf{J}$. Let $y = (y_1, \ldots, y_n)$ where $y_j = w_j$ if $j \notin \mathbf{J} \cup \{1\}$, or z_j otherwise. Let A, $\overline{A}$ be the complementary bases corresponding to y, in the original LCP (3.1) and in the canonical tableau for (3.1) obtained after the single principal pivot step in position 1, respectively. So $\overline{A} = PA$. Let Δ_1 be the principal subdeterminant of M corresponding to the subset $\{1\} \cup \mathbf{J}$. We have $\Delta = (-1)^r$ (determinant of $\overline{A}$) $= (-1)^r$ (determinant of PA) $= (-1)^r$ (determinant of P) (determinant of A) $= (-1)^r(-1/m_{11})(-1)^{r+1}\Delta_1 = (\Delta_1/m_{11}) > 0$, because $m_{11} > 0$ and $\Delta_1 > 0$ since M is a P-matrix.

Case 2: $1 \in \mathbf{J}$. In this case let $y = (y_1, \ldots, y_n)$ where $y_j = z_j$ if $j \in \mathbf{J} \setminus \{1\}$, or w_j otherwise. Let A, $\overline{A}$ be the complementary bases corresponding to y, in the original LCP (3.1), and in the canonical tableau for (3.1) obtained after the single principal pivot step in position 1, respectively. Then $\overline{A} = PA$. Let Δ_2 be the principal subdeterminant of M determined by the subset $\mathbf{J} \setminus \{1\}$. As in Case 1, we have $\Delta = (-1)^r$ (determinant of $\overline{A}$) $= (-1)^r$ (determinant of P) (determinant of A) $= (-1)^r(-1/m_{11})(-1)^{r-1}\Delta_2 = (\Delta_2/m_{11}) > 0$, since both Δ_2, m_{11} are strictly positive because M is a P-matrix.

Hence the principal subdeterminant of M' corresponding to the subset $\mathbf{J}$ is strictly positive. This holds for all subsets $\mathbf{J} \subset \{1, \ldots, n\}$. So M' is itself a P-matrix.

Thus the property of being a P-matrix is preserved in the PPTs of M obtained after a single principal pivot step on (3.1). By Theorem 3.3 any PPT of M can be obtained by making a sequence of single principal pivot steps on (3.1). So, applying the above result repeatedly after each single principal pivot step, we conclude that every PPT of M is also a P-matrix.

□

Theorem 3.6 *If all the diagonal entries in every PPT of M are strictly positive, M is a P-matrix.*

Proof. By the hypothesis of the theorem all principal subdeterminants of M of order 1 are strictly positive.

Induction Hypothesis: Under the hypothesis of the theorem, all principal subdeterminants of M of order less than or equal to r are strictly positive.

We will now prove that under the hypothesis of the theorem, the induction hypothesis implies that any principal subdeterminant of M of order $r+1$ is also strictly positive. Let Δ_1 be the principal subdeterminant of M corresponding to the subset $\{j_1, \ldots, j_r, j_{r+1}\} \subset \{1, 2, \ldots, n\}$. Carry out a single principal pivot step in position j_{r+1} in (3.1) and let M' be the PPT of M obtained after this step. Since M' is a PPT of M it also satisfies the hypothesis of the theorem. So by the induction hypothesis, all principal subdeterminants of M' of order r or less are strictly positive, and so Δ, the principal subdeterminant of M' corresponding to the subset $\{j_1, \ldots, j_r\}$, is > 0. As in

the proof of Theorem 3.5 we have $\Delta = \Delta_1/m_{j_{r+1},j_{r+1}}$, that is $\Delta_1 = m_{j_{r+1},j_{r+1}}\Delta$, and since $m_{j_{r+1},j_{r+1}} > 0$, $\Delta > 0$, we have $\Delta_1 > 0$. So under the hypothesis of the theorem, the induction hypothesis implies also that all principal subdeterminants of M of order $r+1$ are strictly positive. Hence by induction, all principal subdeterminants of M are strictly positive, and hence M is a P-matrix.

□

Corollary 3.6 *The following conditions (i) and (ii) are equivalent*

(i) all principal subdeterminants of M are strictly positive

(ii) the diagonal entries in every PPT of M are strictly positive.

Proof. Follows from Theorem 3.5, 3.6.

□

Corollary 3.7 *If M is a P-matrix, in making any sequence of single principal pivot steps on (3.1), the pivot element will always be strictly negative.*

□

Theorem 3.7 *Let M' be a PPT of M obtained after carrying out exactly one single principal pivot step. Then M' is* PD *if M is* PD. *And M' is* PSD *if M is* PSD.

Proof. Let $M = (m_{ij})$. Let $u = (u_1, \ldots, u_n)^T \in \mathbf{R}^n$. Define $v = (v_1, \ldots, v_n)^T$ by

$$v - Mu = 0\,. \tag{3.5}$$

Suppose $M' = (m'_{ij})$ is the PPT of M obtained after making a single principal pivot step in (3.5) in position r. So $m_{rr} \neq 0$. After this single principal pivot step in position r, (3.5) becomes

$$(v_1, \ldots, v_{r-1}, u_r, v_{r+1}, \ldots, v_n)^T - M'(u_1, \ldots, u_{r-1}, v_r, u_{r+1}, \ldots, u_n)^T = 0\,. \tag{3.6}$$

For any $u \in \mathbf{R}^n$ and v defined by (3.5), let $\xi = (u_1, \ldots, u_{r-1}, v_r, u_{r+1}, \ldots, u_n)$, $\eta = (v_1, \ldots, v_{r-1}, u_r, v_{r+1}, \ldots, v_n)$. Since $v_r = M_{r.}u$ and $m_{rr} \neq 0$, as u varies over all of $\mathbf{R}^n$, ξ also varies over all of $\mathbf{R}^n$. Also, as u varies over all the nonzero points in $\mathbf{R}^n$, ξ does the same. Since (3.6) is obtained from (3.5) by a pivot step, they are equivalent. So for any $u \in \mathbf{R}^n$ and v defined by (3.5), (3.6) also holds. Now $u^T Mu = u^T v = \xi^T \eta = \xi^T M' \xi$. These facts imply that $\xi^T M' \xi \geqq 0$ for all $\xi \in \mathbf{R}^n$ iff $u^T Mu \geqq 0$ for all $u \in \mathbf{R}^n$ and $\xi^T M' \xi > 0$ for all $\xi \neq 0$ iff $u^T Mu > 0$ for all $u \neq 0$. Hence M is PD iff M' is PD. And M' is PSD iff M is PSD.

□

Theorem 3.8 *Let M'' be a PPT of M obtained after carrying out exactly one double principal pivot step. Then M'' is* PD *if M is* PD. *And M'' is* PSD *if M is* PSD.

Proof. Let $M = (m_{ij})$. Let $u = (u_1, \ldots, u_n)^T \in \mathbf{R}^n$. Define $v = (v_1, \ldots, v_n)^T$ by (3.5). Suppose $M'' = (m''_{ij})$ is the PPT of M obtained after making a double principal pivot step in positions r and s. This implies that

$$\Delta = \text{ determinant } \begin{pmatrix} -m_{ss} & -m_{sr} \\ -m_{rs} & -m_{rr} \end{pmatrix} \neq 0\,,$$

as otherwise the double principal pivot step in positions r and s cannot be carried out on (3.5). For any $u \in \mathbf{R}^n$ and v defined by (3.5) define $\xi = (u_1, \ldots, u_{s-1}, v_s, u_{s+1}, \ldots, u_{r-1}, v_r, u_{r+1}, \ldots, u_n)^T$, $\eta = (v_1, \ldots, v_{s-1}, u_s, v_{s+1}, \ldots, v_{r-1}, u_r, v_{r+1}, \ldots, v_n)^T$. Then after this double principal pivot step in positions r and s, (3.5) gets transformed into

$$\eta - M''\xi = 0\,. \tag{3.7}$$

Since (3.7) is obtained by performing two pivots on (3.5), they are equivalent. So for any $u \in \mathbf{R}^n$ and v defined by (3.5), (3.7) holds and we have $u^T M u = u^T v = \xi^T \eta = \xi^T M'' \xi$. Also, since $\Delta \neq 0$, as u varies over all of $\mathbf{R}^n$, so does ξ; and as u varies over all nonzero points in $\mathbf{R}^n$ so does ξ. These facts imply that $\xi^T M'' \xi \geqq 0$ for all $\xi \in \mathbf{R}^n$ iff $u^T M u \geqq 0$ for all $u \in \mathbf{R}^n$ and $\xi^T M'' \xi > 0$ for all $\xi \neq 0$ iff $u^T M u > 0$ for all $u \neq 0$. Hence M'' is PD iff M is PD, and M'' is PSD iff M is PSD.

□

Theorem 3.9 *If M is a* PD *matrix, all its PPTs are also* PD.

Proof. By Theorem 3.3 when M is PD, every PPT of M can be obtained by carrying out a sequence of single principal pivot steps on (3.1). By applying the argument in Theorem 3.7 repeatedly after each single principal pivot step in the sequence, we conclude that all PPTs of M are also PD, if M is.

□

Theorem 3.10 *If M is* PSD, *any PPT of M can be obtained by making a sequence of single or double principal pivot steps on (3.1). Also, all these PPTs of M are also* PSD.

Proof. Let $y = (y_1, \ldots, y_n)$ be a complementary basic vector of (3.1). Starting with the complementary basic vector w, perform single principal pivot steps in position j for as many $j \in \mathbf{Z}(y)$ as possible in any possible order. If this leads to the complementary basic vector y, we are done by repeated use of the result in Theorem 3.7 after each single principal pivot step. Suppose y has not yet been obtained and no more single principal pivot steps can be carried out in the remaining positions $j \in \mathbf{Z}(y)$. Let $u = (u_1, \ldots, u_n)$ be the complementary basic vector at this stage. Let $\mathbf{U} = \{j : j \text{ such that } u_j \neq y_j\}$. So $\mathbf{U} \neq \emptyset$, $\mathbf{U} \subset \mathbf{Z}(y)$. And for each $j \in \mathbf{U}$, we have $u_j = w_j$, $y_j = z_j$. Let t_j denote the complement of u_j, $j = 1$ to n. Let the canonical tableau of (3.1) at this stage be

basic vector	$u_1, \ldots, u_n$	$t_1, \ldots, t_n$	q
u	I	$-M'$	q'

(3.8)

M' is the PPT of M corresponding to $\mathbf{U}$, it is PSD by repeated use of Theorem 3.7. We have $-m'_{jj} = 0$ for each $j \in \mathbf{U}$ (as single principal pivot steps cannot be carried out in these positions). If $\mathbf{U}$ is a singleton set, this would imply that the set of column vectors corresponding to y in (3.8) is linearly dependent, a contradiction, since y is a complementary basic vector. So cardinality of $\mathbf{U}$ is $\geqq 2$. Let $r \in \mathbf{U}$. Since $m'_{rr} = 0$ and

M' is PSD, by Result 1.6 we have $m'_{ri} + m'_{ir} = 0$ for all $i = 1$ to n. Search for an $s \in \mathbf{U}$ such that $m'_{sr} \neq 0$. If an s like this does not exist, again the set of column vectors corresponding to y in (3.8) is linearly dependent, and y is not a complementary basic vector, a contradiction. So there always exists an $s \in \mathbf{U}$ such that $m'_{sr} \neq 0$. Since $m'_{rs} + m'_{sr} = 0$, $m'_{rs} \neq 0$ too. So the determinant

$$\begin{pmatrix} m'_{rr} & m'_{rs} \\ m'_{sr} & m'_{ss} \end{pmatrix}$$

is nonzero, and a double principal pivot step can be carried out in (3.8) in positions r, s. The complementary basic vector obtained after this double principal pivot step contains two more variables in common with y than u does, and the PPT of M corresponding to it is also PSD by Theorem 3.8. Delete r, s from $\mathbf{U}$. In the resulting canonical tableau, make as many single principal pivot steps in positions $j \in \mathbf{U}$ as possible, deleting such j from $\mathbf{U}$ after each step. Or make another double principal pivot step in positions selected from $\mathbf{U}$ as above, and continue the same way until $\mathbf{U}$ becomes empty. At that stage we reach the canonical tableau with respect to y. By repeated use of Theorems 3.7, 3.8, the PPT of M with respect to y is also PSD.

□

3.2.1 Principal Rearrangements of a Square Matrix

Let M be a given square matrix of order n. Let $p = (i_1, \ldots, i_n)$ be a permutation of $(1, \ldots, n)$. The square matrix P of order n whose rows are $I_{i_1}., I_{i_2}., \ldots, I_{i_n}.$ in that order, is the permutation matrix corresponding to the permutation p. P is obtained essentially by permuting the rows of the unit matrix I of order n using the permutation p. The matrix $M' = PMP^T$ is known as the principal rearrangement of M according to the permutation p. Clearly M' is obtained by first rearranging the rows of M according to the permutation p, and in the resulting matrix, rearranging the columns again according to the same permutation p.

As an example let $n = 3$, and

$$p = (3, 1, 2)\,, \quad M = \begin{pmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{pmatrix}\,, \quad P = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

then

$$PM = \begin{pmatrix} m_{31} & m_{32} & m_{33} \\ m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \end{pmatrix}\,, \quad M' = PMP^T = \begin{pmatrix} m_{33} & m_{31} & m_{32} \\ m_{13} & m_{11} & m_{12} \\ m_{23} & m_{21} & m_{22} \end{pmatrix}$$

and M' here is the principal rearrangement of M according to the permutation p.

The following results can be obtained directly using the definition. Let M' be the principal rearrangement of M according to the permutation p associated with the permutation matrix P. Then M' is a P-matrix, iff M is. For all $y \in \mathbf{R}^n$, $y^T My = (Py)^T M'(Py)$. So M' is a PSD, or PD, or NSD, or ND matrix iff M has the same property. Also, M' is principally degenerate (or nondegenerate) iff M has the same property.

3.3 LCPs ASSOCIATED WITH P-MATRICES

Properties of P-Matrices

The following Theorems 3.11, 3.12 are important properties of P-matrices due to D. Gale and H. Nikaido (see reference [3.24]).

Theorem 3.11 *Let $F = (f_{ij})$ be a P-matrix of order n. Then the system of linear inequalities*

$$\begin{aligned} Fx &\leqq 0 \\ x &\geqq 0 \end{aligned} \tag{3.9}$$

has "$x = 0$" as its unique solution.

Proof. The theorem is easily verified to be true for $n = 1$. We will prove the theorem for all n by induction.

Induction Hypothesis: If T is a P-matrix of order $r \leqq n-1$, then the system of inequalities $T\xi \leqq 0$, $\xi \geqq 0$, $\xi \in \mathbf{R}^r$ has "$\xi = 0$" as its unique solution.

Under the induction hypothesis we will now prove that the statement of the theorem holds for the matrix F which is a P-matrix of order n. Since F is a P-matrix, it is nonsingular, and hence F^{-1} exists. Let $B = F^{-1} = (b_{ij})$. From standard results in the theory of determinants (for example, see Chapter 3 in F. E. Hohn, *Elementary Matrix Algebra*, Macmillan, 2nd edition, 1964) it is known that $b_{ii} =$ (principal subdeterminant of F corresponding to the subset $\{1, \ldots, i-1, i+1, \ldots, n\}$)/determinant of F. So $b_{ii} > 0$ for all i, since F is a P-matrix. Thus each column of B has at least one positive entry. Let $\overline{x} \in \mathbf{R}^n$ satisfy (3.9). Select a column of B, say $B_{.1}$. Let $\theta =$ minimum$\{\overline{x}_i/b_{i1} : i$ such that $b_{i1} > 0\}$, and suppose this minimum is attained by $i = s$. So $\theta = \overline{x}_s/\overline{b}_{s1} \geqq 0$, and $(\overline{x}_j/b_{j1}) \geqq \theta$, for all j such that $b_{j1} > 0$. From this and the fact that $\overline{x} \geqq 0$, we have $\overline{\eta} = (\overline{\eta}_1, \ldots, \overline{\eta}_n)^T = \overline{x} - \theta B_{.1} \geqq 0$ and $\overline{\eta}_s = 0$. Also $F\overline{\eta} = F\overline{x} - \theta FB_{.1} = F\overline{x} - \theta I_{.1} \leqq 0$. Let T be the matrix of order $n-1$ obtained by striking off the sth row and the sth column from F. Since F is a P-matrix, its principal submatrix T is also a P-matrix. Let $\overline{\xi} = (\overline{\eta}_1, \ldots, \overline{\eta}_{s-1}, \overline{\eta}_{s+1}, \ldots, \overline{\eta}_n)^T$. Since $\overline{\eta}_s = 0$ and $F\overline{\eta} \leqq 0$, we have $T\overline{\xi} \leqq 0$. Also since $\overline{\eta} \geqq 0$, $\overline{\xi} \geqq 0$ too. So $T\overline{\xi} \leqq 0$, $\overline{\xi} \geqq 0$. Since T is a P-matrix of order $n-1$, by the induction hypothesis, $\overline{\xi} = 0$. $\overline{\xi} = 0$, $\overline{\eta}_s = 0$ together imply that $\overline{\eta} = 0$. So $F\overline{\eta} = 0$, that is $F(\overline{x} - \theta I_{.1}) = 0$. Then $F\overline{x} = \theta I_{.1} \geqq 0$. However from (3.9), $F\overline{x} \leqq 0$. So $F\overline{x} = 0$, and since F is nonsingular, $\overline{x} = 0$.

Thus under the induction hypothesis the statement of the theorem also holds for F which is a P-matrix of order n. The statement of the theorem is easily verified for $n = 1$. Hence, by induction, the statement of the theorem is true for all n.

□

Theorem 3.12 The Sign Nonreversal Property: *Let F be a square matrix of order n. For $x \in \mathbf{R}^n$ let $y = Fx$. Then F is said to reverse the sign of x if $x_i y_i \leqq 0$ for all i. If F is a P-matrix it reverses the sign of no vector except zero.*

Proof. For this proof we need only to consider the case $x \geq 0$. For if F reverses the sign of an $\overline{x} \ngeqq 0$, let $\mathbf{J} = \{j : \overline{x}_j < 0\}$, let D be the diagonal matrix obtained from the unit matrix by multiplying its jth column by -1 for each $j \in \mathbf{J}$. The matrix $F^* = DFD$ is again a P-matrix, since F^* is obtained by simply changing the signs of rows and columns in F for each $j \in \mathbf{J}$. And F^* reverses the sign of $\hat{x} = D\overline{x}$, where $\hat{x} \geqq 0$.

Now suppose that $x \geqq 0$ and that F reverses the sign of x. Let $\mathbf{P} = \{j : x_j > 0\}$. Assume that $\mathbf{P} \neq \emptyset$. Let A be the principal submatrix of F corresponding to $\mathbf{P}$. Let χ be the vector of x_j for $j \in \mathbf{P}$. The fact that F reverses the sign of x implies that A reverses the sign of χ. Since $\chi > 0$, this implies that $A\chi \leqq 0$. Since A is a P-matrix $A\chi \leqq 0$, $\chi \geqq 0$ implies $\chi = 0$ by Theorem 3.11, a contradiction. So x must be zero. □

Unique Solution Property of LCPs Associated with P-Matrices

Theorem 3.13 *Let M be a P-matrix. The LCP (q, M) has a unique solution for each $q \in \mathbf{R}^n$. Also, when the complementary pivot algorithm of Section 2.2 is applied on the LCP (q, M), it finds the solution.*

Proof. Suppose when the complementary pivot algorithm is applied on the LCP (q, M) it ends in ray termination. As in the proof of Theorem 2.1 this implies that there exists a $z^h \geq 0$, $w^h \geqq 0$, $z_0^h \geqq 0$ satisfying $w^h = Mz^h + e_n z_0^h$; $w_i^h z_i^h = 0$ for all i. So $z_i^h(M_{i.}z^h) + z_i^h z_0^h = 0$. This implies that $z_i^h(M_{i.}z^h) = -z_i^h z_h^0 \leqq 0$ for all i. So M reverses the sign of $z^h \geq 0$, which is a contradiction to Theorem 3.12. So, when the complementary pivot method is applied on the LCP (q, M) associated with a P-matrix, it cannot end in ray termination, it has to terminate with a solution of the LCP. This also proves that every P-matrix is a Q-matrix.

Now we will prove that if M is a P-matrix, for any $q \in \mathbf{R}^n$, the LCP (q, M) has exactly one solution, by induction on n, the order of the problem.

Suppose $n = 1$. $M = (m_{11})$ is a P-matrix, iff $m_{11} > 0$. In this case $q = (q_1)$. If $q_1 \geqq 0$, $(w = (w_1) = (q_1); z = (z_1) = (0))$ is the only solution to the LCP (q, M). If $q_1 < 0$, $(w = (w_1) = (0); z = (z_1) = (-q_1/m_{11}))$ is the only solution to the LCP (q, M). Hence the theorem is true for $n = 1$.

Induction Hypothesis: Suppose any LCP of order $(n-1)$ or less, associated with a P-matrix, has a unique solution for each of its right hand side constant vectors.

Now we will prove that under the induction hypothesis, the LCP (q, M) where M is a P-matrix of order n, has a unique solution for any $q \in \mathbf{R}^n$. We have shown above that it has at least one solution, say $(\tilde{w}; \tilde{z})$. For each $j = 1$ to n let $u_j = z_j$, if

$\tilde{z}_j > 0$; or w_j otherwise; and let v_j be the complement of u_j. Then $u = (u_1, \ldots, u_n)$ is a complementary feasible basic vector of variables associated with the BFS $(\tilde{w}; \tilde{z})$ for (3.1). Obtain the canonical tableau for (3.1) with respect to the complementary feasible basic vector u, and suppose it is

$u_1, \ldots, u_n$	$v_1, \ldots, v_n$	q
I	$-\widetilde{M}$	$\tilde{q}$

(3.10)

$\tilde{q} \geqq 0$ by our assumptions here. (3.10) can itself be viewed as the LCP $(\tilde{q}, \widetilde{M})$, one solution of this LCP is $(u = \tilde{u} = \tilde{q}; v = \tilde{v} = 0)$. $\widetilde{M}$ is a PPT of M, by Theorem 3.5, $\widetilde{M}$ is also a P-matrix. So all the principal submatrices of $\widetilde{M}$ are also P-matrices. So the principal subproblem of the LCP $(\tilde{q}, \widetilde{M})$ in the variables $(u_1, \ldots, u_{i-1}, u_{i+1}, \ldots, u_n)$; $(v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n)$ is an LCP of order $(n-1)$ associated with a P-matrix, and by the induction hypothesis this principal subproblem has a unique solution. One solution of this principal subproblem is $(\tilde{u}_1, \ldots, \tilde{u}_{i-1}, \tilde{u}_{i+1}, \ldots, \tilde{u}_n; \tilde{v}_1, \ldots, \tilde{v}_{i-1}, \tilde{v}_{i+1}, \ldots, \tilde{v}_n) = (\tilde{q}_1, \ldots, \tilde{q}_{i-1}, \tilde{q}_{i+1}, \ldots, \tilde{q}_n; 0, \ldots, 0, 0, \ldots, 0)$. If the LCP $(\tilde{q}, \widetilde{M})$, (3.10), has an alternate solution $(\hat{u}; \hat{v}) \neq (\tilde{u}; \tilde{v})$ in which $\hat{v}_i = 0$, its principal subproblem in the variables $(u_1, \ldots, u_{i-1}, u_{i+1}, \ldots, u_n)$; $(v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n)$ will have an alternate solution $(\hat{u}_1, \ldots, \hat{u}_{i-1}, \hat{u}_{i+1}, \ldots, \hat{u}_n; \hat{v}_1, \ldots, \hat{v}_{i-1}, \hat{v}_{i+1}, \ldots, \hat{v}_n)$, a contradiction. So, if the LCP $(\tilde{q}, \widetilde{M})$ has an alternate solution $(\hat{u}; \hat{v}) \neq (\tilde{u}; \tilde{v})$, then $\hat{v}_i$ must be strictly positive in it, and by complementarity $\hat{u}_i$ must be zero. Since this holds for each $i = 1$ to n, $\hat{v} > 0$, $\hat{u} = 0$. So $\hat{u} - \widetilde{M}\hat{v} = \tilde{q}$, $\hat{u} = 0$, $\hat{v} > 0$. Since $\tilde{q} \geqq 0$, this implies that $\widetilde{M}\hat{v} = -\tilde{q} \leqq 0$, $\hat{v} > 0$, a contradiction to Theorem 3.11, since $\widetilde{M}$ is a P-matrix. Hence under the induction hypothesis the LCP $(\tilde{q}, \widetilde{M})$ has a unique solution, which implies that the equivalent LCP (q, M) has a unique solution also. Since this holds for any $q \in \mathbf{R}^n$, under the induction hypothesis, the LCP (q, M) of order n has a unique solution for each $q \in \mathbf{R}^n$ when M is a P-matrix. Hence, by induction the theorem is true.

□

Theorem 3.14 *Let M be a given square matrix of order n. Suppose the LCP (q, M) has at most one solution for each $q \in \mathbf{R}^n$. Then M is a P-matrix.*

Proof. So, the number of solutions of the LCP (q, M) is either 1 or 0 and hence is finite for all q, which implies that M is nondegenerate by Theorem 3.2. So the determinant of M is nonzero, and hence M^{-1} exists.

Proof is by induction on n, the order of the matrix M. We first verify that the theorem is true if $n = 1$. In this case $q = (q_1)$, $M = (m_{11})$. Since M is shown to be nondegenerate under the hypothesis of the theorem, $m_{11} \neq 0$. If $m_{11} < 0$; when $q_1 > 0$, $(w = (q_1), z = 0)$, $(w = 0, z = q_1/(|m_{11}|))$ are two distinct solutions of the LCP (q, M). Hence under the hypothesis of the theorem $m_{11} \not< 0$. So, $m_{11} > 0$, which implies that the theorem is true when $n = 1$.

Induction Hypothesis: If F is a square matrix of order $r \leqq n - 1$, such that the LCP (γ, F) has at most one solution for each $\gamma \in \mathbf{R}^r$, then F is a P-matrix.

Under the hypothesis of the theorem, and the induction hypothesis, we will now prove that M has to be a P-matrix too.

Consider the principal subproblem of the LCP (q, M) in the variables $\omega = (w_2, \ldots, w_n)$, $\xi = (z_2, \ldots, z_n)$. This is an LCP of order $n-1$ associated with the principal submatrix of M determined by the subset $\{2, \ldots, n\}$. If there exists a $\tilde{q} = (\tilde{q}_2, \ldots, \tilde{q}_n)^T$ for which this principal subproblem has two distinct solutions, namely, $(\overline{\omega}, \overline{\xi})$ and $(\hat{\omega}, \hat{\xi})$, choose $\tilde{q}_1$ to satisfy $\tilde{q}_1 > \text{Maximum}\{|\sum_{j=2}^{n} \overline{z}_j m_{1j}|, |\sum_{j=2}^{n} \hat{z}_j m_{1j}|\}$, and let $\overline{w}_1 = \tilde{q}_1 + \sum_{j=2}^{n} \overline{z}_j m_{1j}$, $\overline{z}_1 = 0$, $\hat{w}_1 = \tilde{q}_1 + \sum_{j=2}^{n} \hat{z}_j m_{1j}$, $\hat{z}_1 = 0$, $\overline{w} = (\overline{w}_1, \overline{w}_2, \ldots, \overline{w}_n)$, $\overline{z} = (\overline{z}_1, \overline{z}_2, \ldots, \overline{z}_n)$, $\hat{w} = (\hat{w}_1, \hat{w}_2, \ldots, \hat{w}_n)$, $\hat{z} = (\hat{z}_1, \hat{z}_2, \ldots, \hat{z}_n)$, $\tilde{q} = (\tilde{q}_1, \tilde{q}_2, \ldots, \tilde{q}_n)^T$. Then $(\overline{w}; \overline{z})$, $(\hat{w}; \hat{z})$ are two distinct solutions of the LCP $(\tilde{q}, M)$, contradicting the hypothesis of the theorem. So the principal subproblem of the LCP (q, M) in the variables ω, ξ has at most one solution for each of its right hand side constant vectors. By the induction hypothesis this implies that the principal submatrix of M determined by the subset $\{2, \ldots, n\}$ is a P-matrix.

A similar argument can be made for each principal subproblem of the LCP (q, M) of order $(n-1)$, and this implies that all principal submatrices of M of order $(n-1)$ are P-matrices, by the induction hypothesis. Hence all the principal subdeterminants of M of order $\leqq (n-1)$ are strictly positive. In particular, the diagonal entries of M are strictly positive. It only remains to be proved that the determinant of M itself is strictly positive. We have already seen that M^{-1} exists. The canonical tableau of (3.1) with respect to the complementary basic vector $(z_1, \ldots, z_n)$ is

z	w	
I	$-\overline{M}$	$\overline{q}$

(3.11)

where $\overline{M} = M^{-1}$ and $\overline{q} = -M^{-1}q$. The LCP in (3.11) has at most one solution for each $\overline{q} \in \mathbf{R}^n$. So by the previous arguments all diagonal entries in the matrix $\overline{M}$ have to be strictly positive. However since $\overline{M} = (\overline{m}_{ij}) = M^{-1}$, $\overline{m}_{11} =$ (principal subdeterminant of M corresponding to the subset $\{2, \ldots, n\}$)/(determinant of M). Since the principal subdeterminant of M corresponding by the subset $\{2, \ldots, n\}$ has been shown to be strictly positive, $\overline{m}_{11} > 0$ implies that the determinant of M is strictly positive. Hence under the hypothesis of the theorem, and the induction hypothesis, the matrix M of order n has to be a P-matrix. So, by induction the theorem is true in general.

□

Corollary 3.8 *Let M be a given square matrix of order n. If the LCP (q, M) has at most one solution for each $q \in \mathbf{R}^n$, then it has exactly one solution for each $q \in \mathbf{R}^n$. This follows from Theorems 3.13, 3.14.*

□

Theorem 3.15 *Let M be a given square matrix of order n. The LCP (q, M) has a unique solution for each $q \in \mathbf{R}^n$ iff M is a P-matrix.*

□

Proof. Follows from Theorems 3.13, 3.14.

Strict Separation Property

The strict separation property is a property of the matrix M, and does not depend on the right hand side constants vector q. An LCP associated with the matrix M (or the class of complementary cones $\mathcal{C}(M)$) is said to satisfy the **strict separation property** if the following conditions are satisfied.

(i) Every subcomplementary set of column vectors is linearly independent.

(ii) If $(A_{.1}, \ldots, A_{.i-1}, A_{.i+1}, \ldots, A_{.n})$ is any subcomplementary set of column vectors, the hyperplane which is its linear hull strictly separates the points represented by the left out complementary pair of column vectors $(I_{.i}, -M_{.i})$.

From (i) and (ii), it is clear that every complementary set of column vectors has to be linearly independent for the strict separation property to be satisfied.

Example 3.4

Let $M = \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix}$. Here $n = 2$. The points representing the column vectors of I, $-M$ are plotted in Figure 3.1.

Since $n = 2$ here, in this case each subcomplementary set consists of exactly one of the column vectors from $\{I_{.1}, I_{.2}, -M_{.1}, -M_{.2}\}$. The linear hull of any subcomplementary set of vectors in this example is the straight line through the vector in that subcomplementary set and the origin.

Consider the subcomplementary set of column vectors $\{I_{.1}\}$. The left out complementary pair of column vectors in this set is $(I_{.2}, -M_{.2})$. The linear hull of $\{I_{.1}\}$, which is the horizontal axis in Figure 3.1, strictly separates the points $I_{.2}, -M_{.2}$, since neither of these points is on this straight line and they are on opposite sides of it. In a similar manner it can be verified that both properties (i) and (ii) discussed above are satisfied in this example. Hence any LCP associated with the matrix M in this example satisfies the strict separation property.

Example 3.5

Let $M = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. Here again, $n = 2$. The points representing the column vectors of I, $-M$ in this case are plotted in Figure 1.3. Consider the subcomplementary set of column vectors $\{I_{.2}\}$ in this example. Its linear hull is the vertical axis in Figure 1.3, and it strictly separates the left-out complementary pair of column vectors $(I_{.1}, -M_{.1})$. In a similar manner, it can be verified that the strict separation property holds in this case.

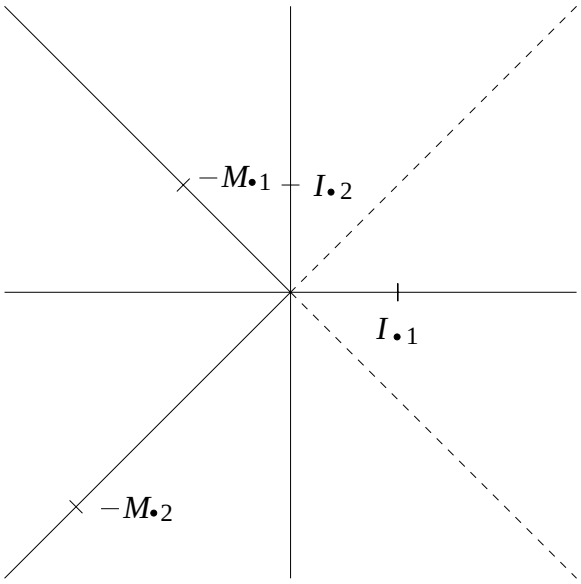


Figure 3.1 Illustration of Strict Separation

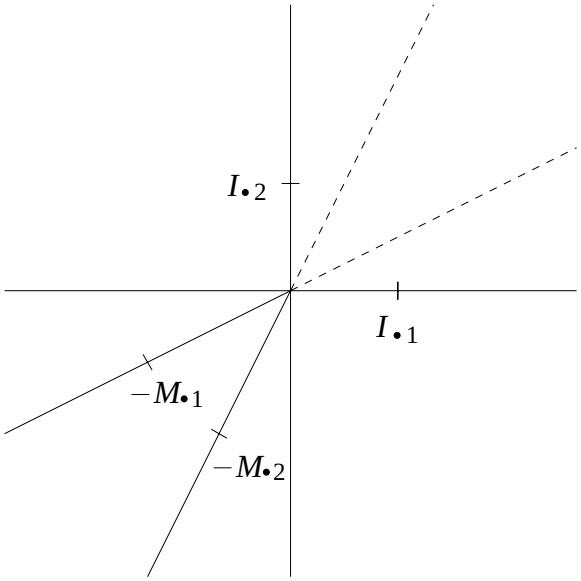


Figure 3.2 Violation of the Strict Separation Property

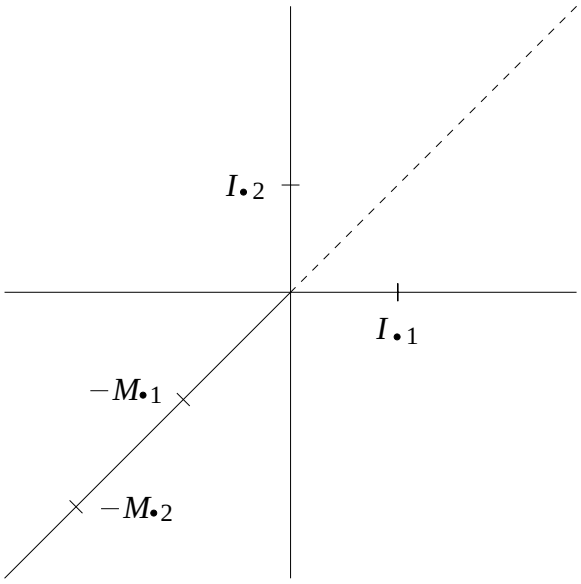


Figure 3.3 Another Example of Violation of the Strict Separation Property.

Example 3.6

Let $M = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}$. Here $n = 2$, and the column vectors of I, $-M$ are plotted in Figure 3.2. Consider the subcomplementary set of column vectors $\{-M_{.1}\}$ here. Both the points in the left-out complementary pair $(I_{.2}, -M_{.2})$ are on the same side of the linear hull of $\{-M_{.1}\}$ here, and hence the strict separation property is not satisfied by the LCPs associated with the matrix M here.

Example 3.7

Let $M = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}$. See Figure 3.3. Consider the subcomplementary set of column vectors $\{-M_{.1}\}$ here. The point $-M_{.2}$ from the left-out complementary pair $(I_{.2}, -M_{.2})$ lies on the straight line which is the linear hull of the subcomplementary set of column vectors $\{-M_{.1}\}$. So the strict separation property is not satisfied in this example.

Corollary 3.9 *If an LCP associated with the matrix M satisfies the strict separation property, M is nondegenerate. This follows from the definitions.*

□

Theorem 3.16 *The LCP associated with a matrix M satisfies the strict separation property iff M is a P-matrix.*

Proof. Suppose M is a P-matrix. Property (i) required for strict separation property is obviously satisfied because M is nondegenerate (Corollary 3.1).

Let $(A_{.1}, \ldots, A_{.i-1}, A_{.i+1}, \ldots, A_{.n})$ be any subcomplementary set of column vectors where $A_{.j} \in \{I_{.j}, -M_{.j}\}$ for each $j \neq i$. Let $\mathbf{H}$ be the hyperplane which is the linear hull of $\{A_{.1}, \ldots, A_{.i-1}, A_{.i+1}, \ldots, A_{.n}\}$. By Corollary 3.1, the complementary sets of column vectors $(A_{.1}, \ldots, A_{.i-1}, I_{.i}, A_{.i+1}, \ldots, A_{.n})$ and $(A_{.1}, \ldots, A_{.i-1}, -M_{.i}, A_{.i+1}, \ldots, A_{.n})$ are both linearly independent, and hence neither $I_{.i}$ nor $-M_{.i}$ lie on the hyperplane $\mathbf{H}$. Suppose both $I_{.i}$ and $-M_{.i}$ are on the same side of the hyperplane $\mathbf{H}$ in $\mathbf{R}^n$. See Figure 3.4. In this case the interiors of the complementary cones $\text{Pos}(A_{.1}, \ldots, A_{.i-1}, I_{.i}, A_{.i+1}, \ldots, A_{.n})$ and $\text{Pos}(A_{.1}, \ldots, A_{.i-1}, -M_{.i}, A_{.i+1}, \ldots, A_{.n})$ have a nonempty intersection, and if $\overline{q}$ is a point in the intersection, then $\overline{q}$ is in the interior of two complementary cones, and the LCP $(\overline{q}, M)$ has two distinct solutions; a contradiction to Theorem 3.13, since M is a P-matrix. So $I_{.i}$ and $-M_{.i}$ cannot be on the same side of the hyperplane $\mathbf{H}$. Since neither of these points is on $\mathbf{H}$, and they are not on the same side of $\mathbf{H}$, these points are on either side of $\mathbf{H}$, that is $\mathbf{H}$ separates them strictly. Since this holds for any subcomplementary set of column vectors and the corresponding left-out complementary pair of column vectors, the strict separation property holds when M is a P-matrix.

Suppose the strict separation property is satisfied. By Corollary 3.9 M is nondegenerate. So all the principal subdeterminants of M are nonzero. It remains to be proved that they are all positive. Let $y = (y_1, \ldots, y_n)$ be any complementary vector of variables for the LCP (q, M). Let t_j be the complement of y_j for $j = 1$ to n. Since M is nondegenerate, $(y_1, \ldots, y_n)$ is a complementary basic vector of variables by Corollary 3.1. Obtain the canonical tableau of (3.1), with respect to the complementary basic vector y. Suppose it is

$y_1 \ldots y_n$	$t_1 \ldots t_n$	q
I	$-M'$	q'

(3.12)

where $M' = (m'_{ij})$ is the PPT of M corresponding to the complementary basic vector y. Now look at the subcomplementary vector of variables $(y_1, \ldots, y_{i-1}, y_{i+1}, \ldots, y_n)$.

The column corresponding to y_j in (3.12) is $I_{.j}$, for $j = 1$ to n. For convenience, call the coordinates along the axis of coordinates, as $x_1, \ldots, x_n$. Since the column of y_j in (3.12) is $I_{.j}$, the hyperplane in $\mathbf{R}^n$ which contains the columns of y_j in (3.12) for all $j = 1, \ldots, i-1, i+1, \ldots, n$, is the coordinate hyperplane $\mathbf{H} = \{x : x_i = 0\}$.

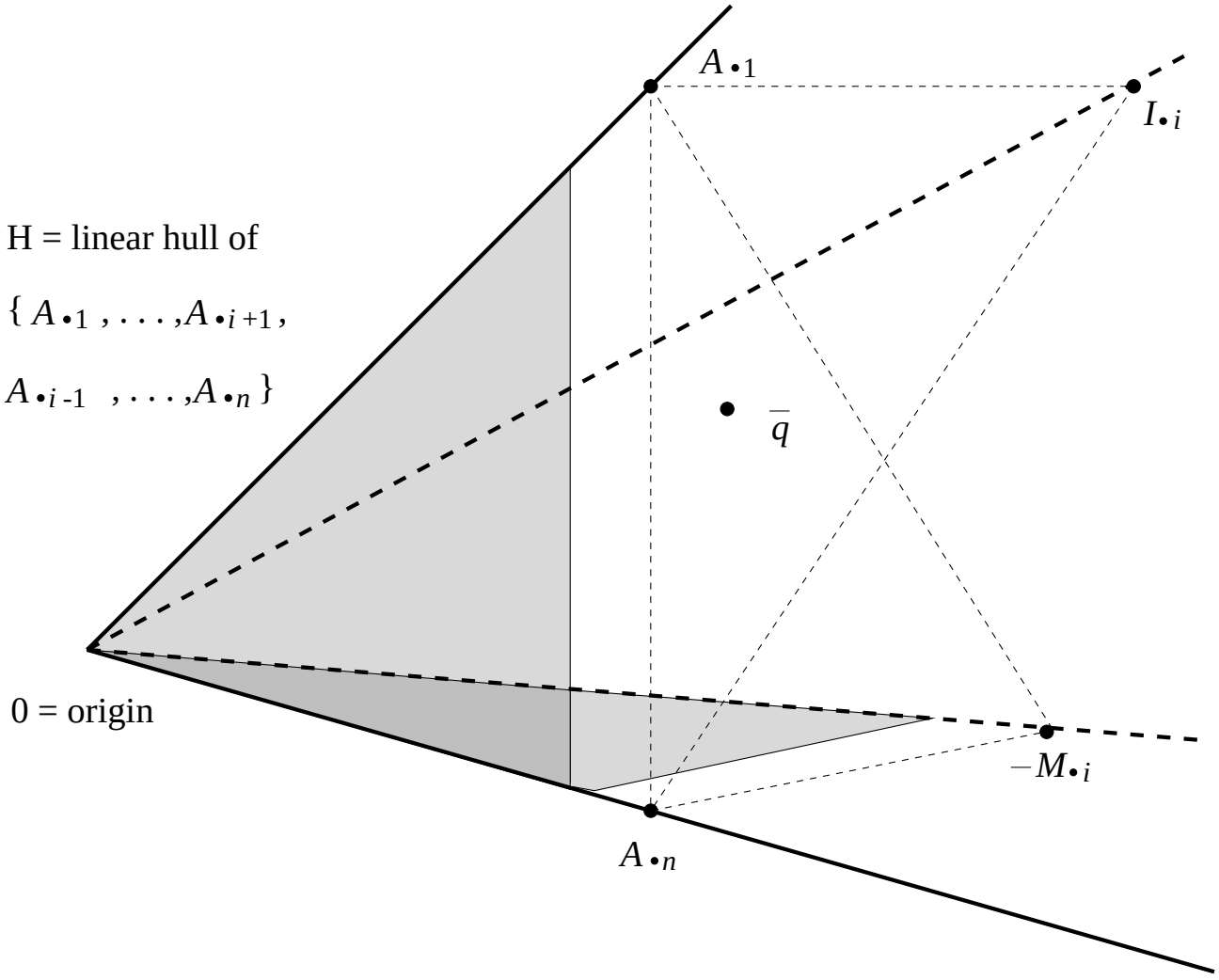


Figure 3.4 $I_{.i}$ and $-M_{.i}$ are both on the same side of $\mathbf{H}$. Interiors of the complementary cones $\text{Pos}(A_{.1}, \ldots, A_{.i-1}, I_{.i}A_{.i+1}, \ldots, A_{.n})$ and $\text{Pos}(A_{.1}, \ldots, A_{.i-1}, -M_{.i}A_{.i+1}, \ldots, A_{.n})$ have a nonempty intersection.

Among the left-out complementary pair of column vectors $(I_{.i}, -M'_{.i})$, since the ith component in the column vector $I_{.i}$ is $+1$, it is on the side on $\mathbf{H}$ corresponding to the inequality $x_i > 0$. So by the strict separation property, the point $-M'_{.i}$ is on the side of $\mathbf{H}$ corresponding to the inequality $x_i < 0$, which implies that $-m'_{ii} < 0$, or $M'_{ii} > 0$. Thus the ith diagonal element in M' is strictly positive. In a similar manner we see that if the strict separation property holds, then all the diagonal elements in all PPTs of M are strictly positive. By Theorem 3.6 this implies that M is a P-matrix.

□

A class of convex polyhedral cones in $\mathbf{R}^n$ is said to **partition** $\mathbf{R}^n$ if

a) Every cone in the class has a nonempty interior.

b) The union of the cones in the class is $\mathbf{R}^n$.

c) The interiors of any pair of cones in the class are disjoint.

Theorem 3.17 *Let M be a given square matrix of order n. The class of complementary cones $\mathcal{C}(M)$ partitions $\mathbf{R}^n$ iff M is a P-matrix.*

Proof. If M is a P-matrix, the result that the class of complementary cones $\mathcal{C}(M)$ partitions $\mathbf{R}^n$ follows from Corollary 3.1 and Theorem 3.13.

To prove the converse, suppose that $\mathcal{C}(M)$ partitions $\mathbf{R}^n$. Since every complementary cone in $\mathcal{C}(M)$ has a nonempty interior, by Corollary 3.2, M must be nondegenerate. Hence all complementary sets of column vectors are linearly independent. If the strict separation property is not satisfied, there exists a subcomplementary set of column vectors, say $(A_{.1}, \ldots, A_{.i-1}, A_{.i+1}, \ldots, A_{.n})$ such that the hyperplane H which is its linear hull contains both the points in the left out complementary pair $(I_{.i}, -M_{.i})$ on the same side of it. As in the proof of Theorem 3.16, this implies that the interiors of the complementary cones $\text{Pos}(A_{.1}, \ldots, A_{.i-1}, I_{.i}, A_{.i+1}, \ldots, A_{.n})$ and $\text{Pos}(A_{.1}, \ldots, A_{.i-1}, -M_{.i}, A_{.i+1}, \ldots, A_{.n})$ have a nonempty intersection; a contradiction, since $\mathcal{C}(M)$ partitions $\mathbf{R}^n$. Hence, if $\mathcal{C}(M)$ partitions $\mathbf{R}^n$, the strict separation property is satisfied, and by Theorem 3.16 this implies that M is a P-matrix.

Hence the class of complementary cones $\mathcal{C}(M)$ partitions $\mathbf{R}^n$ iff M is a P-matrix. □

Example 3.8

Let $M = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. The complementary cones are the quadrants in $\mathbf{R}^2$, drawn in Figure 1.3, and obviously this class of complementary cones partitions $\mathbf{R}^n$. For any n in general $\mathcal{C}(I)$ is the class of orthants of $\mathbf{R}^n$, and these obviously partition $\mathbf{R}^n$. As mentioned earlier the class of complementary cones is a generalization of the class of orthants of $\mathbf{R}^n$ (orthants of $\mathbf{R}^n$ are the special class of complementary cones obtained by taking $M = I$), and $\mathcal{C}(M)$ possesses the property of partitioning $\mathbf{R}^n$ as long as M is a P-matrix. This was first proved by Samelson, Thrall and Wesler in [3.69].

Corollary 3.10 *Let M be a given square matrix of order n. The following conditions are mutually equivalent.*

i) All principal subdeterminants of M are strictly positive.

ii) The LCP (q, M) has a unique solution for each $q \in \mathbf{R}^n$.

iii) The LCP (q, M) has at most one solution for each $q \in \mathbf{R}^n$.

iv) The diagonal entries in all PPTs of M are strictly positive.

v) LCPs associated with M satisfy the strict separation property.

vi) The class of complementary cones $\mathcal{C}(M)$ forms a partition of $\mathbf{R}^n$.

Proof. Follows from Theorems 3.15, 3.16, 3.17, 3.6 and Corollaries 3.6, 3.8. □

Theorem 3.18 *Consider the LCP (3.1) in which M is a P-matrix. Suppose $(\overline{w}, \overline{z})$ is the unique solution of the LCP with $\overline{z}_1 = 0$. Let $\omega = (w_2, \ldots, w_n)$, $\xi = (z_2, \ldots, z_n)$. If $(y_2, \ldots, y_n)$, with $y_j \in \{w_j, z_j\}$ for $j = 2$ to n, is a complementary feasible basic vector for the principal subproblem of (3.1) in ω, ξ; $(w_1, y_2, \ldots, y_n)$ is a complementary feasible basic vector for (3.1).*

Proof. By Result 2.2 and Theorem 3.13, $\overline{\omega} = (\overline{w}_2, \ldots, \overline{w}_n)$, $\overline{\xi} = (\overline{z}_2, \ldots, \overline{z}_n)$ is the unique solution of the principal subproblem in ω, ξ. Since $\overline{w}$, $\overline{z}$ is the unique solution of (3.1), and $\overline{z}_1 = 0$, we have $\sum_{j=2}^{n} m_{2j}\overline{z}_j + q_1 = \overline{w}_1 \geqq 0$. Under degeneracy, there may be several complementary feasible basic vectors (all differing in the zero valued basic variables) for the principal subproblem in ω, ξ, but the BFS corresponding to each of them must be $\overline{\omega}$, $\overline{\xi}$ by the uniqueness of the solution. Also, the column vector of w_1 in (3.1) is $I_{.1}$. So, when we compute the basic solution of (3.1) corresponding to the basic vector $(w_1, y_2, \ldots, y_n)$, we get $w_j = \overline{w}_j$, $z_j = \overline{z}_j$ for $j = 2$ to n, $\overline{z}_1 = 0$ and $w_1 = \sum_{j=2}^{n} m_{2j}\overline{z}_j + q_1 = \overline{w}_1 \geqq 0$, which is the solution $(\overline{w}, \overline{z})$ of (3.1). So, $(w_1, y_2, \ldots, y_n)$ is a complementary feasible basic vector for (3.1).

□

Higher Order Separation Theorems

Theorem 3.19 *Let M be a P-matrix of order n and let $\mathbf{J}$, $\overline{\mathbf{J}}$ be a partition of $\{1, \ldots, n\}$ with $\mathbf{J}$, $\overline{\mathbf{J}}$ both being nonempty. Let $\{A_{.j} : j \in \mathbf{J}\}$, $\{A_{.j} : j \in \overline{\mathbf{J}}\}$ be the corresponding partition of a complementary set of vectors. Let $\{B_{.j} : j \in \overline{\mathbf{J}}\}$ be the complement of the subcomplementary set $\{A_{.j} : j \in \overline{\mathbf{J}}\}$. If $\mathbf{H}$ is a hyperplane in $\mathbf{R}^n$ satisfying*

i) $\mathbf{H}$ contains the origin 0 and all the vectors in the subcomplementary sets $\{A_{.j} : j \in \mathbf{J}\}$.

ii) All the vectors in the subcomplementary set $\{A_{.j} : j \in \overline{\mathbf{J}}\}$ lie in one of the closed half-spaces, $\mathbf{H}^{\geqq}$, defined by $\mathbf{H}$, then at least one of the vectors in $\{B_{.j} : j \in \overline{\mathbf{J}}\}$ lies strictly on the other side of $\mathbf{H}$ in the other open half-space $\mathbf{H}^{<}$ defined by $\mathbf{H}$.

Proof. Consider the system (3.13)

$$w - Mz = 0\,. \tag{3.13}$$

Perform principal pivot steps in (3.13) to transform the complementary set of vectors $\{A_{.j} : j \in \mathbf{J} \cup \overline{\mathbf{J}}\}$ into the set of unit vectors. This is a nonsingular linear transformation that preserves separation properties. If u_j denotes the variable in (3.13) associated with $A_{.j}$, and v_j denotes its complement, this transforms (3.13) into

$$u - \overline{M}v = 0 \tag{3.14}$$

where $\overline{M}$ is also a P-matrix because it is a principal pivot transform of the P-matrix M. Let $\overline{M}_{\overline{\mathbf{J}}\overline{\mathbf{J}}}$ denote the principal submatrix of $\overline{M}$ corresponding to the subset $\overline{\mathbf{J}}$. Let

$\overline{\mathbf{H}} = \{x : \sum_{j=1}^{n} a_j x_j = 0\}$ be the transform of $\mathbf{H}$. Since $A_{.j}$ is transformed into $I_{.j}$, by (i) we have $a_j = 0$ for each $j \in \mathbf{J}$, and by (ii) we have $a_{\overline{\mathbf{J}}} = (a_j : j \in \overline{\mathbf{J}}) \geqq 0$. So the row vector $a = (a_j) \geqq 0$ and since $\overline{\mathbf{H}}$ is a hyperplane $a \geq 0$, that is $a_{\overline{\mathbf{J}}} \geq 0$. (A vector $y = (y_j) \geq 0$ means that each y_j is nonnegative and at least one y_j is strictly positive). For $j \in \overline{\mathbf{J}}$, $B_{.j}$ is now transformed into $-\overline{M}_{.j}$. The vector $(a(-\overline{M}_{.j}) : j \in \overline{\mathbf{J}}) = -a_{\overline{\mathbf{J}}}\overline{M}_{\overline{\mathbf{J}}\overline{\mathbf{J}}}$. Since $\overline{M}_{\overline{\mathbf{J}}\overline{\mathbf{J}}}$ is itself a P-matrix and $a_{\overline{\mathbf{J}}} \geq 0$, by Theorem 3.11 at least one of the components of $a_{\overline{\mathbf{J}}}\overline{M}_{\overline{\mathbf{J}}\overline{\mathbf{J}}}$ is strictly positive, that is $a(-\overline{M}_{.j}) < 0$ for at least one $j \in \overline{\mathbf{J}}$. That is, at least one of the $-\overline{M}_{.j}$ for $j \in \overline{\mathbf{J}}$ lies in the open half-space $\overline{\mathbf{H}}^{<} = \{x : \sum_{j=1}^{m} a_j x_j < 0\}$ not containing the unit vectors. In terms of the original space this implies that at least one of the $B_{.j}$, $j \in \overline{\mathbf{J}}$ is contained in the open half-space $\mathbf{H}^{<}$ defined by $\mathbf{H}$ not containing the complementary set of vectors $\{A_{.j} : j \in \mathbf{J} \cup \overline{\mathbf{J}}\}$. □

Theorem 3.20 *Let M be a P-matrix of order n, $\mathbf{J}$ a nonempty proper subset of $\{1, \ldots, n\}$ and let $\{A_{.j} : j \in \mathbf{J}\}$ be a subcomplementary set of vectors. Let $\mathbf{H}$ be a hyperplane in $\mathbf{R}^n$ that contains the origin 0 and all the vectors in the set $\{A_{.j} : j \in \mathbf{J}\}$. Then $\mathbf{H}$ strictly separates at least one pair of the left out complementary pairs of vectors $\{I_{.j}, -M_{.j}\}$ for $j \in \overline{\mathbf{J}} = \{1, \ldots, n\} \setminus \mathbf{J}$.*

Proof. Choose the subcomplementary set $\{A_{.j} : j \in \overline{\mathbf{J}}\}$ arbitrarily and transform the system (3.13) into (3.14) as in the proof of Theorem 3.19. Using the notation in the proof of Theorem 3.19, suppose this transforms $\mathbf{H}$ into $\overline{\mathbf{H}} = \{x : \sum_{j=1}^{n} a_j x_j = 0\}$. Since $A_{.j}$ is transformed into $I_{.j}$ and $\mathbf{H}$ contains $A_{.j}$ for $j \in \mathbf{J}$, $\overline{\mathbf{H}}$ must contain $I_{.j}$ for $j \in \mathbf{J}$, that is $a_j = 0$ for all $j \in \mathbf{J}$. Since $\overline{\mathbf{H}}$ is a hyperplane, we must have $a \neq 0$, that is $a_{\overline{\mathbf{J}}} = (a_j : j \in \overline{\mathbf{J}}) \neq 0$. Define $\overline{M}_{\overline{\mathbf{J}}\overline{\mathbf{J}}}$ as in the proof of Theorem 3.19, it is a P-matrix as noted there. By the sign nonreversal theorem for P-matrices of D. Gale and H. Nikaido, Theorem 3.12, if $(y_j : j \in \overline{\mathbf{J}}) = a_{\overline{\mathbf{J}}}\overline{M}_{\overline{\mathbf{J}}\overline{\mathbf{J}}}$, $a_j y_j > 0$ for at least one $j \in \overline{\mathbf{J}}$. Since $a_j = 0$ for $j \in \mathbf{J}$, these facts imply that there exists at least one $j \in \overline{\mathbf{J}}$ satisfying the property that $aI_{.j}$ and $a(-\overline{M}_{.j})$ have strictly opposite signs, that is $\overline{\mathbf{H}}$ separates the complementary pair of vectors $\{I_{.j}, -\overline{M}_{.j}\}$ strictly. In terms of the original space, this implies that $\mathbf{H}$ strictly separates the complementary pair of vectors $\{I_{.j}, -M_{.j}\}$ for that $j \in \overline{\mathbf{J}}$.

□

Comment 3.1 Theorem 3.2 is from K. G. Murty [3.47, 3.48]. Theorem 3.5 is due to A. W. Tucker [3.78]. The proofs of Theorems 3.7, 3.8 given here are attributed to P. Wolfe. The fact that the LCP (q, M) of order n has a unique solution for all $q \in \mathbf{R}^n$ is originally established [3.69]. The inductive proof of Theorem 3.13 given here, and Theorems 3.14, Corollary 3.6 are from K. G. Murty [3.47, 3.49].

A Variant of the LCP

We now discuss some results from K. G. Murty [3.51]. Let M be a given square matrix of order n and q a given column vector of order n. Let $\mathbf{J}$ be a given subset of $\{1, \ldots, n\}$.

The **generalized** LCP with data q, M, $\mathbf{J}$ is the problem of finding column vectors $w \in \mathbf{R}^n$, $z \in \mathbf{R}^n$ satisfying:

$$\begin{array}{lll} w - Mz = q & & \\ w_j z_j = 0 & \text{for all } j = 1 \text{ to } n & \\ w_j, z_j \geqq 0 & \text{for all } j \notin \mathbf{J} & \\ w_j, z_j \leqq 0 & \text{for all } j \in \mathbf{J}\,. & \end{array} \tag{3.15}$$

We will use the notation $(q, M, \mathbf{J})$ to denote this generalized LCP. Notice that if $\mathbf{J} = \emptyset$, the generalized LCP $(q, M, \emptyset)$ is the same as the usual LCP (q, M) that we have been discussing so far. We will now prove some results about the uniqueness of the solution to this generalized LCP.

Theorem 3.21 *Let M be a given square matrix of order n, and $\mathbf{J}$ a given subset of $\{1, \ldots, n\}$. With M, $\mathbf{J}$ fixed, the generalized LCP $(q, M, \mathbf{J})$ has a unique solution for each $q \in \mathbf{R}^n$ iff M is a P-matrix.*

Proof. In (3.15), make the following transformation of variables: $w_i = u_i$ for $i \notin \mathbf{J}$, $-u_i$ for $i \in \mathbf{J}$; $z_i = v_i$ for $i \notin \mathbf{J}$, $-v_i$ for $i \in \mathbf{J}$. After making these substitutions, multiply both sides of the ith equation in it by -1 for each $i \in \mathbf{J}$. Let $u = (u_1, \ldots, u_n)^T$, $v = (v_1, \ldots, v_n)^T$. After these transformation the problem becomes:

$$\begin{array}{r} u - \overline{M}v = \overline{q} \\ u \geqq 0,\ v \geqq 0 \\ u^T v = 0 \end{array} \tag{3.16}$$

where $\overline{M}$ is the matrix obtained from M by multiplying each entry in the ith row of M by -1 for each $i \in \mathbf{J}$, and then multiplying each entry in the ith column of the resulting matrix by -1 for each $i \in \mathbf{J}$. So the value of a principal subdeterminant of $\overline{M}$ is exactly equal to the corresponding principal subdeterminant of M. Thus $\overline{M}$ is a P-matrix, iff M is. The column vector $\overline{q}$ is obtained by multiplying the ith entry in q by -1 for each $i \in \mathbf{J}$. (3.16) is equivalent to (3.15). If $(\hat{w}, \hat{z})$ is a solution of (3.15), then the corresponding (u, v) obtained as above is a solution of (3.16) and vice versa. But (3.16) is the usual LCP $(\overline{q}, \overline{M})$, and hence by Theorem 3.13 it has a unique solution for each $\overline{q} \in \mathbf{R}^n$ iff $\overline{M}$ is a P-matrix. Consequently (3.15) has a unique solution for each $q \in \mathbf{R}^n$ iff M is a P-matrix.

□

Now let M be a given square matrix of order n, and consider the usual LCP (q, M), (3.1), again. The column vector q is **nondegenerate** in (3.1), if q is not in the linear hull of any set of $(n-1)$ columns of $(I, -M)$. There are $2^{\mathbf{n}}$ complementary sets of column vectors in the LCP (q, M), and number these sets in some order, from $l = 1$ to $2^{\mathbf{n}}$. Let A_l denote the matrix whose columns are the columns in the lth complementary set of column vectors (in that order), for $l = 1$ to $2^{\mathbf{n}}$. If M is a P-matrix, by Corollary 3.1, A_l is nonsingular and hence is a complementary basis for

(3.1), for each $l = 1$ to $2^{\mathbf{n}}$. Let $\mathcal{A}$ denote the set of all these complementary bases, that is $\mathcal{A} = \{A_l : l = 1, 2, \ldots, 2^{\mathbf{n}}\}$.

It is clear from the definitions, that if q is nondegenerate in the LCP (q, M) and A is a complementary basis for the LCP (q, M) and $\hat{q} = (\hat{q}_j) = A^{-1}q$, then $\hat{q}_j \neq 0$ for each $j = 1$ to n. (Since $\hat{q} = A^{-1}q$, we have $q = A\hat{q} = \sum_{j=1}^{n} \hat{q}_j \ A_{.j}$. If $\hat{q}_j = 0$ for some j, then q can be expressed as a linear combination of $(n-1)$ column vectors of $(I \ \vdots \ -M)$, contradicting the hypothesis that q is nondegenerate in (3.1)).

We will now discuss some important results on the LCP (q, M) when M is a P-matrix and q is nondegenerate, from [3.51].

Theorem 3.22 *Let M be a given P-matrix of order n, and let q be nondegenerate in the LCP (q, M). Then for each subset $\mathbf{J} \subset \{1, \ldots, n\}$, there exists a unique complementary basis $A \in \mathcal{A}$ satisfying the property that if $\hat{q} = (\hat{q}_j) = A^{-1}q$, then $\hat{q}_j < 0$ for all $j \in \mathbf{J}$ and $\hat{q}_j > 0$ for all $j \notin \mathbf{J}$.*

Proof. Since q is nondegenerate, for any $A \in \mathcal{A}$ all the components in $A^{-1}q$ are nonzero. Suppose $\hat{q} = A^{-1}q$ is such that $\hat{q}_j < 0$ for all $j \in \mathbf{J}$ and $\hat{q}_j > 0$ for all $j \notin \mathbf{J}$. Let $(y_1, \ldots, y_n)$ be the complementary vector of variables corresponding to the complementary basis A. Let $(\hat{w}; \hat{z})$ be the solution defined by:

$$\begin{array}{rll} & y_j = \hat{q}_j, & \text{for } j = 1 \text{ to } n \\ \text{Complement of} & y_j = 0, & \text{for } j = 1 \text{ to } n \, . \end{array}$$

Then $(\hat{w}, \hat{z})$ is a solution of the generalized LCP $(q, M, \mathbf{J})$. However, by Theorem 3.21, the generalized LCP $(q, M, \mathbf{J})$ has a unique solution, since M is a P-matrix. This implies that there exists a unique complementary basis $A \in \mathcal{A}$ such that if $\hat{q} = A^{-1}q$, then $\hat{q}_j < 0$ for all $j \in \mathbf{J}$ and $\hat{q}_j > 0$ for all $j \notin \mathbf{J}$.

□

Example 3.9

Let

$$\widetilde{M}(3) = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 2 & 2 & 1 \end{pmatrix} \qquad q = \begin{pmatrix} -1 \\ -1 \\ -1 \end{pmatrix} .$$

Here $n = 3$, and there are eight complementary bases. Verify that $\widetilde{M}(3)$ is a P-matrix. The LCP $(q, \widetilde{M}(3))$ corresponding to this data will be discussed in Example 4.1 of Chapter 4. From there, we see that for $A \in \mathcal{A}$, $\overline{q} = A^{-1}q$, the updated right hand side constants vector is as tabulated below.

Complementary Basic Vector Corresponding to the Complementary Basis	$\overline{q}^T =$ Transpose of the Updated Right Hand side Constants Vector
(w_1, w_2, w_3)	$(-1, -1, -1)$
(w_1, w_2, z_3)	$(-1, -1, \ \ 1)$
(w_1, z_2, z_3)	$(-1, \ \ 1, -1)$
(w_1, z_2, w_3)	$(-1, \ \ 1, \ \ 1)$
(z_1, z_2, w_3)	$(\ \ 1, -1, -1)$
(z_1, z_2, z_3)	$(\ \ 1, -1, \ \ 1)$
(z_1, w_2, z_3)	$(\ \ 1, \ \ 1, -1)$
(z_1, w_2, w_3)	$(\ \ 1, \ \ 1, \ \ 1)$

As an example let $\mathbf{J} = \{2\}$. We verify that the complementary basis corresponding to the complementary basic vector (z_1, z_2, z_3) is the unique complementary basis in this problem satisfying the property that the jth updated right hand side constant is negative for $j \in \mathbf{J}$ and positive for $j \notin \mathbf{J}$. In a similar manner, the statement of Theorem 3.22 can be verified to be true in this example for all subsets $\mathbf{J} \subset \{1, 2, 3\}$.

3.3.1 One-to-One Correspondence Between Complementary Bases and Sign Vectors

Given any vector of "+" and "−" sign symbols in $\mathbf{R}^n$, Theorem 3.22 states that if M is a P-matrix of order n and q is nondegenerate in the LCP (q, M), then there exists a unique complementary basis for the LCP (q, M) satisfying the property that the signs of the components in the updated right hand sides constants vector with respect to that complementary basis, are exactly the given vector of signs.

Corollary 3.11 *Let M be a given P-matrix of order n, and let q be a given column vector which is nondegenerate for the LCP (q, M). The number of complementary basis $A \in \mathcal{A}$ such that if $\hat{q} = (\hat{q}_i) = A^{-1}q$, then exactly r of the q_i are strictly negative, is $\binom{n}{r}$. This follows from Theorem 3.22.*

Corollary 3.12 *Let M be a given P-matrix of order n, and let q be a given column vector which is nondegenerate for the LCP (q, M). There is a* **one-to-one correspondence** *between the $2^{\mathbf{n}}$ complementary basic vectors for this problem, and the $2^{\mathbf{n}}$ sign vectors for the components in the updated q. This follows from Theorem 3.22.*

The result in Theorem 3.22 and Corollary 3.12 displays the nice combinatorial structure of the LCP (q, M) when M is a P-matrix and q is nondegenerate. As we move from one complementary basic vector to another, the sign pattern of the components in the updated q vector changes distinctly. The problem of solving the LCP (q, M) in this case, is the same as that of finding the complementary basic vector that corresponds to the sign vector consisting of all + signs under this one-to-one correspondence.

3.4 OTHER CLASSES OF MATRICES IN THE STUDY OF THE LCP

In this section we provide a brief summary of some of the other classes of matrices used by many researchers in the study of the LCP.

The Weak Separation Property

This is a property of the matrix M, and does not depend on the right hand side constants vector q. An LCP associated with the matrix M (or the class of complementary cones $\mathcal{C}(M)$) is said to satisfy the weak separation property if, given any subcomplementary set of column vectors $(A_{.1}, \ldots, A_{.i-1}, A_{.i+1}, \ldots, A_{.n})$, there exists a hyperplane $\mathbf{H}$ in $\mathbf{R}^n$ which contains the points 0, and $A_{.t}$, $t = 1, \ldots, i-1, i+1, \ldots, n$, and separates (not necessarily strictly) the points represented by the left out complementary pair of column vectors $I_{.i}$, $-M_{.i}$. See reference [3.48]. As an example let $M = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. The corresponding complementary cones are drawn in Figure 3.5, verify that the weak separation property holds, but not the strict separation property. Also see Figure 3.6.

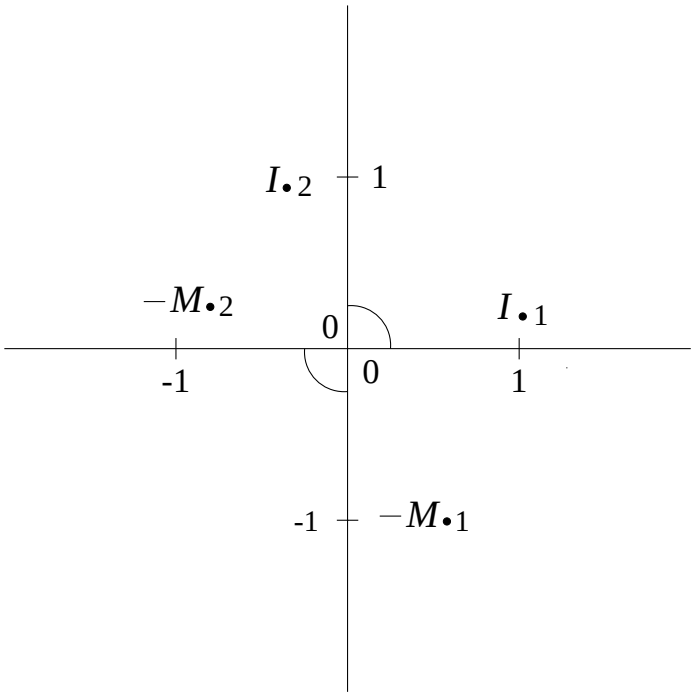


Figure 3.5 The Complementary Cones when $M = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. The Complementary Cones $\text{Pos}\{I_{.1}, -M_{.2}\}$, $\text{Pos}\{-M_{.1}, I_{.2}\}$ are both degenerate, they are the coordinate lines. The Weak Separation Property Holds.

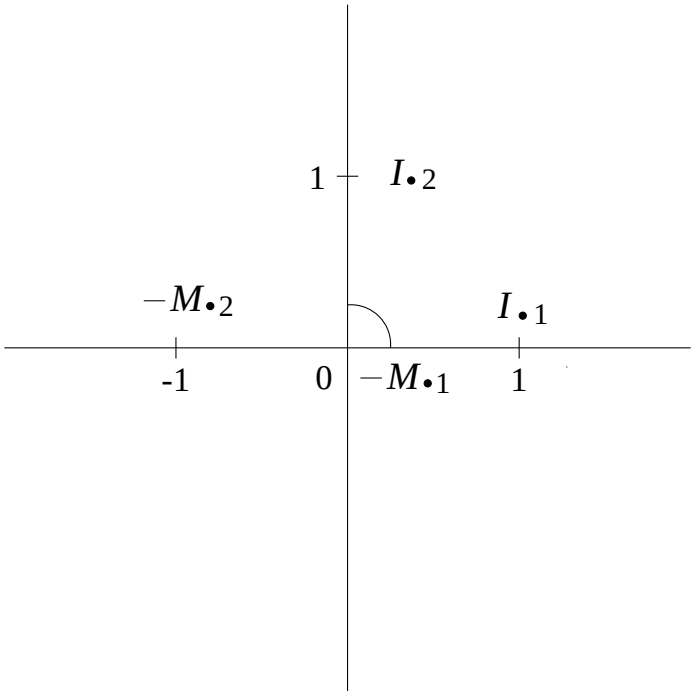


Figure 3.6 The Complementary Cones when $M = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. The Cones $\text{Pos}(I_{.1}, -M_{.2})$, $\text{Pos}(-M_{.1}, -M_{.2})$, $\text{Pos}(-M_{.1}, I_{.2})$ are all degenerate and their Union is the Horizontal Coordinate Line, and the Nonnegative Half of the Vertical Coordinate Line.

The square matrix M of order n is said to be a **weak separation matrix** if it satisfies the weak separation property. Using arguments similar to those in the proof of Theorem 3.16, it can be verified that M is a weak separation matrix iff the diagonal entries in M and all the PPTs of M are nonnegative. See reference [3.48], and also Exercise 3.1.

P_0-Matrices: A square matrix M of order n belongs to this class iff all its principal subdeterminants are $\geqq 0$.

The union of all the complementary cones in $\mathcal{C}(M)$ may not even be convex when M is a P_0-matrix. For example, consider $M = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. The complementary cones for this case are plotted in Figure 3.6. The complementary pivot algorithm may not be able to process the LCP (q, M) when M is a P_0-matrix. For example, on the LCP in which M is the matrix given above, and $q = (-1, 0)^T$, the complementary pivot algorithm ends up in ray termination, even though the LCP has a solution.

Z-Matrices: A square matrix $M = (m_{ij})$ of order n is said to be a Z-matrix iff $m_{ij} \leqq 0$ for every $i \neq j$. A very efficient special algorithm for solving the LCP (q, M) when M is a Z-matrix has been developed by R. Chandrasekaran, and this is discussed in Section 8.1.

Matrices with Dominant Principal Diagonal: A square matrix $M = (m_{ij})$ of order n belongs to this class if $|m_{ii}| > \sum_{\substack{j=1 \\ j \neq i}}^{n} |m_{ij}|$ for each $i = 1$ to n.

Generalized Diagonally Dominant: A square matrix M is said to be a generalized diagonally dominant if there exists a positive diagonal matrix T such that AT is strictly diagonally dominant.

M-Matrices: A square matrix M of order n is said to be an M-matrix if it is a Z-matrix which is also a P_0-matrix. In the literature these matrices are also called K_0-matrices in some references (see S. R. Mohan [3.46]). Nonsingular M-matrices are precisely Z-matrices which are also P-matrices (in the literature these are also known as Minkowski-matrices or K-matrices and some authors refer to these as M-matrices. See the paper [3.22] by M. Fiedler and V. Ptak for the properties of these matrices. If M is a nonsingular M-matrix, then its inverse $M^{-1} \geqq 0$.)

Comparison Matrix: Given a square matrix $M = (m_{ij})$, its comparison matrix is $A = (a_{ij})$ where $a_{ii} = |m_{ii}|$ for $i = 1$ to n and $a_{ij} = -|m_{ij}|$ for all $i \neq j$, $i, j = 1$ to n.

H-Matrix: A square matrix M is said to be a H-matrix if its comparison matrix (which is a Z-matrix) is a P-matrix.

Semi-Monotone Matrices (E_0-Matrices): The square matrix M of order n is said to be semi-monotone iff for all $x \in \mathbf{R}^n$, $x \geq 0$, there exists an index i such that $x_i > 0$ and $M_i.x \geqq 0$. This class of matrices has also been called the class of L_1-matrices. The matrix M belongs to this class iff the LCP (q, M) has a unique solution whenever $q > 0$. If M is symmetric, then it is semi-monotone iff it is copositive.

Strictly Semi-Monotone Matrices: The square matrix M of order n belongs to this class if for every $x \in \mathbf{R}^n$, $x \geq 0$, there exists an index i such that $x_i > 0$ and $M_i.x > 0$. Equivalently, let $\widetilde{M}$ refer to any nonempty principal submatrix of M, or M itself. Then M is strictly semi-monotone, iff the system

$$\begin{aligned} \widetilde{M}\tilde{z} &< 0 \\ \tilde{z} &\geq 0 \end{aligned}$$

has no solution $\tilde{z}$, for all such $\widetilde{M}$. B. C. Eaves [3.21] calls this class of matrices L_*. See also the papers [1.3] of R. W. Cottle and G. B. Dantzig, [1.16] by S. Karamardian, and [3.40] of C. E. Lemke (Lemke calls this class of matrices E).

If M is symmetric, M is strictly semi-monotone iff it is strictly copositive. A matrix M is strictly semi-monotone if the LCP (q, M) has a unique solution whenever $q \geqq 0$. This class is the same as the class of $\overline{Q}$ or completely Q-matrices.

Fully Semi-Monotone: A square matrix of M of order n belongs to this class if M and all its PPTs are semi-monotone. See R. W. Cottle and R. E. Stone [3.13]. The square matrix M is fully semi-monotone iff the LCP (q, M) has a unique solution whenever q is in the interior of any nondegenerate complementary cone.

S-Matrix: A matrix M, not necessarily square, belongs to this class if the system

$$Mx > 0,\ x \geq 0$$

has a solution x. See [3.40] by C. E. Lemke.

$\overline{Q}$, or Completely Q-Matrices: A square matrix of order n belongs to this class if the matrix, and all its principal submatrices are Q-matrices. In [3.9] R. W. Cottle has proved that this class is exactly the same as the class of strictly semi-monotone matrices. See Exercises 3.10, 3.11.

V-Matrices: The square matrix M of order n belongs to this class if every principal submatrix $\widetilde{M}$ of M has the property that there is no positive column vector z such that the last coordinate of $\widetilde{M}\tilde{z}$ is nonpositive and the remaining ones are zero. In [2.38] L. Van der Heyden constructed a new algorithm for the LCP and showed that it will always obtain a solution to the LCP (q, M), provided M is a V-matrix. In [3.9] R. W. Cottle has proved that this class of matrices is the same as the class of strictly semi-monotone matrices, or the class of $\overline{Q}$-matrices. See Exercises 3.10, 3.11.

Q_0-Matrices: A square matrix M of order n belongs to this class if the union of all the complementary cones in $\mathcal{C}(M)$ is convex. In some early papers on the LCP this class was denoted by K. We have the following theorem on this class of matrices.

Theorem 3.23 *If M is a Q_0-matrix, the union of all the complementary cones in $\mathcal{C}(M)$ is* $\text{Pos}(I, -M)$.

Proof. Let $\mathbf{K}(M)$ denote the union of all the complementary cones in $\mathcal{C}(M)$. Every solution of the LCP (q, M) is a $(\overline{w}, \overline{z})$ satisfying $\overline{w} = M\overline{z} + q$, $\overline{w}, \overline{z} \geqq 0$ and $\overline{w}^T\overline{z} = 0$, and hence $(\overline{w}, \overline{z})$ give the coefficients in an expression for q as a nonnegative linear combination of the columns of $(I \ \vdots \ -M)$. So if $q \in \mathbf{K}(M)$, then $q \in \text{Pos}(I \ \vdots \ -M)$, that is, $\mathbf{K}(M) \subset \text{Pos}(I \ \vdots \ -M)$. Now, let $\mathbf{\Gamma} \subset \{I_{.j}, -M_{.j}, j = 1 \text{ to } n\}$. For any $j = 1$ to n, if $q = I_{.j}$, $(w = I_{.j}, z = 0)$ is a solution of the LCP (q, M); and if $q = -M_{.j}$, $(w = 0, z = I_{.j})$ is a solution of the LCP (q, M). So $\mathbf{\Gamma} \subset \mathbf{K}(M)$. Since M is a Q_0-matrix by hypothesis $\mathbf{\Gamma} \subset \mathbf{K}(M)$ implies that $\text{Pos}(\mathbf{\Gamma}) \subset \mathbf{K}(M)$, that is, $\text{Pos}(I \ \vdots \ -M) \subset \mathbf{K}(M)$. All these facts together imply that $\mathbf{K}(M) = \text{Pos}(I \ \vdots \ -M)$.

□

$\overline{Q}_0$-Matrices: The square matrix M of order n belongs to this class if it, and all its principal subdeterminants are Q_0-matrices.

Adequate Matrices: A square matrix of order n belongs to this class if it is a P_0-matrix, and whenever a principal submatrix of M corresponding to a subset $\{i_1, \ldots, i_r\} \subset \{1, \ldots, n\}$ is singular, the sets of vectors $\{M_{i.} : i \in \{i_1, \ldots, i_r\}\}$, $\{M_{.i} : i \in \{i_1, \ldots, i_r\}\}$ are both linearly dependent. This class of matrices has been defined by A. W. Ingleton [3.31]. He proved that if M is adequate, for any $q \in \mathbf{R}^n$, there exists at most one w such that (w, z) is a solution of the LCP (q, M). Also, if M is invertible and adequate, it is a P-matrix.

L_2-Matrices: A square matrix of order n is said to be an L_2-matrix if for each $z \geq 0$ satisfying $w = Mz \geqq 0$ and $w^T z = 0$; there exists a $\hat{z} \neq 0$ satisfying $\hat{w} = -(\hat{z}^T M)^T$, $w \geqq \hat{w} \geqq 0$, $z \geqq \hat{z} \geqq 0$.

$E^*(d)$-Matrices: Let $d \in \mathbf{R}^n$ be given. The square matrix M of order n belongs to this class if $z = 0$ in every solution of the LCP (d, M). Thus if M is an $E^*(d)$ matrix, the LCP (d, M) has the unique solution $(w = d,\ z = 0)$ if $d \geqq 0$, and no solutions if $d \not\geqq 0$.

$E(d)$-Matrices: Let $d \in \mathbf{R}^n$ be given. The square matrix M of order n belongs to this class, if whenever $(\overline{w}, \overline{z})$ is a solution of the LCP (d, M) with $\overline{z} \neq 0$, there exists an $x \geqq 0$ such that $y = -M^T x \geqq 0$, and $\overline{z} \geqq x$, $\overline{w} \geqq y$.

$L(d)$-Matrices: Let $d \in \mathbf{R}^n$ be given. The square matrix M of order n belongs to this class if it is both an $E(d)$-matrix and also an $E(0)$-matrix.

$L^*(d)$-Matrices: Let $d \in \mathbf{R}^n$ be given. The square matrix M of order n belongs to this class if it is both an $E^*(d)$-matrix and also an $E^*(0)$-matrix.

The classes of matrices $E(d)$, $E^*(d)$, $L(d)$, $L^*(d)$ have been defined by C. B. Garcia [3.25]. He has shown that if $d > 0$, and M is an $L(d)$ matrix, then the LCP (q, M) can be processed by the variant of the complementary pivot algorithm in which the original column of the artificial variable z_0 is taken to be $-d$.

Regular Matrices: The square matrix M of order n is said to be a regular matrix (denoted by R-matrix) if there exists no $z \in \mathbf{R}^n$, $t \in \mathbf{R}^1$ satisfying

$$\begin{aligned} z \geqq 0, t \geqq 0 & \\ M_{i.}z + t = 0 & \quad \text{if } i \text{ is such that } z_i > 0 \\ M_{i.}z + t \geqq 0 & \quad \text{if } i \text{ is such that } z_i = 0. \end{aligned}$$

So the matrix M is a regular matrix iff for all $\lambda \geqq 0$, the only solution to the LCP $(\lambda e, M)$ is $(w = \lambda e,\ z = 0)$. S. Karmardian [1.16] introduced this class of matrices and proved that all regular matrices are Q-matrices.

R_0-Matrices: These are matrices M for which the LCP $(0, M)$ has a unique solution. This is exactly the class $E^*(0)$ defined earlier. These matrices have also been called **superregular matrices**. If M belongs to this class there exists no $z \in \mathbf{R}^n$ satisfying

$$\begin{aligned} z \geq 0 & \\ M_{i.}z = 0 & \quad \text{for } i \text{ is such that } z_i > 0 \\ M_{i.}z \geqq 0 & \quad \text{for } i \text{ is such that } z_i = 0. \end{aligned}$$

This class includes all regular matrices. In particular the matrix $M = \begin{pmatrix} -1 & -1 \\ -1 & -1 \end{pmatrix}$ is an R_0-matrix, but not regular.

A degenerate complementary cone $\text{Pos}(A_{.1}, \ldots, A_{.n})$ is said to be **strongly degenerate** if there exists $\alpha = (\alpha_1, \ldots, \alpha_{.n}) \geq 0$ satisfying $\sum_{j=1}^{n} \alpha_j A_{.j} = 0$, **weakly**

degenerate if no such α exists. As an example, let

$$M = \begin{pmatrix} 1 & -1 & -1 \\ 1 & -1 & -1 \\ -1 & 1 & 0 \end{pmatrix} .$$

For this matrix M, the degenerate complementary cone $\text{Pos}(-M_{.1}, -M_{.2}, I_{.3})$ is strongly degenerate because $0 = (-M_{.1}) + (-M_{.2})$. The degenerate complementary cone $\text{Pos}(I_{.1}, I_{.2}, -M_{.3})$, is weakly degenerate since it is impossible to express 0 as $\alpha_1 I_{.1} + \alpha_2 I_{.2} + \alpha_3(-M_{.3})$ with $\alpha_1, \alpha_2, \alpha_3 \geqq 0$ and at least one of $\alpha_1, \alpha_2, \alpha_3$ strictly > 0.

Clearly, a square matrix M is an R_0-matrix iff there exists no strongly degenerate complementary cone in$\mathcal{C}(M)$.

***N*-Matrix:** A square matrix of order n belongs to this class if all its nonempty principal subdeterminants are strictly negative. See M. Kojima and R. Saigal [3.39] in which they prove that if M is an N-matrix, then the LCP (q, M) has either 0, 1, 2 or 3 solutions for any q.

***U*-Matrix:** A square matrix of order n belongs to this class iff the LCP (q, M) has a unique solution whenever q is in the interior of $\mathbf{K}(M)$ = the union of all complementary cones in $\mathcal{C}(M)$. See R. W. Cottle and R. E. Stone [3.13].

INS-Matrices: A square matrix M of order n is said to be an INS-Matrix (Invariant Number of Solutions) iff the number of solutions of the LCP (q, M) is the same for all q contained in the interior of $\mathbf{K}(M)$. See R. W. Cottle and R. E. Stone [3.13], R. E. Stone [3.70, 3.71].

INS$_k$-Matrices: A square matrix M of order n is called an INS$_k$-Matrix if for every q in the interior of $\mathbf{K}(M)$, the LCP (q, M) has exactly k distinct solutions.

***W*-Matrices:** Let M be a given real square matrix of order n. For any $\mathbf{J} \subset \{1, \ldots, n\}$ define the complementary matrix $A(\mathbf{J})$ associated with the subset $\mathbf{J}$ to be the square matrix of order n in which

$$(A(\mathbf{J}))_{.j} = \begin{cases} -M_{.j}, & \text{if } j \in \mathbf{J} \\ I_{.j}, & \text{if } j \notin \mathbf{J} . \end{cases}$$

The matrix M is said to be a W-matrix iff

$$\text{Pos}(A(\mathbf{J})) \cap \text{Pos}(A(\overline{\mathbf{J}})) = \{0\}$$

for every $\mathbf{J} \subset \{1, \ldots, n\}$ and $\overline{\mathbf{J}} = \{1, \ldots, n\} \setminus \mathbf{J}$. This definition is due to M. W. Jeter and W. C. Pye, they have shown that every W-matrix is a U-matrix.

3.5 Exercises

3.1 Let M be a given square matrix of order n. Let $\mathbf{\Gamma} = \{1, \ldots, n\}$. If $\mathbf{S} \subset \mathbf{\Gamma}$ define

$$f(\mathbf{S}) = 1, \quad \text{if } \mathbf{S} = \emptyset$$
$$= \text{principal subdeterminant of } M \text{ corresponding to } \mathbf{S}, \text{ if } \mathbf{S} \neq \emptyset.$$

Prove that M is a weak separation matrix iff there exists no nonempty subset $\mathbf{S} \subset \mathbf{\Gamma}$ satisfying the property that for some $j \in \mathbf{S}$, $f(\mathbf{S})$ and $f(\mathbf{S} \setminus \{j\})$ are both non-zero and have strictly opposite signs. Using it, prove that a square matrix is a weak separation matrix iff the diagonal entries of all its PPTs are $\geqq 0$.

Prove that every nondegenerate weak separation matrix is a P-matrix and that every square matrix which is not a weak separation matrix must have a negative principal subdeterminant. Show that all P_0-matrices are weak separation matrices.

Prove that if the LCP (q, M) has more than one solution, and M is a weak separation matrix, then $q \geqq 0$ (K. G. Murty [3.48, 1.26]).

3.2 Prove that the two definitions given for strictly semi-monotone matrices are equivalent.

3.3 Prove that every copositive plus matrix which contains a strictly positive column vector, is a Q-matrix.

3.4 Prove that all PPTs of a P_0-matrix are P_0-matrices.

3.5 Prove that the square matrix M of order n is a P_0-matrix iff for all $y \in \mathbf{R}^n$, $y \neq 0$, there exists an i such that $y_i \neq 0$ and $y_i(M_{i.}y) \geqq 0$ (Fiedler and Ptak [3.23]).

3.6 If M is a P_0-matrix, prove that there exists an $x \geq 0$ such that $Mx \geqq 0$ (B. C. Eaves [3.21]).

3.7 If M is a P_0-matrix and $x > 0$ satisfies $Mx = 0$, prove that there exists a $y \geq 0$ such that $y^T M = 0$.

3.8 If M is a P_0-matrix and (q, M) has a nondegenerate complementary BFS, then prove that it is the unique complementary feasible solution. Construct a numerical example to show that the converse could be false (B. C. Eaves [3.21]).

3.9 Prove that every Q-matrix is an S-matrix (C. E. Lemke [3.40]).

3.10 Prove that if M is a square matrix of order n which is an S-matrix, and every $(n-1) \times (n-1)$ principal submatrix of M is strictly semi-monotone then M itself is strictly semi-monotone; using this prove that the class of strictly semi-monotone matrices is the same as the class of completely Q-matrices (R. W. Cottle [3.9]).

3.11 Prove that the classes of matrices, strictly semi-monotone, $\overline{Q}$, V, are the same (R. W. Cottle [3.9]).

3.12 If M is a square symmetric matrix of order n, prove that the following conditions are equivalent.

(i) M is strictly copositive,

(ii) M is strictly semi-monotone,

(iii) for all $q \geqq 0$, the LCP (q, M) has a unique solution (F. Pereira [3.59]).

3.13 If M is a square matrix of order n which is principally nondegenerate, prove that the number of complementary feasible solutions for the LCP (q, M) has the same parity (odd or even) for all $q \in \mathbf{R}^n$ which are nondegenerate. As an example, when

$$M = \begin{pmatrix} -1 & 2 & 2 \\ 2 & -1 & 2 \\ 2 & 2 & -1 \end{pmatrix}$$

show that the number of complementary feasible solutions for the LCP (q, M) is always an even number $\geqq 2$ whenever q is nondegenerate (K. G. Murty [1.26, 3.47]).

3.14 Prove that if the number of complementary feasible solutions for the LCP (q, M) is a constant for all q which are nondegenerate, then that constant must be equal to 1, and M must be a P-matrix. (K. G. Murty [1.26, 3.47]).

3.15 If $y^T q + y^T M y$ is bounded below on the set $y \geqq 0$, then prove that the LCP (q, M) has a solution and it can be computed by using the complementary pivot algorithm (B. C. Eaves [3.21]).

3.16 Let q, M be matrices of orders $n \times 1$, $n \times n$ respectively. If there exists an $x \in \mathbf{R}^n$, $x \geq 0$ such that $q^T x < 0$, $M^T x \leqq 0$, prove that the LCP (q, M) has no solution (C. B. Garcia [3.25]).

3.17 Prove that the classes of matrices $E(d)$ and $E^*(d)$ are the same whenever either $d > 0$, or $d < 0$ (C. B. Garcia [3.25]).

3.18 Prove that the semi-monotone class of matrices is $\bigcap_{d>0} E(d)$. Also, prove that

the class L of matrices is $\bigcap_{d>0} L(d)$. Verify that the matrix

$$M = \begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & 1 \\ -1 & -1 & 0 \end{pmatrix}$$

is an $L(d)$ matrix for $d = (2, 3, 1)^T$, but not an L-matrix (C. B. Garcia [3.25]).

3.19 Let $d > 0$ and suppose M is an $L^*(d)$ matrix. For any $q \in \mathbf{R}^n$, prove that when the variant of the complementary pivot algorithm in which the original column of the artificial variable z_0 is taken to be $-d$, is applied on the LCP (q, M), it terminates with a solution of the LCP (S. Karamardian, [1.16], C. B. Garcia [3.25]).

3.20 Let M be a copositive plus matrix. Prove that the set of solutions of the LCP (q, M) is nonempty and bounded iff the optimum objective value in the following LP is zero

$$\begin{array}{lr} \text{Maximize} & e^T u \\ \text{Subject to} & M^T u \geqq 0 \\ & q^T u \leqq 0 \\ & u \geqq 0 \,. \end{array}$$

In particular, prove that if M is copositive plus and the LCP (q, M) has a nondegenerate complementary BFS, then the set of solutions of the LCP (q, M) is bounded (O. L. Mangasarian [3.42]).

3.21 Let M be a copositive plus matrix. If the system: $Mx > 0$, $x \geqq 0$ has a solution $x \in \mathbf{R}^n$, prove that the set of solutions of the LCP (q, M) is nonempty and bounded, for every $q \in \mathbf{R}^n$ (O. L. Mangasarian [3.42], J. Parida and K. L. Roy [3.56]).

3.22 Prove that every regular matrix is a Q-matrix (S. Karamardian [1.16]).

3.23 Prove that if M is a P_0-matrix then the following are equivalent

(i) M is an R_0-matrix,

(ii) M is a regular matrix,

(iii) M is a Q-matrix.

(M. Aganagic and R. W. Cottle [3.2]).

3.24 If M is a P-matrix, prove that the system $Mx > 0$, $x > 0$ has a feasible solution.

3.25 Let M be a P-matrix of order n and let $q \in \mathbf{R}^n$. Consider the quadratic program:

$$\begin{array}{lll} \text{minimize} & z^T(Mz+q) & \\ \text{subject to} & Mz+q \geqq 0 & \\ & z \geqq 0\,. & \end{array} \tag{3.17}$$

Prove the following

(i) (3.17) has a unique local minimum $\overline{z}$ which is the global minimum with objective value 0. In this case $(\overline{w} = M\overline{z}+q, \overline{z})$ is the unique solution of the LCP (q, M).

(ii) If $\overline{z}$ is the unique local minimum for (3.17), let $\pi = \overline{z}^T$, $\mu = (M\overline{z}+q)^T$. Then $(\overline{z}, \overline{\pi}, \overline{\mu})$ is the unique KKT point for (3.17) (Y. C. Chang [3.7]).

3.26 The square matrix M of order n is a nonsingular M-matrix iff the following property holds. Let $(\overline{w}, \overline{z})$ be the solution of the LCP (q, M). Then $\overline{z}$ is the unique vector in the region $\mathbf{X} = \{z : Mz+q \geqq 0, z \geqq 0\}$ satisfying $\overline{z} \in \mathbf{X}$ and $z \geqq \overline{z}$ for any $z \in \mathbf{X}$ (R. W. Cottle and A. F. Veinott, Jr. [3.14]).

3.27 Let M be a Z-matrix which is also a P-matrix of order n, and $q^1, q^2 \in \mathbf{R}^n$ satisfying $q^2 \geqq q^1$. If (w^i, z^i) is a solution of the LCP (q^i, M) for $i = 1, 2$, prove that $z^1 \geqq z^2$ (R. W. Cottle, G. H. Golub, and R. S. Sacher [3.11]).

3.28 Let M be an N-matrix. Then prove that either $M < 0$ or there exists a $d > 0$ such that $Md > 0$. Also prove that a square matrix M is an N-matrix iff all proper principal subdeterminants of M^{-1} are positive and the determinant of M^{-1} is < 0 (M. Kojima and R. Saigal [3.39]).

3.29 Let M be an N-matrix. Prove the following. If $M < 0$, (q, M) has no solutions for $q \ngeqq 0$ and exactly two solutions for $q > 0$. If $M \not< 0$, and $q \not> 0$, the LCP (q, M) has a unique solution. If $M \not< 0$, and $q > 0$, the LCP (q, M) has 2 or 3 solutions. If $M \not< 0$, $q \geq 0$ and $q_i = 0$ for at least one i, the LCP (q, M) has exactly two solutions (M. Kojima and R. Saigal [3.39]).

3.30 If M is an M-matrix prove that the union of all the degenerate complementary cones is the set of all $q \in \mathbf{R}^n$ for which the LCP (q, M) has an infinite number of solutions. Also, in this case, prove that the LCP (q, M) has infinitely many solutions iff q is in the boundary of $\mathbf{K}(M)$, which is the union of all complementary cones in $\mathcal{C}(M)$ (S. R. Mohan [3.46]).

3.31 Prove that every U-matrix is a fully semi-monotone matrix (R. W. Cottle and R. E. Stone [3.13]).

3.32 Prove that the LCP (q, M) has an even number of solutions for each $q \in \mathbf{R}^n$ which is nondegenerate, if there exists a $z > 0$ such that $zM < 0$, or equivalently if $(x = 0, y = 0)$ is the only solution to the system

$$Ix - My = 0$$
$$x, y \geqq 0$$

(R. Saigal [3.63]).

3.33 Consider the LCP (q, M) where M is an adequate matrix. If $(\overline{w}, \overline{z})$, $(\hat{w}, \hat{z})$ are any two solutions of this LCP, prove that $\overline{w} = \hat{w}$ (A. W. Ingleton [3.31]).

3.34 Let M be a square nondegenerate matrix of order n. For some $q^* \in \mathbf{R}^n$, if the LCP (q^*, M) has a unique solution (w^*, z^*) and $w^* + z^* > 0$, then prove that M is a Q-matrix (A. W. Ingleton [3.31]).

3.35 If M is an L-matrix and an R_0-matrix prove that it must also be an R-matrix and a Q-matrix.

Prove that if M is R_0-matrix which is copositive, then it must be an R-matrix and a Q-matrix.

If M is an L_2-matrix and a Q-matrix, prove that it must be an R_0-matrix.

If M is an L-matrix, prove that the following are equivalent:

(i) M is a Q-matrix,
(ii) M is an R-matrix,
(iii) M is an R_0-matrix, and
(iv) M is an S-matrix.

Is every Q-matrix which is an L_1-matrix, also an R_0-matrix? (J. S. Pang [3.53]).

3.36 Prove that copositive plus and strictly copositive matrices are L-matrices.

3.37 Prove that every P_0-matrix is semi-monotone, and that every Q-matrix is an S-matrix.

3.38 If M is an L-matrix, prove that it is a Q-matrix iff it is an S-matrix (B. C. Eaves [3.21]).

3.39 Prove that the system: $Mx = 0, x > 0$, is inconsistent if either M is an L_1-matrix and a Q-matrix, or M is a Q-matrix which is copositive.

If M is an L_1-matrix and a Q-matrix, prove that every nonzero $\overline{z}$ that leads to solution of the LCP $(0, M)$ must have at least two nonzero components.

If M is a Q-matrix which is copositive, prove that any vector $\overline{z}$ satisfying $\overline{z}^T M\overline{z} = 0$ and $(M + M^T)\overline{z} = 0$, and leads to a solution of the LCP $(0, M)$ must be the zero vector.

If M is a Q-matrix which is symmetric and copositive, prove $x = 0$ is the only feasible solution to the system: $Mx = 0$, $x \geqq 0$.

If M is a Q-matrix which is symmetric and copositive plus, prove that it must be strictly copositive.

If M is a copositive plus matrix prove that the following are equivalent:

(i) M is a Q-matrix,
(ii) M is a R-matrix,
(iii) M is a R_0-matrix,
(iv) M is an S-matrix.

In addition, if M is also symmetric, then prove that each of the above is equivalent to

(v) M is strictly copositive,
(vi) $x = 0$ is the only feasible solution of the system: $Mx = 0$, $x \geqq 0$

(J. S. Pang [3.53]).

3.40 Let M be a nondegenerate Q-matrix of order n. Prove that the number of distinct solutions of the LCP (q, M) is $\leqq 2^n - 1$ for any $q \in \mathbf{R}^n$ (A. Tamir [3.75]).

3.41 If M is a square matrix all of whose principle subdeterminants are negative and there exists an $x > 0$ such that $Mx > 0$, then M is a Q-matrix (R. Saigal [3.65]).

3.42 Prove that any square matrix of order 2 with all diagonal entries zero cannot be a Q-matrix. Show that this result is not true for higher order matrices by considering

$$M = \begin{pmatrix} 0 & 3 & -1 & 0 \\ 3 & 0 & 0 & -1 \\ -1 & -1 & 0 & 1 \\ -1 & -1 & 1 & 0 \end{pmatrix}$$

which is a Q-matrix since $M^{-1} > 0$ (M. Jetter and W. Pye [3.33]).

3.43 If M is a square matrix of order n such that there exists a $z > 0$ satisfying $z^T M < 0$ then the LCP (q, M) has an even number of solutions for all nondegenerate q (R. Saigal [3.63]).

3.44 If M is copositive plus and the LCP (q, M) has a solution $(\overline{w}, \overline{z})$ which is a nondegenerate BFS of "$w - MZ = q$, $w \geqq 0$, $z \geqq 0$", prove that the set of solutions of the LCP (q, M) is a bounded set. However, show that the existence of a nondegenerate

BFS solution is not necessary for the set of solutions of the LCP (q, M) to be bounded. (Hint: try $q = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$, $M = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ (O. L. Mangasarian [3.42]).

3.45 If M is a copositive plus matrix of order n, for any $q \in \mathbf{R}^n$, the set of solutions of the LCP (q, M) is nonempty and bounded if the following system has a solution $x \in \mathbf{R}^n$.

$$Mx + q > 0, \; x \geqq 0 \tag{3.18}$$

(O. L. Mangasarian [3.42]).

3.46 If M is a copositive Q-matrix, prove that the system

$$\begin{aligned} Mx &= 0 \\ x &> 0 \end{aligned}$$

is inconsistent.

3.47 If M is a symmetric, copositive plus Q-matrix, prove that M must be strictly copositive (J. S. Pang [3.53]).

3.48 If M is a copositive plus matrix of order n, the solution set of the LCP (q, M) is nonempty and bounded for each $q \in \mathbf{R}^n$ iff M is a Q-matrix. This happens iff the system "$Mx > 0, \; x \geqq 0$" has a solution $x \in \mathbf{R}^n$ (O. L. Mangasarian [3.42]).

3.49 If the nondegenerate matrix M is the limit of a convergent sequence of nondegenerate Q-matrices, prove that M is a Q-matrix (M. Aganagic and R. W. Cottle [3.2]).

3.50 Suppose M is a Q-matrix of order n. Let $\mathbf{J} \subset \{1, \ldots, n\}$ be such that $M_{j.} \geqq 0$ for a $j \in \mathbf{J}$. Then the principal submatrix of M determined by the subset $\{1, \ldots, n\} \setminus \mathbf{J}$ must be a Q-matrix.

3.51 Let M be a Q-matrix of order n. If $\{A_{.1}, \ldots, A_{.j-1}, A_{.j+1}, \ldots, A_{.n}\}$ is a subcomplementary set, there exists a hyperplane $\mathbf{H}$ in $\mathbf{R}^n$ containing 0 and all the vectors in this subcomplementary set such that $I_{.j}$ and $-M_{.j}$ do not lie in the same open halfspace corresponding to this hyperplane $\mathbf{H}$. Also, if M is a nondegenerate Q-matrix, there exists a hyperplane $\mathbf{H}$ of the type described above, which strictly separates $I_{.j}$ and $-M_{.j}$ (M. Aganagic and R. W. Cottle [3.2]).

3.52 If M is a Q_0-matrix satisfying the property that the LCP (q, M) has a unique solution for each q in the interior of $\mathbf{K}(M)$, prove that M must be a P_0-matrix. Also,

if M is a P_0-matrix with only one zero principal subdeterminant and has the property that $\mathbf{K}(M) \neq \mathbf{R}^n$, then prove that $\mathbf{K}(M)$ is a closed half-space and that the LCP (q, M) has a unique solution whenever q is in the interior of $\mathbf{K}(M)$ (R. W. Cottle and R. E. Stone [3.13]).

3.53 If M is a symmetric matrix of order n satisfying

$$m_{ii} > 0 \quad \text{for all } i$$
$$m_{ij} \leqq 0 \quad \text{for all } j \neq i$$

prove that M is copositive iff it is PSD.

3.54 Prove that the LCP (q, M) has a unique solution for all $q > 0$ iff for all $x \geq 0$ there exists an i such that $x_i > 0$, $y = (y_1, \ldots, y_n)^T = Mx$ and $y_i \geqq 0$.

3.55 If M is a symmetric matrix of order n, the following are equivalent

(i) M is copositive;
(ii) for all $x \geq 0$ there exists an i such that $x_i > 0$ and $y = (y_1, \ldots, y_n)^T = Mx$, $y_i \geqq 0$;
(iii) (q, M) has a unique solution for all $q > 0$.

3.56 If M is a symmetric matrix of order n, the following are equivalent

(i) M is strictly copositive;
(ii) M is a Q-matrix and the LCP (q, M) has a unique solution for all $q \in \{I_{.1}, \ldots, I_{.n}\}$ (F. J. Pereira [3.59]).

3.57 Prove that a H-matrix with positive diagonals is a P-matrix (J. S. Pang [3.55]).

3.58 Prove that M-matrices and generalized diagonally dominant matrices are H-matrices.

3.59 Prove that if M is a strictly semi-monotone matrix and q is nondegenerate in the LCP (q, M), then the LCP (q, M) has an odd number of solutions (B. C. Eaves [3.21]).

3.60 Prove that a square matrix M of order n is a Z-matrix iff for each $q \in \mathbf{R}^n$ for which the set $\mathbf{X}(q, M) = \{x : Mx + q \geqq 0,\ x \geqq 0\} \neq \emptyset$, there exists a least element $\tilde{x} \in \mathbf{X}(q, M)$ (given $\mathbf{K} \subset \mathbf{R}^n$, an element $\overline{x} \in \mathbf{K}$ is said to be a least element in $\mathbf{K}$ if $\overline{x} \leqq x$ for all $x \in \mathbf{K}$. If a least element exists, it is clearly unique) satisfying $\tilde{x}^T(M\tilde{x} + q) = 0$ (A. Tamir [3.73]).

3.61 Prove that a square matrix M of order n is a nonsingular M-matrix (i. e., a Z-matrix which is also a P-matrix) iff for each $q \in \mathbf{R}^n$, the set $\mathbf{X}(q, M) = \{x : Mx + q \geqq 0, x \geqq 0\}$ has a least element $\tilde{x}$ which is the only vector in $\mathbf{X}(q, M)$ satisfying $x^T(Mx + q) = 0$ (R. W. Cottle and A. F. Veinott, Jr. [3.14]).

3.62 Prove that a square matrix which has either a zero row or a zero column cannot be a Q-matrix.

3.63 If M is a Q-matrix and PSD, is M^T also a Q-matrix? (Hint: Check $\begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}$).

3.64 Let M be a PSD matrix and A a PD matrix of order n. Let $(w(\varepsilon), z(\varepsilon))$ denote the solution of the LCP $(q, M + \varepsilon A)$ for some $q \in \mathbf{R}^n$ and $\varepsilon > 0$. If the LCP (q, M) has a solution, prove that the $\text{limit}_{\varepsilon \to 0^+} z(\varepsilon)$ exists, and if this limit is $\overline{z}$, it is the point that minimizes the norm $\|Az\|$ in the set $\{z : (w = Mz + q, z)$ is a solution of the LCP $(q, M)\}$. If the LCP (q, M) has no solution, prove that $\text{limit}_{\varepsilon \to 0^+} \|z(\varepsilon)\| = +\infty$ (A. Gana [5.6]).

3.65 Let $-M$ be a Z matrix. A well-known theorem states that if there exists an $x \geqq 0$ such that $x^T M < 0$ in this case, then M^{-1} exists and $-M^{-1} \geqq 0$. Using this theorem, prove the following:

(a) If M satisfies all the above properties, there exist $y_{ij} \geqq 0$ for all i, j such that

$$I_{.j} = \sum_{i=1}^{n} (-y_{ij}) M_{.i}, \quad \text{for all } j\,.$$

(Hint: Use the fact that $M^{-1} \leqq 0$.)

(b) Under the same conditions on M, $\text{Pos}(I \vdots -M) = \text{Pos}(-M)$.

(c) Under the same conditions on M the LCP (q, M) has a solution iff $-M^{-1}q \geqq 0$. Also, if $-M^{-1}q \geqq 0$, then a solution to the LCP is $(w, z) = (0, -M^{-1}q)$ (R. Saigal).

3.66 Let M be a square matrix of order n satisfying the property "if $Mx \leqq 0$, then x must be nonnegative". Prove the following.

(a) M^{-1} must exist.

(b) $-M^{-1} \geqq 0$. (Hint: Use the fact that $(M(M^{-1}))_{.j} = I_{.j} \geqq 0$.)

(c) In this case $\text{Pos}(-M) \supset \text{Pos}(I)$.

3.67 Let M be an arbitrary square matrix of order n. Consider the LCP (q, M). Prove that the following property "the LCP has a solution whenever q is such that the system

$w - Mz = q, w \geqq 0, z \geqq 0$ has a feasible solution and for all such q the LCP has a solution in which $w = 0$" holds iff $\text{Pos}(-M) \supset \text{Pos}(I)$ [i. e., $\text{Pos}(I \vdots -M) = \text{Pos}(-M)$].

Also prove that this property holds iff for all x such that $Mx \leqq 0$, x must be nonnegative (A. K. Rao).

3.68 Let M be a square matrix of order n with non-positive off-diagonal elements. If M is a P-matrix, prove that it has a nonnegative inverse (M. Fiedler and V. Ptak [3.22]).

3.69 Let M be a square matrix of order n. Let $q \in \mathbf{R}^n$. The matrix M is said to be a Q_0-matrix if the LCP (q, M) has a complementary feasible solution whenever the system

$$w - Mz = q$$
$$w \geqq 0, z \geqq 0$$

has a feasible solution.

i) Prove that M is a Q_0-matrix iff the union of all the complementary cones in $\mathcal{C}(M)$ is a convex set.

ii) Prove that the matrix M is a Q_0-matrix iff the LCP (q, M) satisfies: "if $q^1, q^2 \in \mathbf{R}^n$ are such that (q^1, M) has a complementary feasible solution, and $q^2 \geqq q^1$, then (q^2, M) also has a complementary feasible solution" (A. K. Rao).

3.70 If M is a square matrix which is positive semidefinite, and q is nondegenerate in the LCP (q, M), prove that the number of solutions of the LCP (q, M) is either 0 or 1.

3.71 If M is a square matrix of order n which is positive semidefinite, prove that the intersection of the interiors of any pair of complementary cones in $\mathcal{C}(M)$ is empty.

3.72 If M is a square matrix of order n which is positive semidefinite, and q lies in the interior of a complementary cone in $\mathcal{C}(M)$, prove that the LCP (q, M) has a unique solution.

3.73 Let M be a M-matrix (i. e., a Z-matrix which is also a P_0-matrix). Let $w(\varepsilon)$, $z(\varepsilon)$ be the solution of the LCP $(q, M + \varepsilon I)$. If the LCP (q, M) has a solution, prove that $\lim_{\varepsilon \to 0^+} z(\varepsilon)$ exists, and if this limit is $\overline{z}$, it is the least element of $\{z : z \geqq 0, Mz + q \geqq 0\}$ (i. e., $\overline{z} \leqq z$ for all z in this set). If the LCP (q, M) does not have a solution, then $\lim_{\varepsilon \to 0^+} \|z(\varepsilon)\|$ is $+\infty$ (A. Gana [5.6]).

3.74 Consider the LCP (q, M) of order n. Suppose the matrix M is not a P-matrix, but its principal submatrix of order $n - 1$ obtained by deleting row i and column i

from it is a P-matrix for a given i. Discuss an efficient algorithm for computing all the solutions of this LCP (V. C. Prasad and P. K. Sinha [3.60]).

3.75 Let M be a square nondegenerate matrix. Prove that the number of complementary feasible solutions for the LCP (q, M), is either even for all q that are nondegenerate, or odd for all q that are nondegenerate (K. G. Murty [3.50]).

3.76 Given $q \in \mathbf{R}^n$ and a square matrix M of order n, q is said to be **nondegenerate with respect to** M, if q does not lie in the linear hull of any set of $n-1$ or less column vectors of $(I \ \vdots \ -M)$.

Let M be a nondegenerate Q-matrix of order n satisfying the property for some $q \in \mathbf{R}^n$ which is nondegenerate with respect to M, the LCP (q, M) has an odd number of solutions. Prove that small perturbations in the entries of M still leave it as a nondegenerate Q-matrix (A. Tamir).

3.77 Let M be a square matrix of order 2 and let I be the identity matrix of order 2. Prove that M is a Q-matrix iff the LCPs $(-I_{.1}, M)$ and $(-I_{.2}, M)$ both have complementary feasible solutions (L. M. Kelly and L. T. Watson [3.38]).

3.78 Let

$$M = \begin{pmatrix} 1 & -1 & 4 \\ 4 & -3 & 1 \\ 1 & 0.4 & -0.1 \end{pmatrix}, \quad \hat{q} = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$$

and let I be the identity matrix of order 3. Show that the LCPs $(-I_{.1}, M)$, $(-I_{.2}, M)$, $(-I_{.3}, M)$ all have complementary feasible solutions, but the LCP $(\hat{q}, M)$ does not have a complementary feasible solution. This clearly shows that the result in Exercise 3.77 cannot be generalized for $n > 2$ (L. M. Kelly and L. T. Watson [3.38]).

3.79 Consider the following matrix

$$M(\varepsilon) = \begin{pmatrix} 21 & 25 & -27 & -36 - \varepsilon \\ 7 & 3 & -9 & 36 + \varepsilon \\ 12 & 12 & -20 & 0 \\ 4 & 4 & -4 & -8 \end{pmatrix}$$

and let I be the identity matrix of order 4.

(a) Show that $M(\varepsilon)$ is a nondegenerate matrix for all $0 \leqq \varepsilon < 1$.

(b) Show that $M(0)$ is a Q-matrix.

(c) Show that $M(\varepsilon)$ is not a Q-matrix for $0 < \varepsilon < 1$. In particular, let $q(\varepsilon) = (1 - \varepsilon)32I_{.3} + \varepsilon(0.26, -0.02, 30.8, -0.08)^T$. Show that the LCP $(q(\varepsilon), M(\varepsilon))$ has no complementary feasible solution when $0 < \varepsilon < 1$.

These results clearly establish that small perturbations in its elements might change a nondegenerate Q-matrix into a nondegenerate non Q-matrix (L. M. Kelly and L. T. Watson [3.38]).

3.80 Let M be a given square matrix of order n. Prove that the set of complementary feasible solutions for the LCP (q, M) is a bounded set for every $q \in \mathbf{R}^n$, iff $(w, z) = (0, 0)$ is the unique solution of the LCP $(0, M)$.

3.81 The set of nondegenerate Q-matrices is closed in the relative topology of the set of nondegenerate matrices.

Let M be a given nondegenerate Q-matrix of order n. Let $\beta > 0$, and let ∂M be a square matrix of order n satisfying the properties that

a) $M + \lambda\partial M$ is a nondegenerate Q-matrix for all $0 \leqq \lambda < \beta$,
b) $M + \beta\partial M$ is nondegenerate.

Then prove that $M + \beta\partial M$ is also a Q-matrix.

Using the same arguments, prove the following: Suppose $M^1, M^2, \ldots$ is a given infinite sequence of nondegenerate Q-matrices satisfying the property that it converges to a limit, $\overline{M}$. If $\overline{M}$ is also nondegenerate, prove that $\overline{M}$ is a Q-matrix (L. T. Watson [3.79], and M. Aganagic and R. W. Cottle [3.2]).

3.82 Let M be a square matrix of order n satisfying the following properties:

a) $m_{ij} \geqq 0$ for all $i \neq j$, and $m_{ii} \leqq 0$.
b) There exists a row vector $\pi \in \mathbf{R}^n$ satisfying $\pi > 0$ and $\pi M < 0$.

Property b) is easily satisfied by $\pi = e$, if a) holds and $|m_{ii}| > \sum_{j \neq i} m_{ij}$ for each i. Prove the following:

i) If M satisfies properties a), b) above, then $\text{Pos}(I) \subset \text{Pos}(-M)$.
ii) If M satisfies properties a), b) above, then either the LCP (q, M) has a solution in which $w = 0$, or it has no solution at all.
iii) If M satisfies properties a), b) above, the LCP (q, M) has a solution iff

$$\begin{aligned} -Mz &= q \\ z &\geqq 0 \end{aligned}$$

has a solution. And if $\overline{z}$ is a feasible solution of the above system then $(\overline{w} = 0, \overline{z})$ is a solution of the LCP (q, M).
iv) If M satisfies conditions a), b) above, and if $q \geq 0$, the the LCP (q, M) has $2^{\mathbf{n}}$ distinct solutions (R. Saigal [3.64]).

3.83 Consider the LCP (q, M) where M is a square matrix of order n all of whose nonempty principal subdeterminants are strictly negative. Prove the following:

i) The matrix

$$\begin{pmatrix} -1 & 2 & -2 & 2 & \cdots \\ 2 & -1 & 2 & -2 & \cdots \\ -2 & 2 & -1 & 2 & \cdots \\ 2 & -2 & 2 & -1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & -1 \end{pmatrix}$$

satisfies the property that all its nonempty principal subdeterminants are strictly negative.

ii) If all the nonempty principal subdeterminants of M are strictly negative, either $M < 0$ or there exists an $x > 0$ satisfying $Mx > 0$.

(iii) All the nonempty principal subdeterminants of M are strictly negative iff all the proper principal subdeterminants of M^{-1} are strictly positive and the determinant of M^{-1} is strictly negative.

(iv) If all the nonempty principal subdeterminants of M are strictly negative and $M < 0$, then the LCP (q, M) has a solution whenever $q \geqq 0$, and no solution whenever $q \ngeqq 0$. Also when $q > 0$, it has exactly two solutions.

v) If all the nonemtpy principal subdeterminants of M are strictly negative and $M \not< 0$, then the LCP (q, M)

a) has a unique solution whenever $q \ngeqq 0$,

b) has exactly three solutions whenever $q > 0$,

c) has exactly two solutions, with one solution degenerate, whenever $q \geqq 0$ with at least one $q_i = 0$.

Hence establish that any matrix $M \not< 0$ whose nonempty principal subdeterminants are strictly negative, is a Q-matrix.

Also prove that in this case, if $q \geq 0$, and $w_i = 0$ in some solution of the LCP (q, M), then that $w_i > 0$ in all other solutions of the LCP (q, M).

vi) Whenever M is such that all the nonempty principal subdeterminants of M are strictly negative, the LCP (q, M) has either 0, 1, 2 or 3 solutions for any $q \in \mathbf{R}^n$ (M. Kojima and R. Saigal [3.39]).

3.84 If M is a Q-matrix, prove that the system

$$\begin{aligned} Mz &> 0 \\ z &\geqq 0 \end{aligned}$$

has a solution z.

3.85 Let M be a given square matrix of order n. For $j = 1$ to n, let $A_{.j} \in \{I_{.j}, -M_{.j}\}$. Then $(A_{.1}, \ldots, A_{.n})$ is a complementary set of column vectors for the LCP (q, M) and we call the matrix with $A_{.1}, \ldots, A_{.n}$ as its columns in this order, a **complementary submatrix** of $(I \ \vdots \ -M)$. Obviously there are 2^n such matrices, and let these be $A^1, \ldots, A^{2^n}$. On these, some may be nonsingular and some singular. Let there be

l nonsingular complementary submatrices, and let all the $2^n - l$ remaining complementary submatrices be singular. Rearrange the complementary submatrices in the sequence $A^1, \ldots, A^{2^n}$, so that the first l of these are nonsingular, and all the remaining are singular. So the complemenatry cone $\text{Pos}(A^t)$ has a nonempty interior iff $1 \leqq t \leqq l$, and has an empty interior if $l+1 \leqq t \leqq 2^n$.

Prove that M is a Q-matrix iff

$$\bigcup_{t=1}^{l} \text{Pos}(A^t) = \mathbf{R}^n$$

that is, iff the union of all the complementary cones with a nonempty interior is $\mathbf{R}^n$.

3.86 Using the same notation as in Exercise 3.85 for any fixed i between 1 to n, the subcomplementary set of column vectors $(A^t_{.1}, \ldots, A^t_{.i-1}, A^t_{.i+1}, \ldots, A^t_{.n})$ is linearly independent for $1 \leqq t \leqq l$, and let $\mathbf{H}^t_i$ denote the hyperplane in $\mathbf{R}^n$ which is the subspace of $\mathbf{R}^n$ containing all the column vectors in this subcomplementary set.

If there exists an i between 1 to n such that $I_{.i}$ and $-M_{.i}$ are both in one of the open half-spaces determined by $\mathbf{H}^t_i$, for each $t = 1$ to l, then prove that M is not a Q-matrix.

3.87 A Finite Procedure for Checking Whether a Given Square Matrix M of Order n is a Q-Matrix

Using the same notation as in Exercise 3.85, let D^t be $(A^t)^{-1}$ for $t = 1$ to l. For each $t = 1$ to l, select one of the rows of D^t, for example the i_tth for $t = 1$ to l, leading to the set of row vectors $\{D^t_{i_t.} : t = 1 \text{ to } l\}$. For each t, i_t can be chosen in n different ways, and hence there are n^l different sets of row vectors $\{D^t_{i_t.} : t = 1 \text{ to } l\}$ obtained in this manner. For each such sets define the following system of linear inequalities in the variables $q = (q_1, \ldots, q_n)^T$

$$D^t_{i_t.} q < 0, \ t = 1 \text{ to } l\,. \tag{3.19}$$

So there are n^l different systems of inequalities of the form (3.19) depending on the choice of the rows from the matrices D^t.

(i) (3.19) is a system of l strict linear inequalities in n variables $q_1, \ldots, q_n$. Prove that the system (3.19) has a feasible solution q, iff the following system (3.20) is infeasible:

$$\begin{aligned} \sum_{t=1}^{l} \pi_t D^t_{i_t.} &= 0 \\ \sum_{t=1}^{l} \pi_t &= 1 \\ \pi_t &\geqq 0 \text{ for all } t = 1 \text{ to } l \end{aligned} \tag{3.20}$$

that is, it has no feasible solution $\pi = (\pi_t)$.

(ii) Prove that M is a Q-matrix iff each of the n^l systems of the form (3.19) is infeasible, that is, none of them has a feasible solution q.

(iii) Remembering that $l \leq 2^n$, construct a finite procedure for checking whether a given square matrix M of order n is a Q-matrix, using the above results. Comment on the practical usefulness of such a procedure (D. Gale, see [3.2]).

3.88 A square matrix M is called a Q_0-matrix if the union of all complementary cones in $\mathcal{C}(M)$ is a convex set.

(i) Prove that M is a Q_0-matrix iff $w - Mz = q$, $w \geqq 0$, $z \geqq 0$ has a feasible solution implies that the LCP (q, M) has a complementary feasible solution.

(ii) Prove that M is a Q_0-matrix iff

(q^0, M) has a complemenatry feasible solution

implies

(q, M) has a complemenatry feasible solution for all $q \geqq q^0$.

(iii) Prove that every 1×1-matrix is a Q_0 matrix. Also develop necessary and sufficient condition for a 2×2 matrix to be a Q_0-matrix.

(iv) Consider the matrices

$$M = \begin{pmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix} , \quad q = \begin{pmatrix} -10 \\ 2 \\ 5 \end{pmatrix} .$$

Show that $w - Mz = q$, $w \geqq 0$, $z \geqq 0$ has a feasible solution, but the LCP (q, M) has no complementary feasible solution. Also, in this case verify that all the proper principal submatrices of M are Q_0-matrices (by (i), this implies that there are matrices which are not Q_0-matrices, but all of whose proper submatrices are Q_0-matrices).

3.89 A Finite Characterization for Q_0-Matrices

Given a square matrix M of order n, using the notation and results in Exercises 3.85, 3.87, prove that M is a Q_0-matrix iff

$$\text{Pos}(I \ \vdots \ -M) = \bigcup_{t=1}^{l} \text{Pos}(A^t) .$$

Using this, show that M is a Q_0-matrix, iff each of the following n^l systems

$$\begin{aligned} \sum_{t=1}^{l} \pi_t D^t_{i_t .} - \mu &= 0 \\ -\mu M &\geqq 0 \\ \sum_{t=1}^{l} \pi_t &= 1 \\ \pi \geqq 0, \ \mu &\geqq 0 \end{aligned}$$

are infeasible (i. e., none of them have a feasible solution (π, μ)). This provides a method for checking whether a given square matrix of order n, is a Q_0-matrix or not, using at most a finite amount of computation.

3.90 Prove that every PPT of a Q-matrix is a Q-matrix.

3.91 Let M be a square matrix of order n. Prove that all nonempty principal submatrices of M are Q-matrices iff any of the following three equivalent conditions hold.

i) For all nonempty principal submatrices $\overline{M}$ of M (including M itself), the system

$$\begin{aligned} \overline{M}y &\leqq 0 \\ y &\geq 0 \end{aligned}$$

has no solution.

ii) For every vector $x \geq 0$, there exists an index j such that $x_j > 0$ and $(Mx)_j > 0$.

iii) For every $q \geqq 0$ the LCP (q, M) admits the unique solution $(w; z) = (q; 0)$ (R. W. Cottle [3.9]).

3.92 Row and Column Scalings of Matrices

Given a square matrix of order n, multiply its rows by positive numbers $\alpha_1, \ldots, \alpha_n$ respectively. Multiply the columns of the resulting matrix by positive numbers $\beta_1, \ldots, \beta_n$ respectively. The final matrix M', is said to have been obtained from M by row scaling using the positive vector of scales $\alpha = (\alpha_1, \ldots, \alpha_n)$, and column scaling using the positive vector of scales $\beta = (\beta_1, \beta_2, \ldots, \beta_n)$.

(i) Prove that, to every LCP associated with the matrix M; there is a corresponding LCP associated with the matrix M', that can be obtained by dividing each constraint by a suitable positive number and appropriate scaling of the variables (i. e., choose appropriate units for measuring it); and vice versa.

(ii) Prove that M is a P-matrix iff M' is.

(iii) Assume that M is an asymmetric P-matrix which is not a PD matrix. It is possible that M' is PD (e. g., let $M = \begin{pmatrix} 1 & 0 \\ -10 & 1 \end{pmatrix}$. Obtain M' using $\alpha = (100, 1)$, $\beta = (1, 1)$ and verify that the resulting matrix is PD). If M is either a lower triangular or an upper triangular P-matrix, show that positive scale vectors α, β exists, such that the resulting matrix is PD.

(iv) Let

$$M = \begin{pmatrix} 1 & -1 & -3 \\ 1 & 1 & 1 \\ 1 & -3 & \varepsilon \end{pmatrix}$$

where ε is a positive number. Verify that M is a P-matrix. When ε is sufficiently small, prove that there exist no positive scale vectors α, β which will transform this matrix into a PD matrix by scaling.

If M is a P-matrix which is not PSD, the LCP (q, M) is equivalent to the nonconvex quadratic program

$$\begin{array}{lll} \text{Minimize} & z^T(Mz+q) & \\ \text{Subject to} & z & \geqq 0 \\ & Mz+q & \geqq 0\,. \end{array}$$

And yet, if we can find positive row and column scale vectors α, β that will convert M into a PD matrix M' by scaling, this problem can be transformed into an equivalent convex quadratic programming problem. For this reason, the study of scalings of P-matrices that transform them into PD matrices is of interest. Prove that every P-matrix of order 2 can be scaled into a PD-matrix. Characterize the class of P-matrices which can be transformed into PD matrices by scaling (R. Chandrasekaran and K. G. Murty).

3.93 Let D be a given square matrix of order n and let I be the unit matrix of order n. Let c, b be given column vectors in $\mathbf{R}^n$. Let

$$q = \begin{pmatrix} c \\ b \end{pmatrix}\,, \quad M = \begin{pmatrix} D & I \\ -I & 0 \end{pmatrix}\,.$$

With this data, prove that LCP (q, M) always has a solution, and that the solution is unique if D is a P-matrix (B. H. Ahn [9.4]).

3.94 Let M be a Z-matrix of order n. Prove that M is a P-matrix if the LCPs $(0, M)$ and (e_n, M) have unique solutions.

3.95 Let M be a given square matrix of order n, and let D be an arbitrary diagonal matrix with positive diagonal elements. Prove that the following are equivalent.

i) M is a P-matrix.

ii) $(I-E)D+EM$ is a P-matrix for all diagonal matrices $E=(E_{ij})$ of order n satisfying $0 \leqq E_{ii} \leqq 1$ for all i.

iii) $(I-E)D+EM$ is nonsingular for all diagonal matrices $E=(E_{ij})$ of order n satisfying $0 \leqq E_{ii} \leqq 1$ for all i (M. Aganagic [3.1]).

3.96 Develop an efficient method based on the complementary pivot algorithm to check whether a given square matrix is an M-matrix (K. G. Ramamurthy [3.61]).

3.97 Prove that a Z-matrix which is also a Q-matrix must be a P-matrix. Also prove that every M-matrix is a U-matrix.

3.98 Prove that a symmetric matrix is semi-monotone iff it is copositive. Prove that a symmetric matrix M is strictly semi-monotone iff it is strictly copositive.

3.99 If M is a fully semi-monotone matrix and $(\overline{w}, \overline{z})$ is a solution of the LCP (q, M) and $\overline{w} + \overline{z} > 0$, prove that $(\overline{w}, \overline{z})$ is the unique solution of this LCP.

3.100 (Research Problem) Given a square matrix M of order n, develop finite sets of points $\mathbf{\Gamma}_1$ and $\mathbf{\Gamma}_2$ in $\mathbf{R}^n$, constructed using the data in M, satisfying the properties

(i) M is a Q-matrix if the LCP (q, M) has a solution for each $q \in \mathbf{\Gamma}_1$,

(ii) M is a Q_0-matrix if the LCP (q, M) has a solution for each $q \in \mathbf{\Gamma}_2$.

3.101 Let M be a P-matrix of order n. Let $\mathbf{J} \subset \{1, 2, \ldots, n\}$, $\overline{\mathbf{J}} = \{1, 2, \ldots, n\} \setminus \mathbf{J}$. Let $(A_{.j} : j \in \mathbf{J})$ be a subcomplementary vector corresponding to $\mathbf{J}$. For each $j \in \overline{\mathbf{J}}$, let $\{A_{.j}, B_{.j}\} = \{I_{.j}, -M_{.j}\}$. Is the following conjecture — "there exists a hyperplane containing the linear hull of $(A_{.j} : j \in \mathbf{J})$ which separates the convex hull of $\{A_{.j} : j \in \overline{\mathbf{J}}\}$ from the convex hull of $\{B_{.j} : j \in \overline{\mathbf{J}}\}$", — true?

3.102 Let M be a square matrix of order n. M is said to be totally principally degenerate iff all its principal subdeterminants are zero. Prove that M is totally principally degenerate iff it is a principal rearrangement of an upper triangular matrix with zero diagonal elements. Use this to develop an efficient algorithm to check whether a matrix is totally principally degenerate (T. D. Parsons [4.15]).

3.103 Let M be a square matrix of order n which is not an R_0-matrix (i. e., the LCP $(0, M)$ has $(w = 0,\ z = 0)$ as the unique solution). Show that there exists a square matrix $\hat{M} = (\hat{m}_{ij})$ of order n, satisfying

$$\begin{aligned} \hat{m}_{nn} &= 0 \text{ and} \\ \hat{m}_{in} &= 0 \text{ or } 1 \text{ for all } i = 1 \text{ to } n-1 \end{aligned}$$

such that for any $q \in \mathbf{R}^n$, the LCP (q, M) can be transformed into an equivalent LCP $(\hat{q}, \hat{M})$, by performing a block principal pivot step, some principal rearrangements, and row scalings.

Use this to show the following

a) Every Q-matrix of order 2 must be an R_0-matrix.

b) Every Q-matrix which is also a PSD matrix, must be an R_0-matrix.

Verify that the result in (a) does not generalize to $n > 2$, using the matrix

$$M = \begin{pmatrix} -1 & 2 & 1 \\ 2 & -1 & 1 \\ 10 & 10 & 0 \end{pmatrix}.$$

3.6 References

3.1 M. Aganagic, "Newton's Method for Linear Complementary Problems", *Mathematical Programming*, 28 (1984) 349–362.

3.2 M. Aganagic and R. W. Cottle, "A Note on Q-Matrices", *Mathematical Programming*, 16 (1979) 374–377.

3.3 A. Berman, "Matrices and the Linear Complementarity Problem", *Linear Algebra and Its Applications*, 40 (1981) 249–256.

3.4 A. Berman and R. J. Plemons, *Nonnegative Matrices in the Mathematical Sciences,* Academic Press, New York, 1979.

3.5 R. Chandrasekaran, S. N. Kabadi and K. G. Murty, "Some $\mathcal{NP}$-Complete Problems in Linear Programming", *Operations Research Letters*, 1 (1982) 101–104.

3.6 Y. C. Cheng, "Iterative Methods for Solving Linear Complementarity and Linear Programming Problems", Ph. D. Dissertation, Department of Computer Science, University of Wisconsin-Madison, Wisconsin, 1981.

3.7 Y. C. Cheng, "On the Gradient Projection Method for Solving the Nonsymmetric Linear Complementarity Problems", Technical Report, University of Maryland, Baltimore County, Catonsville, Maryland, 1982.

3.8 R. W. Cottle, "Solution Rays for a Class of Complementarity Problems", *Mathematical Programming Study*, 1 (1974) 59–70.

3.9 R. W. Cottle, "Completely Q-Matrices", *Mathematical Programming Study*, 19 (1980) 347–351.

3.10 R. W. Cottle, *Quadratic Programming and Linear Complementarity,* Academic Press, New York (to appear).

3.11 R. W. Cottle, G. H. Golub and R. S. Sacher, "On the Solution of Large Structured Linear Complementarity Problems", *Applied Mathematics and Optimization*, 4 (1978) 347–363.

3.12 R. W. Cottle, G. J. Habetler and C. E. Lemke, "On the Classes of Copositive Matrices", *Linear Algebra and Its Applications*, 3 (1970) 295–310.

3.13 R. W. Cottle and R. E. Stone, "On the Uniqueness of Solutions to Linear Complementarity Problems", *Mathematical Programming*, 27 (1983) 191–213.

3.14 R. W. Cottle and A. F. Veinott, "Polyhedral Sets Having a Least Element", *Mathematical Programming*, 3 (1969) 238–249.

3.15 R. W. Cottle, R. von Randow and R. E. Stone, "On Spherically Convex Subsets and Q-Matrices", *Linear Algebra and Its Applications*, 41 (1981) 73–80.

3.16 G. B. Dantzig, "On Positive Principal Minors", Technical Report 67-1, O. R. House, Stanford University, Stanford, California, 1967.

3.17 R. D. Doverspike, "A Cone Approach to the Linear Complementarity Problem", Ph. D. Dissertation, Department of Mathematical Sciences, Rensselaer Polytechnic Institute, Troy, New York, 1979.

3.18 R. D. Doverspike, "Some Perturbation Results for the Linear Complementarity Problem", *Mathematical Programming*, 23 (1982) 181–192.

3.19 R. D. Doverspike and C. E. Lemke, "A Partial Characterization of a Class of Matrices Defined by Solutions to the Linear Complementarity Problem", *Mathematics of Operations Research*, 7 (1982) 272–294.

3.20 B. C. Eaves, "On the Basic Theorem of Complementarity", *Mathematical Programming*, 1 (1971) 68–75.

3.21 B. C. Eaves, "The Linear Complementarity Problem", *Management Science*, 17 (1971) 612–634.

3.22 M. Fiedler and V. Pták, "On Matrices With Nonpositive Off-Diagonal Elements and Positive Principal Minors", *Czechoslovak Math. J.*, 12 (1962) 382–400.

3.23 M. Fiedler and V. Pták, "Some Generalizations of Positive Definiteness and Monotonicity", *Numerische Mathematik*, 9 (1966) 163–172.

3.24 D. Gale and H. Nikaido, "The Jacobian Matrix and Global Univalence of Mappings", *Mathematische Annalen*, 159 (1965) 81–93.

3.25 G. B. Garcia, "Some Classes of Matrices in Linear Complemenatrity Theory", *Mathematical Progrmaming*, 5 (1973) 299–310.

3.26 G. B. Garcia and F. J. Gould, "Studies in Linear Complemenatrity", Report 8042, Graduate School of Business, University of Chicago, Chicago, 1980.

3.27 G. B. Garcia and W. I. Zangwill, "On Univalence and P-matrices", *Linear Algebra and Its Applications*, 24 (1979) 239–250.

3.28 C. D. Ha, "Stability of the Linear Complementarity Problem at a Solution Point", Virginia Commonwealth University, Virginia, 1983.

3.29 A. J. Hoffman and F. J. R. Pereira, "On Copositive Matrices with -1, 0, 1 Entries", *Journal of Combinatorial Theory A*, 14 (1973) 302–309.

3.30 R. Howe, "Linear Complementarity and the Degree of Mappings", Cowles Foundation Discussion Paper No. 542, Yale University, New Haven, Conneticut, 1980.

3.31 A. W. Ingleton, "A Problem in Linear Inequalities", *Proceedings of the London Mathematical Society, Third Series*, 16 (1969) 519–536.

3.32 A. W. Ingleton, "The Linear Complementarity Problem", *Journal of the London Mathematical Society*, 2 (1970) 330–336.

3.33 M. W. Jeter and W. C. Pye, "Some Properties of Q-Matrices", *Linear Algebra and Its Applications*, 57 (1984) 169–180.

3.34 I. Kaneko, "A Linear Complementarity Problem With an n by $2n$ 'P-Matrix"', *Mathematical Programming Study*, 7 (1978) 120–141.

3.35 I. Kaneko, "The Number of Solutions of a Class of Linear Complementarity Problems", *Mathematical Programming*, 17 (1979) 104–105.

3.36 I. Kaneko, "Isotone Solutions of Parametric Linear Complementarity Problems", *Mathematical Programming*, 12 (1977) 48–59.

3.37 L. M. Kelly and L. T. Watson, "Erratum: Some Perturbation Theorems for Q-Matrices", *SIAM Journal of Applied Mathematics*, 34 (1978) 320–321.

3.38 L. M. Kelly and L. T. Watson, "Q-Matrices and Spherical Geometry", *Linear Algebra and Its Applications*, 25 (1979) 175–189.

3.39 M. Kojima and R. Saigal, "On the Number of Solutions to a Class of Linear Complementarity Problems", *Mathematical Programming*, 17 (1979) 136–139.

3.40 C. E. Lemke, "Recent Results on Complementarity Problems", in: J. B. Rosen, O. L. Mangasarian and K. Ritter (editors), *Nonlinear Programming*, Academic Press, New York, 1970, pp. 349–384.

3.41 O. L. Mangasarian, "Locally Unique Solutions of Quadratic Programs, Linear and Nonlinear Complementarity Problem", *Mathematical Programming*, 19 (1980) 200–212.

3.42 O. L. Mangasarian, "Characterization of Bounded Solutions of Linear Complementarity Problems", *Mathematical Programming Study*, 19 (1982) 153–166.

3.43 N. Megiddo, "On Monotonicity in Parametric Linear Complementarity Problems", *Mathematical Programming*, 12 (1977) 60–66.

3.44 S. R. Mohan, "Parity of Solutions for Linear Complementarity Problem with Z-Matrices", *Opsearch*, 13 (1976) 19–28.

3.45 S. R. Mohan, "Existence of Solution Rays for Linear Complementarity Problems with Z-Matrices", *Mathematical Programming Study*, 7 (1978) 108–119.

3.46 S. R. Mohan, "Degenerate Complementary Cones Induced by a K_0-Matrix", *Mathematical Programming*, 20 (1981) 103–109.

3.47 K. G. Murty, "On the Number of Solutions to the Complementary Quadratic Programming Problem", Ph. D. Dissertation, University of California, Berkeley, California, 1968.

3.48 K. G. Murty, "On Two Related Classes of Complementary Cones", Technical Report 70-7, Department of Industrial Engineering, University of Michigan, Ann Arbor, Michigan, 1970.

3.49 K. G. Murty, "On a Characterization of P-Matrices", *SIAM Journal on Applied Mathematics*, 23 (1972) 47–60.

3.50 K. G. Murty, "On the Number of Solutions of the Complementarity Problem and Spanning Properties of Complementary Cones", *Linear Algebra and Its Applications*, 5 (1972) 65–108.

3.51 K. G. Murty, "On the Linear Complementarity Problem", *Methods of Operations Research*, 31 (1978) 425–439.

3.52 K. G. Murty, "Higher Order Separation Theorems and a Descent Algorithm for P-Matrix LCPs", Technical Report, Department of Industrial and Operations Engineering, University of Michigan, Ann Arbor, Michigan, 1983.

3.53 J. S. Pang, "On Q-Matrices", *Mathematical Programming*, 17 (1979) 243–247.

3.54 J. S. Pang, "A Unification of Two Classes of Q-Matrices", *Mathematical Programming*, 20 (1981) 348–352.

3.55 J. S. Pang, "The Implicit Complementarity Problem: Part II", Management Science Research Report 459, Graduate School of Industrial Administration, Carnegie-Mellon University, Pittsburgh, Pennsylvania, 1980.

3.56 J. Parida and K. L. Roy, "A Note on the Linear Complementarity Problem", Department of Mathematics, Regional Engineering College, Rourkela — 769008, India.

3.57 J. Parida and B. Sahoo, "On the Complementarity Problem Over a Polyhedral Cone", *Indian Journal of Pure and Applied Mathematics*, 11 (1980) 838–844.

3.58 T. Parthasarathy, *On Global Univalence Theorems, Lecture Notes in Mathematics* 977, Springer-Verlag, New York, 1980.

3.59 F. J. Pereira R., "On Characterizations of Copositive Matrices", Technical Report 72-8, Department of Operations Research, Stanford University, Stanford, California, 1972.

3.60 V. C. Prasad and P. K. Sinha, "Determination of all the Solutions of Certain Linear Complementarity Problems", Department of EE, DEI Engineering College, Dayalbagh, Agra, India, 1981.

3.61 K. G. Ramamurthy, "A Polynomial algorithm for testing the nonnegativity of principal minors of Z-matrices", to appear in *Linear Algebra and its Applications*.

3.62 A. K. Rao, "On the Linear Complementarity Problem", *Management Science*, 27 (1975) 427–429.

3.63 R. Saigal, "An Even Theorem", Technical Report, IIM, Ahmedabad, India, 1970.

3.64 R. Saigal, "A Class of Linear Complementarity Problems", Technical Report, IIM, Ahmedabad, India, 1970.

3.65 R. Saigal, "On the Class of Complementary Cones and Lemke's Algorithm", *SIAM Journal of Applied Mathematics*, 23 (1972) 46–60.

3.66 R. Saigal, "A Characterization of the Constant Parity Property of the Number of Solutions to the Linear Complementarity Problem", *SIAM Journal of Applied Mathematics*, 23 (1972) 40–45.

3.67 R. Saigal and C. B. Simon, "Generic Properties of the Complementarity Problem", *Mathematical Programming*, 4 (1973) 324–335.

3.68 R. Saigal and R. E. Stone, "Proper, Reflecting and Absorbing Facets of Complementary Cones", Technical Report, Department of Industrial Egineering and Management Sciences, Northwestern University, Evanston, Illinois, 1983.

3.69 H. Samelson, R. M. Thrall and O. Wesler, "A Partition Theorem for Euclidean n-space", *Proceedings of the American Mathematical Society*, 9 (1958) 805–807.

3.70 R. E. Stone, "Geometric Aspects of the Linear Complementarity Problem", Technical Report SOL 81-6, Department of Operations Research, Stanford University, Stanford, California, 1981.

3.71 R. E. Stone, "Linear Complementarity Problems with Invariant Number of Solutions", Harvard University Technical Report, 1984.

3.72 D. Solow and P. Sengupta, "A Finite Descent Theory for Linear Programming Piecewise Linear Convex Minimization, and the Linear Complementarity Problem", Technical Report, Department of Operations Research, Case Western Reserve University, Cleveland, Ohio, 1984.

3.73 A. Tamir, "The Complementarity Problem in Mathematical Programming", Ph. D. Dissertation, Department of Operations Research, Case Western Reserve University, Cleveland, Ohio, 1973.

3.74 A. Tamir, "On a Characterization of P-Matrices", *Mathematical Programming*, 4 (1973) 110–112.

3.75 A. Tamir, "On the Number of Solutions to the Linear Complementarity Problem", *Mathematical Programming*, 10 (1976) 347–353.

3.76 M. J. Todd, "Orientation in Complementary Pivot Algorithms", *Mathematics of Operations Research*, 1 (1976) 54–66.

3.77 M. J. Todd, "Complementarity in Oriented Matroids", Technical Report No. 547, School of Operations Research and Industrial Engineering, Cornell University, Ithaca, New York, 1983.

3.78 A. W. Tucker, "Principal Pivot Transforms of Square Matrices", *SIAM Review*, 5 (1963) 305.

3.79 L. T. Watson, "Some Perturbation Theorems for Q-Matrices", *SIAM Journal of Applied Mathematics*, 31 (1976) 379–384.

3.80 W. T. Ziemba, C. Parkan and R. Brooks-Hill, "Calculation of Investment Portfolios with Risk Free Borrowing and Lending", *Management Science*, 21 (1974) 209–222.

Chapter 4

PRINCIPAL PIVOTING METHODS FOR LCP

In this chapter we discuss several methods for solving the LCP based on principal pivot steps. One common feature of these methods is that they do not introduce any artificial variable. These methods employ either single or double principal pivot steps, and are guaranteed to process LCPs associated with P-matrices or PSD-matrices or both. We consider the LCP (q, M) of order n, which is the following in tabular form.

w	z	q
I	$-M$	q

$$w, z \geqq 0, \qquad w^T z = 0 \tag{4.1}$$

4.1 PRINCIPAL PIVOTING METHOD I

This method is most useful for solving LCPs (q, M) in which M is a P-matrix. It only moves among complementary basic vectors for (4.1) which are infeasible, and terminates when a complementary feasible basic vector is obtained. It employs only single principal pivot steps. The initial complementary basic vector for starting the method is $w = (w_1, \ldots, w_n)$.

In this method, the variables may change signs several times during the algorithm, before a complementary solution is obtained in the final step.

In a general step, let $\overline{q} = (\overline{q}_1, \ldots, \overline{q}_n)^T$ be the updated right hand side constants vector in the present canonical tableau of (4.1). If $\overline{q} \geqq 0$, the present complementary

basic vector is feasible and the present BFS of (4.1) is a solution of the LCP (q, M), terminate. If $\overline{q} \not\geq 0$, let

$$r = \text{ Maximum } \{i : i \text{ such that } \overline{q}_i < 0\} . \tag{4.2}$$

Make a single principal pivot step in position r, that is, replace the present basic variable in the complementary pair (w_r, z_r) by its complement. If this pivot step cannot be carried out because the pivot element is zero, the method is unable to continue further, and it terminates without being able to solve this LCP. Otherwise the pivot step is carried out and then the method moves to the next step.

Example 4.1

$$M = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 2 & 2 & 1 \end{pmatrix} \quad q = \begin{pmatrix} -1 \\ -1 \\ -1 \end{pmatrix} .$$

The various canonical tableaus obtained in solving this LCP (q, M) by Principal Pivoting Method I are given below. In each tableau the pivot element is inside a box.

Basic Variable	w_1	w_2	w_3	z_1	z_2	z_3	q
w_1	1	0	0	−1	0	0	−1
w_2	0	1	0	−2	−1	0	−1
w_3	0	0	1	−2	−2	[−1]	−1
w_1	1	0	0	−1	0	0	−1
w_2	0	1	0	−2	[−1]	0	−1
z_3	0	0	−1	2	2	1	1
w_1	1	0	0	−1	0	0	−1
z_2	0	−1	0	2	1	0	1
z_3	0	2	[−1]	−2	0	1	−1
w_1	1	0	0	[−1]	0	0	−1
z_2	0	−1	0	2	1	0	1
w_3	0	−2	1	2	0	−1	1

Basic Variable	w_1	w_2	w_3	z_1	z_2	z_3	q
z_1	-1	0	0	1	0	0	1
z_2	2	-1	0	0	1	0	-1
w_3	2	-2	1	0	0	$\boxed{-1}$	-1
z_1	-1	0	0	1	0	0	1
z_2	2	$\boxed{-1}$	0	0	1	0	-1
z_3	-2	2	-1	0	0	1	1
z_1	-1	0	0	1	0	0	1
w_2	-2	1	0	0	-1	0	1
z_3	2	0	$\boxed{-1}$	0	2	1	-1
z_1	-1	0	0	1	0	0	1
w_2	-2	1	0	0	-1	0	1
w_3	-2	0	1	0	-2	-1	1

The solution of this LCP (q, M) is therefore $(w_1, w_2, w_3; z_1, z_2, z_3) = (0, 1, 1; 1, 0, 0)$.

Theorem 4.1 *Suppose M is a given P-matrix of order n. When Principal Pivoting Method I is applied on the LCP (q, M), it terminates with a complementary feasible basic vector for it in a finite number of pivot steps. Also, a complementary basic vector which appeared once in the course of this method never reappears in subsequent steps.*

Proof. The proof is by induction of n. Since M is a P-matrix, and all the pivot steps in the method are principal pivot steps, by Theorem 3.5, and Corollary 3.5 all the pivot steps required are possible, and the pivot element in all the pivot steps is strictly negative. If n is equal to 1, the theorem is easily verified to be true, and the method terminates after at most one pivot step. We now set up an induction hypothesis.

Induction Hypothesis: Suppose F is a P-matrix of order s and $p \in \mathbf{R}^s$. For $s \leq n-1$, Principal Pivoting Method I applied on the LCP (p, F) solves it in a finite number of pivot steps without cycling.

We will now show that the induction hypothesis implies that Principal Pivoting Method I solves the LCP (q, M) of order n in a finite number of steps without cycling. Consider the principal subproblem of the LCP (q, M) in the variables $\omega = (w_2, \ldots, w_n)^T$, $\xi = (z_2, \ldots, z_n)^T$. If Principal Pivoting Method I is applied on this subproblem, by the induction hypothesis, it terminates in a finite number of pivot steps with a complementary feasible basic vector for it. Let $(y_2^l, y_3^l, \ldots, y_n^l)$, $l = 1$ to

k be the sequence of complementary basic vectors for this subproblem obtained under this method.

When Principal Pivoting Method I is applied on the original LCP (q, M), w_1 is a basic variable in the initial complementary basic vector w. The question of replacing w_1 from the basic vector will only arise for the first time when a complementary basic vector $(w_1, y_2, \ldots, y_n)$ associated with a complementary basis B_1 for (4.1) is reached, satisfying the property that if $\overline{q} = B_1^{-1}q$, then $\overline{q}_2 \geqq 0, \ldots, \overline{q}_n \geqq 0$ (i. e., $(y_2, \ldots, y_n)$ must be a complementary feasible basic vector for the principal subproblem in the variables ω, ξ). When such a complementary basis B_1 is obtained for the first time in the method, if $\overline{q}_1 < 0$, w_1 is replaced from the basic vector by z_1, and the method is continued. On the other hand if $\overline{q}_1 \geqq 0$, B_1 is a complementary feasible basis for (4.1) and the method terminates. Hence the first k basic vectors obtained when Principal Pivoting Method I is applied on (4.1) must be $(w_1, y_2^l, \ldots, y_n^l)$, $l = 1$ to k.

By Theorem 3.13, the LCP (q, M) has a unique solution. Suppose it is $(\hat{w}, \hat{z})$. We consider two possible cases separately.

Case 1: $\hat{z}_1 = 0$.

Since $(y_2^k, \ldots, y_n^k)$ is a complementary feasible basic vector for the principal subproblem in the variables ω, ξ, the hypothesis in this case, and Theorem 3.18 imply that $(w_1, y_2^k, \ldots, y_n^k)$ must be a complementary feasible basic vector for the LCP (q, M). Hence in this case the method solves the LCP (q, M) in a finite number of steps, without cycling.

Case 2: $\hat{z}_1 > 0$.

In this case Theorem 3.13 implies that every complementary basic vector of the form $(w_1, y_2, \ldots, y_n)$ must be an infeasible basic vector for (4.1). Let B_1 be the complementary basis for (4.1) corresponding to $(w_1, y_2^k, \ldots, y_n^k)$. If $\overline{q} = (\overline{q}_i) = B_1^{-1}q$, then $\overline{q}_1 < 0$, $\overline{q}_2 \geqq 0, \ldots, \overline{q}_n \geqq 0$, since $(y_2^k, \ldots, y_n^k)$ is a complementary feasible basic vector for the principal subproblem in the variables ω, ξ. Hence, the next basic vector obtained in Principal Pivoting Method I applied on the LCP (q, M) must be $(z_1, y_2^k, \ldots, y_n^k) = (u_1, \ldots, u_n) = u$, say. Let $v = (v_1, \ldots, v_n)$ where v_j is the complement of u_j, $j = 1$ to n. Let the canonical tableau of (4.1) with respect to u be

u	v	q
I	$-\widetilde{M}$	$\tilde{q}$

(4.3)

$\widetilde{M}$ is the PPT of M corresponding to the complementary basic vector u, and by Theorem 3.5, $\widetilde{M}$ is also a P-matrix. By our assumptions in this case, (4.3) is the system of equations in an LCP $(\tilde{q}, \widetilde{M})$ with (u_i, v_i) as the complementary pair of variables for $i = 1$ to n, which has a unique solution in which v_1 is zero. The subsequent complementary basic vectors obtained in the Principal Pivoting Method I applied on the LCP (q, M) are exactly those which will be obtained when the LCP $(\tilde{q}, \widetilde{M})$,

which has the property that $v_1 = 0$ in its unique solution, is solved by the same method starting with u as the initial complementary basic vector. Applying the result established under Case 1 to (4.3), we conclude that when Principal Pivoting Method I is continued from (4.3), $u_1 = z_1$ remains as a basic variable, and after a finite number of principal pivot steps, a complementary feasible vector will be obtained. Also no cycling ever occurs.

This proves that under the induction hypothesis, the statement of the theorem holds for the LCP (q, M) which is of order n. The theorem has been verified for $n = 1$. Hence it holds for all n by induction.

□

When M is a P-matrix, all the pivot elements in pivot steps encountered under Principal Pivoting Method I applied on the LCP (q, M) will be strictly negative, by Corollary 3.7. If row r is the pivot row, just before this pivot step the updated right hand side constant in row r is strictly negative (and this is the bottom most row with this property at this stage) and just after this pivot step, the updated right hand side constant in row r becomes strictly positive.

The pivot row choice rule (4.2) in Principal Pivoting Method I is only one of the rules which guarantee finite termination when M is a P-matrix. Actually, let $(i_1, i_2, \ldots, i_n)$ be any permutation of $(1, 2, \ldots, n)$. Select this permutation at the beginning of the method arbitrarily, but keep it fixed during the method. Suppose, instead of selecting r as in (4.2), it is selected by the following rule:

$$s = \text{ Maximum } \{t : t \text{ such that } \overline{q}_{i_t} < 0\}, r = i_s \,. \tag{4.4}$$

The rule (4.4) selects row r as the last row in which the updated right hand side constant vector is strictly negative when the rows are listed in the fixed order $(i_1, i_2, \ldots, i_n)$. The rule (4.4) becomes rule (4.2) if the permutation $(i_1, \ldots, i_n)$ is $(1, 2, \ldots, n)$. It can be verified that Principal Pivoting Method I with the rule (4.4) for selecting r, instead of (4.2), again solves the LCP (q, M) in a finite number of steps without cycling, if M is a P-matrix. The proof is very similar to the proof of Theorem 4.1. Instead of looking at the principal subproblem of the LCP (q, M) in the variables $((w_2, \ldots, w_n); (z_2, \ldots, z_n))$, look at the principal subproblem in the variables $((w_1, \ldots, w_{i_1-1}, w_{i_1+1}, \ldots, w_n); (z_1, \ldots, z_{i_1-1}, z_{i_1+1}, \ldots, z_n))$; and change the wording of the induction hypothesis to account for the new rule (4.4) of the choice of r in the method. Row i_1 plays the same role as row 1 did in the proof of Theorem 4.1.

Computational experience indicates that by the proper selection of the permutation of the rows $(i_1, \ldots, i_n)$ and the use of (4.4) for choosing r in the Principal Pivoting Method I, its computational efficiency can be improved substantially. Verify that on the problem in Example 4.1 above, if the permutation of rows $(i_1, i_2, i_3) = (2, 3, 1)$ is used together with the rule (4.4) for the choice of r, Principal Pivoting Method I solves that problem after exactly one pivot step, whereas the original version of the method illustrated in Example 4.1 took seven pivot steps. However, no rules have been developed yet for the choice of the row permutation $(i_1, \ldots, i_n)$ depending on the data

in q, M, to guarantee that Principal Pivoting Method I solves the LCP (q, M) most efficiently.

Even with rule (4.2) for pivot row choice, the performance of Principal Pivoting Method I on LCPs (q, M) in which M is a positive definite matrix, was superior to other methods in computational tests. See, for example, reference [4.11] of M. M. Kostreva. The interesting fact is that when M is a P-matrix, Principal Pivoting Method I solves the LCP (q, M), whether q is degenerate or not, in a finite number of pivot steps without cycling, without the explicit use of any techniques for resolving degeneracy, like perturbation of the right hand side constants vector.

It is not necessary to calculate the canonical tableaus of (4.1) in each pivot step to implement Principal Pivoting Method I. Since it does not require the columns of the basis inverse other than the pivot column in any step, an implementation of this method using either the product form of the inverse, or the elimination form of the inverse would be the most convenient to use, when solving problems on a digital computer. Such an implementation improves the numerical stability and also the computational efficiency of the method.

When M is a general matrix (not a P-matrix), Principal Pivoting Method I may be forced to terminate without obtaining a complementary feasible basic vector for the LCP (q, M) if the required single principal pivot step cannot be performed in some step because the corresponding diagonal element in the PPT of M at that stage is zero. However, if M is a nondegenerate matrix (and not a P-matrix), all the required single principal pivot steps in Principal Pivoting Method I can always be carried out by Theorem 3.4. But in this case the pivot elements in some single principal pivot steps under the method may be strictly positive. In such a pivot step, the updated right hand side constant in the pivot row remains negative even after the pivot step, and if the method is continued after such a pivot step, the same complementary basic vector may reappear and cycling occurs. Thus, Principal Pivoting Method I seems to be most useful only for solving LCPs (q, M) where M is a P-matrix.

Comment 4.1 This method and the finiteness proof for it in the case when M is a P-matrix are taken from [4.14] of K. G. Murty.

4.1.1 Extension to an Algorithm for the Nonlinear Complementarity Problem

In [4.8] G. J. Habetler and M. M. Kostreva have extended the Principal Pivoting Method I into an algorithm for solving the nonlinear complementary problem (1.44). Let $f(x) = (f_1(x), \ldots, f_n(x))^T$, where each $f_i(x)$ is a real valued function defined on $\mathbf{R}^n$. f is said to be a **P-function**, if for all $x \neq y \in \mathbf{R}^n$, there exists an i such that $(x_i - y_i)(f_i(x) - f_i(y)) > 0$. Given f and $\mathbf{J} \subset \{1, 2, \ldots, n\}$ define $g^{\mathbf{J}}(x) = (g_j^{\mathbf{J}}(x))$, where

$$\begin{aligned} g_j^{\mathbf{J}}(x) &= x_j \qquad \text{for } j \notin \mathbf{J} \\ &= f_j(x) \text{ for } j \in \mathbf{J}. \end{aligned} \tag{4.5}$$

The P-function f is said to be a **nondegenerate P-function**, if $g^{\mathbf{J}}(x)$ defined as in (4.5) is a function from $\mathbf{R}^n$ onto $\mathbf{R}^n$ for each subset $\mathbf{J} \subset \{1, \ldots, n\}$. If M is a given square matrix of order n, from Theorems 3.11, 3.12 it follows that the affine function $Mx + q$ is a nondegenerate P-function iff M is a P-matrix.

Consider the system of equations $g^{\mathbf{J}}(x) = 0$. If this system has a solution $\overline{x}$, then $\overline{x}$ is said to be a **complementary point** associated with the subset $\mathbf{J}$. A complementary point $\overline{x}$ clearly satisfies the complementary condition $x^T f(x) = 0$ in (1.44). If $f(x)$ is a nondegenerate P-function, it can be shown (see references [4.8, 4.12, 4.13]) that there exists a unique complementary point associated with any subset $\mathbf{J} \subset \{1, \ldots, n\}$; and that the NLCP (1.44) has a unique complementary feasible solution. The algorithm discussed here is guaranteed to solve the NLCP (1.44) when $f(x)$ is a nondegenerate P-function.

For any $\mathbf{J} \subset \{1, \ldots, n\}$, the solution of the system

$$g^{\mathbf{J}}(x) = 0 \tag{4.6}$$

can be found by iterative methods for solving systems of nonlinear equations such as Newton-Raphson method (see Section 2.7.2 and [10.33]). Newton-Raphson method begins with an initial point x^0 and generates a sequence of points by the iteration $x^{r+1} = x^r - (\nabla g^{\mathbf{J}}(x^r))^{-1} g^{\mathbf{J}}(x^r)$. We will denote the solution of (4.6) by the symbol $\overline{x}(\mathbf{J})$.

The Algorithm

Start with $\mathbf{J} = \emptyset$. In a general step suppose $\mathbf{J}$ is the current subset of $\{1, \ldots, n\}$. Find the associated complementary point $\overline{x}(\mathbf{J})$. If $\overline{x}(\mathbf{J}) + f(\overline{x}(\mathbf{J})) \geqq 0$, then the solution of NLCP (1.44) is $\overline{x}_j(\mathbf{J})$, terminate. If $\overline{x}(\mathbf{J}) + f(\overline{x}(\mathbf{J})) \ngeqq 0$, find $r = \min \cdot \{j : \overline{x}_j(\mathbf{J}) + f(\overline{x}(\mathbf{J}))_j < 0\}$. Define $\tilde{\mathbf{J}} = \mathbf{J} \setminus \{r\}$ if $r \in \mathbf{J}$, $\mathbf{J} \cup \{r\}$ otherwise, go to the next step with $\tilde{\mathbf{J}}$ as the new subset and continue.

In [4.8] G. J. Habetler and M. M. Kostreva have proved that if $f(x)$ is a nondegenerate P-function, this algorithm finds the unique solution of the NLCP (1.44) in a finite number of steps. Computational tests have indicated that this algorithm is quite efficient if implemented with an efficient and robust method for solving systems of nonlinear equations of the form (4.6).

4.1.2 Some Methods which Do not Work For LCP

Y. Bard's Method

A method similar to Principal Pivoting Method I was suggested by Y. Bard (see [4.1], pages 157 – 172). His method is the following: start with $w = (w_1, \ldots, w_n)$ as the initial complementary basic vector.

In a general step, let $\overline{q} = (\overline{q}_1, \ldots, \overline{q}_n)^T$ be the updated right hand side constants vector in the canonical tableau of (4.1) with respect to the current complementary basic vector. If $\overline{q} \geqq 0$, the present BFS of (4.1) is a solution of the LCP (q, M), terminate. If $\overline{q} \ngeqq 0$, let r be such that $\overline{q}_r = \text{Minimum } \{\overline{q}_i : i \text{ such that } \overline{q}_i < 0\}$. If there is a tie, select an r among those tied, arbitrarily. Make a single principal pivot step in position r. If this pivot step cannot be carried out because the pivot element is zero, the method is unable to continue further, and it terminates without being able to solve this LCP. Otherwise the pivot step is carried out, and the method moves to the next step.

This method can cycle even when M is a P-matrix, as this following example constructed by L. Watson indicates.

Example 4.2

Let

$$M = \begin{pmatrix} 10 & 0 & -2 \\ 2 & 0.1 & -0.4 \\ 0 & 0.2 & 0.1 \end{pmatrix}, \quad q = \begin{pmatrix} 10 \\ 1 \\ -1 \end{pmatrix}.$$

It can be verified that M is a P-matrix. When this method is applied on the LCP (q, M), (4.1) with this data, the following complementary basic vectors are obtained.

Complementary Basic Vector	$\overline{q}^T =$ Transpose of the Updated Right Hand Side Constants Vector	$r =$ Position of the Single Principle Pivot Step at this Stage
(w_1, w_2, w_3)	$(10, 1, -1)$	3
(w_1, w_2, z_3)	$(-10, -3, 10)$	1
(z_1, w_2, z_3)	$(1, -1, 10)$	2
(z_1, z_2, z_3)	$(-3, 10, -10)$	3
(z_1, z_2, w_3)	$(-1, 10, 1)$	1
(w_1, z_2, w_3)	$(10, -10, -3)$	2
(w_1, w_2, w_3)	$(10, 1, -1)$	3

Hence the method cycles, even though the choice of r in each step in this example was unambiguous. Let

$$M = \begin{pmatrix} 0.01 & -0.1 & 2 \\ -0.2 & 4.1 & -60 \\ -0.4 & -6.0 & 100 \end{pmatrix}, \quad q = \begin{pmatrix} 0.01 \\ -0.7 \\ 1.0 \end{pmatrix}.$$

Verify that M is PD. Apply Y. Bard's method on the LCP (q, M) with this data and verify that the method cycles, even though the choice of r in each step of the method is unambiguously determined.

The Least Recently Considered Pivot Row Choice Rule for Principal Pivoting Method I

Here the pivot row, r, is chosen by the following. Arrange the rows in any specific order at the beginning of the algorithm, say $1, 2, \ldots, n$, and fix this order. In Step 1, choose the pivot row to be the first row with a negative right hand side constant, when the rows are examined in the specific order $1, 2, \ldots, n$. To choose the pivot row in any subsequent step, identify which row was the pivot row in the previous step. Suppose it was row i. Now examine the rows in the specific order $i+1, \ldots, n, 1, \ldots, i-1$, and choose the first one with a negative updated right hand side constant as the pivot row.

This rule circles through the rows in the specific order beginning with the pivot row of the previous step, until it finds the first row eligible to be the pivot row in this step and chooses it. A rule similar to this for choosing the entering column in the primal simplex algorithm for linear programming problems has been found to make it significantly more efficient. Hence this rule was proposed for the pivot row choice in Principal Pivoting Method I, with the hope that it will be computationally more efficient. With this rule, the method does not work, unfortunately. Consider the LCP (q, M) in Example 4.2. When this method is applied on that problem, it can be verified that it goes through exactly the same pivot steps as in Example 4.2 and cycles.

A Block Pivoting Method for the Linear Complementarity Problem

Let M be a square matrix of order n. Consider the following method for solving the LCP (q, M). Start with any complementary basic vector for (4.1), say, $w = (w_1, \ldots, w_n)$.

In a general step let $y = (y_1, \ldots, y_n)$ be the present complementary basic vector, and let $\overline{q} = (\overline{q}_1, \ldots, \overline{q}_n)$ be the updated right hand side constants vector in the canonical tableau of (4.1) with respect to y. If $\overline{q} \geqq 0$, the present BFS is a solution of the LCP (q, M), terminate. If $\overline{q} \ngeqq 0$, define the complementary vector of variables $u = (u_1, \ldots, u_n)$ by

$$\begin{aligned} u_j &= y_j \text{ if } \overline{q}_j \geqq 0 \\ &= \text{ complement of } y_j, \text{ if } \overline{q}_j < 0\,. \end{aligned}$$

If u is not a complementary basic vector (i. e., if the complementary set of column vectors corresponding to u is linearly dependent), the method terminates without being able to solve this LCP. If u is a complementary basic vector, a block pivot is made to obtain the canonical tableau with respect to the new complementary basic vector u, and the method moves to the next step.

Unfortunately this method can cycle even when M is a P-matrix and q is nondegenerate, as illustrated by the following example constructed by L. Watson. Let:

$$M = \begin{pmatrix} 1 & 0 & -2 \\ -2 & 1 & 4 \\ -4 & 2 & 9 \end{pmatrix}, \quad q = \begin{pmatrix} 1 \\ -1 \\ -3 \end{pmatrix}.$$

When this method is applied on the LCP (q, M) beginning with the basic vector $w = (w_1, w_2, w_3)$, we get the following sequence of basic vectors completing a cycle.

Complementary Basic Vector	$\overline{q}^T$ = Transpose of the Updated Right Hand Side Constants Vector
(w_1, w_2, w_3)	$(1, -1, -3)$
(w_1, z_2, z_3)	$(-1, -3, 1)$
(z_1, w_2, z_3)	$(-3, 1, -1)$
(w_1, w_2, w_3)	$(1, -1, -3)$

In the LCP (q, M) if M is a P-matrix, and q is a nondegenerate the results in Theorem 3.22 indicate that the $2^{\mathbf{n}}$ complementary basic vectors for the problem are in **one to one correspondence** with the $2^{\mathbf{n}}$, n dimensional vectors of $+$ and $-$ sign symbols (these are the signs of the components in the updated right hand side constants vector with respect to the complementary basic vector). The LCP (q, M) is equivalent to the problem of finding the complementary basic vector corresponding to the sign vector consisting of all "+" sign symbols, under this one to one correspondence. This gives the problem a combinatorial flavor. It may be possible to develop an efficient algorithm to solve the LCP (q, M) under these conditions, based on this result.

4.2 THE GRAVES' PRINCIPAL PIVOTING METHOD

We will now discuss a **principal pivoting method** for solving LCPs developed by Robert L. Graves in [4.7]. This method is useful for solving LCPs (q, M) in which M is PSD. Consider the LCP (q, M) where M is a given PSD matrix of order n, (4.1). This method deals only with complementary basic vectors for (4.1), beginning with $w = (w_1, \ldots, w_n)$ as the initial complementary basic vector. It uses only single or double principal pivot steps. All the complementary basic vectors obtained in the method, excepting possibly the terminal one, will be infeasible. When a complementary feasible basic vector for (4.1) is obtained, the method terminates. In this method also, variables may change signs several times during the algorithm.

The method requires a nonsingular square matrix of order n, say B, all of whose rows are lexicopositive initially. Any nonsingular square matrix of order n, whose rows are lexicopositive, can be used as the matrix B in the method. Whenever any pivot steps are carried out on (4.1), the same row operations are also carried out on the matrix B. Even though the row vectors of B are lexicopositive initially, they may not

possess this property subsequently, after one or more pivot steps. In our discussion of this method, **we will choose B to be I**, the identity matrix of order n. When B is chosen as I, the updated B at any stage of the method will be the matrix consisting of the columns of w in the canonical tableau of (4.1) at that stage, and clearly, this will be the inverse of the complementary basis at that stage. Thus choosing B to be I is very convenient, because, all the computations in the method can then be performed efficiently using the basis inverse.

Instead of choosing B as I, if it is choosen as some general nonsingular matrix of order n whose rows are lexicopositive, the method is operated in the same way as below, with the exception that $\beta_i.$ is to be replaced by the ith row of the update of the matrix B. In this general method, the lexicopositivity of B is required so that the statement of the corresponding version of Theorem 4.4 discussed below, holds in Step 1 of this general method. We will now describe the method with $B = I$.

The Graves' Principal Pivoting Method

The initial complementary basic vector is $w = (w_1, \ldots, w_n)$. In a general step, let $y = (y_1, \ldots, y_n)$, where $y_j \in \{w_j, z_j\}$ for each $j = 1$ to n, be the present complementary basic vector. Let $\beta = (\beta_{ij})$ be the inverse of the complementary basis corresponding to y. Let $\overline{q}$ be the present updated right hand side constants vector, that is, $\overline{q} = \beta q$. If $\overline{q} \geqq 0$, y is a complementary feasible basic vector for (4.1) and the present BFS is a solution of the LCP (q, M). Terminate. If $\overline{q} \ngeqq 0$, define the row vector $f = (f_1, \ldots, f_n)$ in this step to be $f =$ lexico maximum $\{\beta_i./\overline{q}_i : i \text{ such that } \overline{q}_i < 0\}$. Since $\beta = (\beta_{ij})$ is nonsingular, this lexico maximum is uniquely determined, and suppose it is attained by $i = r$. So $f = (\beta_r.)/\overline{q}_r$. This is known as the ***f*-vector** in this step. Row r in the canonical tableau of (4.1) with respect to the present complementary basic vector, is known as the **crucial row** in this step. Let t_r denote the complement of y_r and let $A_{.r}$ be the column vector corresponding t_r in the original tableau (4.1). The updated column of t_r is $\overline{A}_{.r} = \beta A_{.r} = (\overline{a}_{1r}, \ldots, \overline{a}_{nr})^T$, say. If $\overline{a}_{rr} \neq 0$, perform a single principal pivot step in position r in the present complementary basic vector y and go to the next step. If

$$\overline{a}_{rr} = 0, \text{ and } \overline{a}_{ir} \leqq 0 \text{ for all } i \tag{4.7}$$

under the assumption that M is PSD, (4.1) does not even have a nonnegative solution (this is proved in Theorem 4.2 below) and hence, the LCP (q, M) has no solutions. Terminate. If $\overline{a}_{rr} = 0$ and $\overline{a}_{ir} > 0$ for at least one i, find lexico maximum $\{(\beta_i. - \overline{q}_i(\beta_r./\overline{q}_r))/\overline{a}_{ir} : i \text{ such that } \overline{a}_{ir} > 0\}$. Let s be the i which attains this lexico maximum (it is shown in Theorem 4.3 bleow, that this s is unique). Perform a double principal pivot step in positions r and s in the present complementary basic vector y (we show in Theorem 4.3 below that this is possible under the assumption that M is PSD), and go to the next step.

Example 4.3

Consider the following LCP associated with a PSD matrix.

w_1	w_2	w_3	w_4	z_1	z_2	z_3	z_4	
1	0	0	0	−1	2	−1	1	−4
0	1	0	0	−2	0	2	−1	−4
0	0	1	0	1	−2	0	3	2
0	0	1	1	−2	1	−3	−3	1
$w_j, z_j \geqq 0$, for all j, $w_j z_j = 0$ for all j								

We denote the f-row in the kth step by f^k. We denote the inverses of the various complementary bases obtained in the method as β^k, $k = 1, 2, \ldots$.

The symbol $\overline{A}_{.j}$ represents the present updated column of the entering variable.

First Inverse Tableau

Basic Variable	$\beta^1 =$ Inverse of the Complementary Basis				Updated q
w_1	1	0	0	0	−4
w_2	0	1	0	0	−4
w_3	0	0	1	0	2
w_4	0	0	0	1	1

Step 1: The f-row in this step is lexico maximum $\{-(1,0,0,0)/4, -(0,1,0,0)/4\} = (0, -1/4, 0, 0)$. So $r = 2$ and row 2 is the crucial row. The present basic variable in the crucial row is w_2, its complement z_2 has the updated column vector $\overline{A}_{.2} = (2, 0, -2, 1)^T$. Since $\overline{a}_{22} = 0$, we compute lexico maximum $\{((1,0,0,0) - (-4)(0,-1/4,0,0))/2, ((0,0,0,1) - (0,-1/4,0,0))\}$ and this is attained by $s = 1$. So we carry out a double principal pivot step in positions 2, 1. This leads to

Second Inverse Tableau

Basic Variable	$\beta^2 =$ Inverse of the Complementary Basis				Updated q
z_1	0	$-1/2$	0	0	2
z_2	$1/2$	$-1/4$	0	0	-1
w_3	1	0	1	0	-2
w_4	$-1/2$	$-3/4$	0	1	6

Step 2: The f-vector here is lexico maximum $\{-(1/2, -1/4, 0, 0), -(1, 0, 1, 0)/2\} = (-1/2, 1/4, 0, 0)$. So $r = 2$ and the second row is the crucial row again. The present basic variable in the crucial row is z_2, its complement w_2 has the updated column $\beta^2 I_{.2} = (-1/2, -1/4, 0, -3/4)^T$. Since $\overline{a}_{22} = -1/4 \neq 0$, we perform a single principal pivot step in position 2. This leads to

Third Inverse Tableau

Basic Variable	$\beta^3 =$ Inverse of the Complementary Basis				Updated q
z_1	-1	0	0	0	4
w_2	-2	1	0	0	4
w_3	1	0	1	0	-2
w_4	-2	0	0	1	9

Step 3: From the third inverse tableau we get $f^3 =$ lexico maximum $\{-(1, 0, 1, 0)/2\} = (-1/2, 0, -1/2, 0)$. So $r = 3$ and the crucial row is row 3 in this step. The basic variable in the crucial row is w_3, and the updated column vector of its complement, z_3, is $\overline{A}_{.3} = \beta^3(-1, 2, 0, -3)^T = (1, 4, -1, -1)^T$. Since $\overline{a}_{33} = -1 \neq 0$, we have to carry out a single principal pivot in position 3 in this step. This leads to

Fourth Inverse Tableau

Basic Variable	β^4 = Inverse of the Complementary Basis				Updated q
z_1	0	0	1	0	2
w_2	2	1	4	0	− 4
z_3	−1	0	−1	0	2
w_4	−3	0	−1	1	11

Step 4: From the fourth inverse tableau we get f^4 = lexico maximum $\{-(2,1,4,0)/4\}$ $= (-1/2, -1/4, -1, 0)$. $r = 2$ and row 2 is the crucial row. w_2 is the present basic variable in the crucial row, the updated column vector of its complement, z_2, is $\overline{A}.2 = \beta^4(2,0,-2,1)^T = (-2,-4,0,-3)^T$. Since $\overline{a}_{22} = -4 \neq 0$, we do a single principal pivot in position 2. This leads to

Fifth Inverse Tableau

Basic Variable	β^5 = Inverse of the Complementary Basis				Updated q
z_1	−1	−1/2	−1	0	4
z_2	−1/2	−1/4	−1	0	1
z_3	−1	0	−1	0	2
w_4	−9/2	−3/4	−4	1	14

Since the updated q vector is now nonnegative, (z_1, z_2, z_3, w_4) is a complementary feasible basic vector. The BFS: $(w_1, w_2, w_3, w_4; z_1, z_2, z_3, z_4) = (0,0,0,14;4,1,2,0)$ is a solution of this LCP. Terminate.

Example 4.4

Consider the LCP for which the original tableau is given below (M can be verified to be a PSD matrix in this problem).

w_1	w_2	w_3	w_4	z_1	z_2	z_3	z_4	q
1	0	0	0	−1	1	−1	−1	2
0	1	0	0	−1	−1	0	−2	0
0	0	1	0	1	0	−1	0	−2
0	0	0	1	1	2	0	0	−1
$w_j, z_j \geqq 0$, and $w_j z_j = 0$, for all j								

Step 1: The initial complementary basic vector is (w_1, w_2, w_3, w_4). We compute $f^1 =$ lexico maximum $\{-(0,0,1,0)/2, -(0,0,0,1)\} = (0,0,0,-1)$. So $r = 4$, and the crucial row is row 4. w_4 is the present basic variable in the crucial row, and the updated column vector of its complement, z_4, is: $\overline{A}_{.4} = (-1,-2,0,0)^T$. $\overline{a}_{44} = 0$, and we find that $\overline{a}_{i4} \leqq 0$ for all i. So condition (4.7) is satisfied in this step. The method therefore terminates with the conclusion that the LCP has no solution. Actually, the constraint corresponding to the fourth row is $w_4 + z_1 + 2z_2 = -1$, which by itself has no nonnegative solution. This clearly implies that this LCP (q, M) has no solution.

Proof of the Method

Theorem 4.2 *If M is* PSD *and condition (4.7) is satisfied in some step of the Graves' principal pivoting method applied on (4.1), there exists no feasible solution to $w - Mz = q$, $w \geqq 0$, $z \geqq 0$.*

Proof. Let $y = (y_1, \ldots, y_n)$ where $y_i \in \{w_j, z_j\}$ for each $j = 1$ to n, be the complementary basic vector in the step in which condition (4.7) is satisfied.

Let $t = (t_1, \ldots, t_n)$ where t_j is the complement of y_j for $j = 1$ to n. Let the canonical tableau with respect to the complementary basic vector y be

y	t	
I	$\overline{A}$	$\overline{q}$

Let row r be the crucial row in this step. By (4.7), $\overline{a}_{rr} = 0$ and $\overline{a}_{ir} \leqq 0$ for all i. Since M is PSD, its PPT $-\overline{A}$ is also PSD, and hence by Result 1.6 $\overline{a}_{ir} + \overline{a}_{ri} = 0$ for all i. So $\overline{a}_{ri} \geqq 0$ for all i. So the equation corresponding to the crucial row, row r, in the present canonical tableau, is $y_r + \sum_{i=1}^{n} \overline{a}_{ri} t_i = \overline{q}_r$. Since $\overline{q}_r < 0$ (as row r is the crucial row) and $\overline{a}_{ri} \geqq 0$ for all i, this by itself has no nonnegative solution. This implies that there exists no (w, z) satisfying $w - Mz = q$, $w \geqq 0$, $z \geqq 0$.

□

Theorem 4.3 *In some step of the Graves' principal pivoting method applied on (4.1), if the crucial row is row r, and a single principal pivot in position r cannot be carried out, and if condition (4.7) is not satisfied, then the position s is determined unambiguously. Also, if M is* PSD, *then a double principal pivot in positions r, s is possible in this step.*

Proof. Let y be the complementary basic vector in the step under discussion. Let $\beta = (\beta_{ij})$ be the inverse of the complementary basis associated with y. Let $\overline{q} = \beta q$. Let $-\overline{A}$ be the PPT of M corresponding to y. The hypothesis in the theorem implies that $\overline{a}_{rr} = 0$. Suppose $i = h$, k both tie for the lexico maximum for determining s. Then $(\beta_{h.} - \overline{q}_h(\beta_{r.}/\overline{q}_r))/\overline{a}_{hr} = \beta_{k.} - (\overline{q}_k(\beta_{r.}/\overline{q}_r))/\overline{a}_{kr}$, which is a contradiction to the nonsingularity of the basis inverse β. So s is determined unambiguously.

Now, let $\overline{A}_{.s}$ be the updated column vector associated with the complement of y_s. The double principal pivot step of replacing y_r, y_s in the complementary basic vector y by their complements, is possible iff the order two determinant $\begin{pmatrix} \overline{a}_{ss} & \overline{a}_{sr} \\ \overline{a}_{rs} & \overline{a}_{rr} \end{pmatrix} \neq 0$. Since $\overline{a}_{rr} = 0$, $\overline{a}_{sr} > 0$ in this case, and $\overline{a}_{rs} = -\overline{a}_{sr} \neq 0$, this order two determinant is nonzero. So the double principal pivot in positions r and s is possible in this step. □

Theorem 4.4 *Let M be a* PSD *matrix. Let $\hat{\beta}$ be the inverse of the complementary basis, and $\hat{q}$ the updated right hand side constants vector, in some step of the Graves' principal pivoting method applied on (4.1). If row l is the crucial row in this step, then*

$$\hat{\beta}_{i.} \succ \hat{q}_i(\hat{\beta}_{l.}/\hat{q}_l) \quad \text{for all } i \neq l\,. \tag{4.8}$$

Proof. Since the method begins with w as the initial complementary basic vector, the inverse of the initial complementary basis is I, all of whose rows are lexicopositive. From this, and from the definition of the crucial row in Step 1 of the method, it can be verified that the statement of the theorem holds true in Step 1 of the method. We now show that if the statement of the theorem holds in a step, say step k, then it also holds in the next step $k+1$.

Suppose $\hat{\beta}$ is the inverse of the complementary basis and $\hat{q}$ the updated right hand side constants vector in step $k+1$ of the method, where $k \geqq 1$. In the previous step, step k, let y be the complementary basic vector, and let β be the inverse of the complementary basis corresponding to y. Let row r be the crucial row in step k. Let $\overline{q} = \beta q$, it is the updated right hand side constants vector in step k. Suppose the statement of the theorem holds true in step k, that is:

$$\beta_{i.} \succ \overline{q}_i(\beta_{r.}/\overline{q}_r) \quad \text{for all } i \neq r\,. \tag{4.9}$$

Let t_j be the complement of y_j for $j = 1$ to n and let $-\overline{A}$ be the PPT of M corresponding to the complementary basic vector y. Since M is PSD, by Theorem 3.10, $-\overline{A}$ is also a PSD matrix. So, by Results 1.5, 1.6 we have: $\overline{a}_{ii} \leqq 0$ for all i, and if $\overline{a}_{ii} = 0$, then $\overline{a}_{ij} + \overline{a}_{ji} = 0$ for all j. Since rows r, l are the crucial rows in steps k, $k+1$, we have $\overline{q}_r < 0$, $\hat{q}_l < 0$.

If the pivot step in step k is a single principal pivot step in position r, we have $\overline{a}_{rr} < 0$, $\hat{q}_r = \overline{q}_r/\overline{a}_{rr} > 0$ (which implies that $l \neq r$, by the above facts), $\hat{\beta}_r. = \beta_r./\overline{a}_{rr}$; $\hat{\beta}_i. = \beta_i. - \beta_r.(\overline{a}_{ir}/\overline{a}_{rr})$, for $i \neq r$; $\hat{q}_i = \overline{q}_i - \overline{q}_r(\overline{a}_{ir}/\overline{a}_{rr})$, for $i \neq r$. From (4.9) we have $\beta_i.\overline{q}_r \prec \overline{q}_i\beta_r.$. This implies that for all $i \neq r$, $(\beta_i. - \beta_r.(\overline{a}_{ir}/\overline{a}_{rr}))\overline{q}_r \prec (\overline{q}_i - \overline{q}_r(\overline{a}_{ir}/\overline{a}_{rr}))\beta_r.$, that is, $\hat{\beta}_i.\overline{q}_r \prec \hat{q}_i\beta_r.$. Since $\overline{a}_{rr} < 0$, this implies that for all $i \neq r$, $\hat{\beta}_i.\overline{q}_r/\overline{a}_{rr} \succ \hat{q}_i\beta_r./\overline{a}_{rr}$. So $\hat{\beta}_i.\hat{q}_r \succ \hat{q}_i\hat{\beta}_r.$, or, $\hat{\beta}_i. \succ \hat{q}_i(\hat{\beta}_r./\hat{q}_r)$, since $\hat{q}_r > 0$, for all $i \neq r$. From this we get $(\hat{\beta}_i./\hat{q}_i) \prec (\hat{\beta}_r./\hat{q}_r)$ for all $i \neq r$ satisfying $\hat{q}_i < 0$. Putting $i = l$ in this (since $\hat{q}_l < 0$) we get $(\hat{\beta}_l./\hat{q}_l) \prec (\hat{\beta}_r./\hat{q}_r)$. This and the previously proved statement that $\hat{\beta}_i. \succ \hat{q}_i(\hat{\beta}_r./\hat{q}_r)$ together imply (4.8) for all $i \neq l$ such that $\hat{q}_i \geqq 0$. For $i \neq l$ such that $\hat{q}_i < 0$, (4.8) holds by the definition of the crucial row in step $k+1$. Thus, in this case, (4.8) holds in step $k+1$ if it holds in step k.

If the pivot in step k is a double principal pivot step in positions r, s, we have $\overline{q}_r < 0$, $\overline{a}_{rr} = 0$, $\overline{a}_{sr} > 0$, $\overline{a}_{rs} = -\overline{a}_{sr} < 0$. It can be verified that this pivot step yields

$$\begin{aligned}
\hat{\beta}_r. &= (\beta_s. - \beta_r.(\overline{a}_{ss}/\overline{a}_{rs}))/\overline{a}_{sr}\ ,\ \hat{\beta}_s. = \beta_r./\overline{a}_{rs} \\
\hat{\beta}_i. &= \beta_i. - \beta_r.(\overline{a}_{is}/\overline{a}_{rs}) - (\beta_s. - \beta_r.(\overline{a}_{ss}/\overline{a}_{rs}))(\overline{a}_{ir}/\overline{a}_{sr})\ ,\ \text{for all } i \neq r, s \\
\hat{q}_r &= (\overline{q}_s - \overline{q}_r(\overline{a}_{ss}/\overline{a}_{rs}))/\overline{a}_{sr}\ , \hat{q}_s = \overline{q}_r/\overline{a}_{rs} \\
\hat{q}_i &= \overline{q}_i - \overline{q}_r(\overline{a}_{is}/\overline{a}_{rs}) - (\overline{q}_s - \overline{q}_r(\overline{a}_{ss}/\overline{a}_{rs}))(\overline{a}_{ir}/\overline{a}_{sr})\ ,\ \text{for all } i \neq r, s\ .
\end{aligned} \tag{4.10}$$

We will now prove that, for $i \neq s$:

$$\hat{\beta}_i. \succ \hat{q}_i(\beta_r./\overline{q}_r)\ . \tag{4.11}$$

First consider the case where $i \neq r$ or s. Substituting for $\hat{\beta}_i$, $\hat{q}_i$ and cancelling common terms, we verify that if $\overline{a}_{ir} \neq 0$,

$$\hat{\beta}_i. - \hat{q}_i\Big(\frac{\beta_r.}{\overline{q}_r}\Big) = \overline{a}_{ir}\left[\frac{1}{\overline{a}_{ir}}\Big(\beta_i. - \overline{q}_i\Big(\frac{\beta_r.}{\overline{q}_r}\Big)\Big) - \frac{1}{\overline{a}_{sr}}\Big(\beta_s. - \overline{q}_s\Big(\frac{\beta_r.}{\overline{q}_r}\Big)\Big)\right]\ . \tag{4.12}$$

If $\overline{a}_{ir} < 0$, from the choice of s and the fact that $\overline{a}_{sr} > 0$, we conclude that the right hand side of (4.12) is lexicopositive and hence (4.11) holds. On the other hand if $\overline{a}_{ir} > 0$, then from the choice of s we conclude that the right hand side of (4.12) is lexicopositive, and hence again (4.11) holds. If $\overline{a}_{ir} = 0$, from (4.10) we have $\hat{\beta}_i. - \hat{q}_i(\beta_r./\overline{q}_r) = \beta_i. - \overline{q}_i(\beta_r./\overline{q}_r)$ and by (4.9) this implies that (4.11) holds in this case too. So (4.11) holds for all $i \neq r, s$. Now consider $i = r$. From (4.10) we verify that $\hat{\beta}_r. - \hat{q}_r(\beta_r./\overline{q}_r) = (\beta_s. - \overline{q}_s(\beta_r./\overline{q}_r))/\overline{a}_{sr} \succ 0$ from (4.9) and the fact that $\overline{a}_{sr} > 0$. So (4.11) holds for $i \neq s$. Since l is the crucial row in step $k+1$, and $\hat{q}_s > 0$, we know that $l \neq s$. So from (4.11) we have $\hat{\beta}_l. \succ \hat{q}_l(\beta_r./\overline{q}_r)$ and since $\hat{q}_l < 0$, this yields $(\hat{\beta}_l./\hat{q}_l) \prec (\beta_r./\overline{q}_r)$. Using this in (4.9) we get $\hat{\beta}_i. \succ \hat{q}_i(\hat{\beta}_l./\hat{q}_l)$ for all i such that $i \neq s$ and $\hat{q}_i > 0$, which yields (4.8) for this i. If i is such that $\hat{q}_i < 0$, (4.8) follows from the choice of the crucial row in step $k+1$, since row l is the crucial row in step $k+1$. If i is such that $i \neq s$ and $\hat{q}_i = 0$, (4.8) follows from (4.11). If $i = s$, from (4.10) we conlcude $(\beta_s./\hat{q}_s) = (\beta_r./\overline{q}_r)$. We have already seen above that $(\beta_r./\overline{q}_r) \succ (\hat{\beta}_l./\hat{q}_l)$. So $(\hat{\beta}_s./\hat{q}_s) \succ (\hat{\beta}_l./\hat{q}_l)$ and since $\hat{q}_s > 0$ this implies (4.8) for $i = s$.

Thus whether the pivot step in step k is a single or double principal pivot step, if the statement of this theorem holds in step k, it holds in step $k+1$. We already verified that the statement of the theorem holds in step 1. Hence it holds in all steps of the method.

□

Theorem 4.5 *The f-vector undergoes a strict lexico-decrease in each step of the method, when applied on the LCP (q, M) where M is* PSD.

Proof. We consider a step in the method, say step k. As in the proof of Theorem 4.4, let β, $\hat{\beta}$ denote the inverse of the complementary bases; and let $\overline{q}$, $\hat{q}$ denote the updated right hand side constant vectors, in steps k, $k+1$ respectively. Let rows r, l be the crucial rows; and let f, $\hat{f}$ denote the f-vectors in steps k, $k+1$ respectively. We wish to prove that $\hat{f} \prec f$. From the definition of the crucial row we have: $f = \beta_r./\overline{q}_r$, $\hat{f} = \hat{\beta}_l./\hat{q}_l$. If the pivot in step k is a single principal pivot step, we have already shown in the proof of Theorem 4.4 that $(\hat{\beta}_l./\hat{q}_l) \prec (\hat{\beta}_r./\hat{q}_r) = (\beta_r./\overline{q}_r)$ which implies that $\hat{f} \prec f$. If the pivot in step k is a double principal pivot step, we have already shown in the proof of Theorem 4.4 that $(\hat{\beta}_l./\hat{q}_l) \prec (\beta_r./\overline{q}_r)$ which implies that $\hat{f} \prec f$. So the f-vector undergoes a strict lexico decrease as the algorithm moves from step k to step $k+1$. So it undergoes a strict lexico decrease in each step of the method.

□

Theorem 4.6 *When M is* PSD, *the Graves' principal pivoting method either finds a solution of the LCP (q, M) or determines that it has no solution, in a finite number of steps.*

Proof. Each complementary basic vector for (4.1) corresponds to a unique f-vector. In each step of the method, if it does not terminate by either finding a complementary feasible basic vector, or by determining that the LCP (q, M) has no solution, the f-vector undergoes a strict lexico decrease, by Theorem 4.5. Hence in each step of the method, a new complementary basic vector is obtained, thus a complementary basic vector obtained in a step of the method, cannot reappear later on. Since there are at most $2^{\mathbf{n}}$-complementary basic vectors for (4.1), the method must terminate by either finding a complementary feasible basic vector (the BFS of (4.1) corresponding to which is a solution of the LCP (q, M)) or by determining that (4.1) does not even have a nonnegative solution, after at most $2^{\mathbf{n}}$ steps.

□

The proof of finite convergence of this method is quite novel, and is based on the fact that the f-vector undergoes a strict lexico decrease in each step. There is no objective function in LCPs and the f-vector is really extraneous to the problem, and yet, since the method guarantees that it undergoes a strict lexico decrease in each step, the method must terminate in a finite number of steps, and the only ways the method can terminate is by either finding a solution of the LCP or by determining that the LCP has no solution.

Theorem 4.7 *If the Graves' principal pivoting method is applied on the LCP (q, M) where M is a P-matrix, then the following statements hold:*

i) All pivot steps will be single principal pivot steps.
ii) In each step the pivot element is always strictly negative.
iii) The method terminates with a solution of the LCP in a finite number of steps without cycling.

Proof. (i) and (ii) follow from Corollary 3.5. It can be verified that the proof of Theorem 4.4 holds in this case too, and hence, the conclusion of Theorems 4.5, 4.6 remain valid here also. This implies (iii).

□

Thus the principal pivoting method discussed above can be applied to process LCPs (q, M) when M is either a PSD matrix or a P-matrix. However, when M is a P-matrix, Principal Pivoting Method I discussed in Section 4.1 will probably be much more efficient since it does not require the rows of the explicit basis inverse, or the determination of the lexico maximum of a set of row vectors in each step. The Graves' principal pivoting method has the advantage of processing LCPs (q, M) which M in PSD and not PD, and Principal Pivoting Method I may not be able to process these problems.

Exercises

4.1 Relationship of the Graves' Principal Pivoting Method to the Simplex Algorithm. Consider the LP (1.9) which can be written as

$$\begin{array}{ll} \text{Minimize} & cx \\ \text{subject to} & Ax - v = b \\ & x \geqq 0, v \geqq 0 \end{array} \tag{4.13}$$

where A is a matrix of order $m \times n$, $v = (v_1, \ldots, v_m)^T$, and $-b \geqq 0$. So v is a feasible basic vector for (4.13). The LCP corresponding to this LP is (q, M) with q, M given as in (1.10). Suppose the Graves' principal pivoting method is applied on (4.13). Then prove the following:

(i) All the pivots steps will be double principal pivot steps.
(ii) The columns of the PPT of M obtained in any step can be rearranged so that it has the structure

$$M' = \begin{pmatrix} 0 & -A'^T \\ A' & 0 \end{pmatrix} .$$

(iii) The rows of the inverse of the basis at the end of each step can be rearranged so that it has the following structure:

$$\beta = \begin{pmatrix} \beta^1 & 0 \\ 0 & \beta^2 \end{pmatrix}$$

where β^1, β^2 square nonsingular matrix of orders n and m respectively.

(iv) If $(\overline{c}, -\overline{b}^T)^T$ is the updated right hand side constants vector in any step, then $-\overline{b}$ is nonnegative.

(v) The sequence of basic solutions obtained in the Graves' principal pivoting method applied on (4.13) can be interpreted as the sequence of primal feasible and dual basic solutions obtained in the various steps of the primal simplex algorithm using the lexico minimum ratio rule for pivot row choice in each step, applied on the LP (4.13) beginning with the primal feasible basic vector v (R. L. Graves [4.7]).

4.2 Consider the quadratic program (1.11) discussed in Section 1.3. If $Q(x)$ is a convex function on $\mathbf{R}^n$, prove that the LCP (1.19) corresponding to it, is an LCP (q, M) in which the matrix M is PSD, and so it can be processed by the Graves' Principal Pivoting Method.

4.3 DANTZIG-COTTLE PRINCIPAL PIVOTING METHOD

This method due to G. B. Dantzig and R. W. Cottle [4.5, 4.6] pre-dates the other principal pivoting methods discussed so far, and evolved from a quadratic programming algorithm of P. Wolfe [4.18] who seems to be the first to use a type of complementary pivot choice rule. The method is useful for processing LCPs (q, M) in which M is either a P-matrix or a PSD matrix. The method goes through a sequence of what are called **major cycles**. Each major cycle begins with a complementary basic vector and ends with a complementary basic vector. Intermediate basic vectors in a major cycle are almost complementary basic vectors of the type discussed in Section 2.4. No artificial variable is introduced, but the original problem variables may take negative or nonnegative values during the method. When a nonnegative solution is obtained, it will be a complementary feasible solution of the LCP (q, M) and the method terminates. Once a variable becomes nonnegative in this method, it remains nonnegative in all subsequent steps (this property distinguishes this method from the other principal pivoting methods discussed so far). Also, if M is a P-matrix or a PD matrix, once a component of the updated q becomes nonnegative in this method, that particular component will remain nonnegative in all future updated qs. Each major cycle makes at least one more variable nonnegative. So there can be at most n major cycles when the method is applied to solve an LCP of order n. The first major cycle begins with $w = (w_j)$ as the initial complementary basic vector.

If q is nondegenerate, each component of the updated q remains nonzero throughout and there will never be any ties for the **blocking variable** (this term is defined

below) in each step of any major cycle, thus identifying the blocking variable uniquely and unambiuously in every step. If q is degenerate, there may be ties for the blocking variable. However, as discussed in Section 2.2.8, in this case q can be perturbed to become nondegenerate, treating the perturbation parameter to be positive and small without giving any specific value to it. This requires the use of the lexico minimum ratio test in place of the usual minimum ratio test, whenever it is used, right from the beginning, and this again guarantees that the blocking variable is identified uniquely and unambiguously in each step. If the method can be proved to process the LCP (q, M) in a finite number of steps when q is nondegenerate, using arguments similar to those in Section 2.2.8 it can be proved that it will process it in a finite number of steps even when q is degenerate, if this lexico minimum ratio test is used in place of the minimum ratio test in each step. Because of this, without any loss of generality, we assume that q is nondegenerate, in the description of the method given below.

Case 1: M is a P-Matrix.

The first major cycle begins with $w = (w_1, \ldots, w_n)$ as the initial complementary basic vector.

Let $y = (y_1, \ldots, y_n)$ where $y_j \in \{w_j, z_j\}$ for $j = 1$ to n, be the initial complementary basic vector at the beginning of a major cycle. For $j = 1$ to n, let t_j be the complement of y_j. Let the canonical tableau of (4.1) with respect to y be

basic vector	y	t	
y	I	$-\overline{M}$	$\overline{q}$

$t = 0$ in the current solution, $y = \overline{q}$ (4.14)

If $\overline{q} \geqq 0$, y is a complementary feasible basic vector for the LCP (q, M) and we terminate. Otherwise select an r such that $\overline{q}_r < 0$. y_r will be called the **distinguished variable** in this major cycle. We try to make y_r increase from its present negative value in the solution, to zero, without allowing any variable already nonnegative to become negative. For this, we increase t_r from zero to a λ say. This leads to the new solution

$$\begin{aligned} y_i &= \overline{q}_i + \lambda \overline{m}_{ir} \text{ , } i = 1 \text{ to } n \\ t_r &= \lambda \text{ , all other } t_j = 0. \end{aligned} \tag{4.15}$$

Since M is a P-matrix, by Theorem 3.5, $\overline{m}_{rr} > 0$. Hence, in (4.15), the value of y_r increases as λ increases. So, in this role, t_r is called the **driving variable**. The increase in the value of the driving variable must stop as soon as a positive basic variable decreases to zero, or the distinguished variable increases to zero. The variable which thus limits the increase of the driving variable is called the **blocking variable**. To identify the blocking variable, find minimum $\{(\overline{q}_r/(-\overline{m}_{rr})); (\overline{q}_i/(-\overline{m}_{ir}))$, for all i such that $\overline{q}_i \geqq 0$ and $(-\overline{m}_{ir}) > 0\}$. Suppose this minimum is attained by $i = s$ (if there is a tie for this s, the lexico minimum ratio rule as in Sections 2.2.7, 2.2.8 should be used to break the tie, as discussed above).

If $s = r$, a principal pivot step in position r is carried out in (4.14), this leads to a complementary basic solution in which y_r is positive, and the method moves to the next major cycle with it.

If $s \neq r$, perform a pivot in (4.14) replacing the basic variable y_s by t_r, a non-principal pivot. The new basic vector obtained is almost complementary (as defined in Section 2.4), both the distinguished variable y_s and its complement are basic variables in it, both the blocking variable y_r and its complement are nonbasic. Let $-\overline{\overline{m}}_{is}$, $i = 1$ to n, be the entries in the updated column of t_s after this pivot step. Clearly $-\overline{\overline{m}}_{ss} = -\overline{m}_{ss}/(-\overline{m}_{sr}) < 0$ since $\overline{m}_{ss} > 0$ (since M is a P-matrix) and $(-\overline{m}_{sr}) > 0$ (by the choice of the blocking variable), and $-\overline{\overline{m}}_{rs} = -\overline{m}_{rs} + \overline{m}_{rr}\overline{m}_{ss}/\overline{m}_{sr} < 0$ since $\overline{m}_{sr} < 0$ (by choice of the blocking variable) and $\overline{m}_{rr}\overline{m}_{ss} - \overline{m}_{sr}\overline{m}_{rs} > 0$ (this is the principal subdeterminant of $\overline{M}$ corresponding to the subset $\{s, r\}$ which is positive since $\overline{M}$ is a P-matrix, being a PPT of the P-matrix M). The pivot step has left the distinguished variable basic at a negative value. The next variable to enter the basis, that is, the next driving variable, is the complement of the blocking variable which just became nonbasic; it is t_s here. Since we have shown that $(-\overline{\overline{m}}_{ss}) < 0$, $(-\overline{\overline{m}}_{rs}) < 0$ above, increasing the value of the new driving variable results in the continuing increase of both the distinguished variable and its complement. The increase of the new driving variable is also governed by the same rules as above. Since the value of the distinguished variable has been shown to increase, it is potentially a blocking variable, and hence a blocking variable exists again. Using the properties of P-matrices discussed in Chapter 3, it can be verified that all these properties continue to hold when the major cycle is continued with the same rules. A sequence of almost complementary basic vectors is obtained in the process, which can only terminate when the distinguished variable is driven up to zero, at which time it is the blocking variable, and the corresponding pivot leads to a complementary basic vector. Since the distinguished variable and its complement increase strictly from one pivot step to the next, no basis can be repeated, and hence the sequence is finite, as there are only a finite number of almost complementary basic vectors. The finiteness of the overall method follows since there are at most n major cycles (the number of negative variables decreases by at least one in each major cycle).

In this case it can be verified that once the entry in a row in an updated q becomes nonnegative, it stays nonnegative in all subsequent steps.

Case 2: M is a PSD Matrix, but not a P-Matrix

In this case it is possible that the system $w - Mz = q$, $w, z \geqq 0$ is not even feasible, and the method should be able to detect this possibility. As before let (4.14) be the canonical tableau at the beginning of a major cycle. Select the distinguished variable as in Case 1 to be the basic variable in a row in which the updated right hand side constant is negative, say y_r. Since M is PSD, its PPT $\overline{M}$ is also PSD by Theorem 3.10 and hence its diagonal entries are all nonnegative by Result 1.5. So $\overline{m}_{rr} \geqq 0$, and could be zero here.

Suppose $\overline{m}_{rr} = 0$. In addition, if $(-\overline{m}_{ir}) \leqq 0$ for all i, Result 1.6 implies that $(-\overline{m}_{rj}) \geqq 0$ for all j (since $\overline{M}$ is PSD and $\overline{m}_{rr} = 0$ we will have $\overline{m}_{ir} + \overline{m}_{ri} = 0$ for all i). The equation corresponding to the updated rth row is

$$y_r + \sum_{j=1}^{n} (-\overline{m}_{rj}) t_j = \overline{q}_r \; . \tag{4.16}$$

Under these conditions ($\overline{q}_r < 0$, $-\overline{m}_{rj} \geqq 0$ for all j), (4.16) does not even have a nonnegative solution, which implies that "$w - Mz = q$, $w \geqq 0$, $z \geqq 0$" has no feasible solution. So under these conditions the LCP (q, M) has no solution.

If $\overline{m}_{rr} = 0$, and the infeasibility condition ($-\overline{m}_{ir} \leqq 0$ for all i) is not satisfied; as in Case 1, we increase the value of the driving variable t_r from zero. However, since $\overline{m}_{rr} = 0$, it has no effect on the negative value of the distinguished variable. In addition, if $-\overline{m}_{ir} \leqq 0$ for all i satisfying $\overline{q}_i \geqq 0$, the increase in the value of the driving variable t_r, makes no nonnegative basic variable decrease. But under these conditions $-\overline{m}_{ir} > 0$ for at lest one i satisfying $\overline{q}_i < 0$, and the value of this ith basic variable decreases further from its present negative value as the value of the driving variable is increased. So there is no blocking variable in the sense discussed under Case 1. Also, under these conditions, since there is at least one $\overline{m}_{ir} > 0$, we cannot make the infeasibility conclusion. Thus using the definitions of blocking as under Case 1, these conditions lead to an unblocked driving variable and yet no infeasibility conclusion is possible. In order to force the algorithm to move to a successful conclusion when this occurs, we make the following modifications in the definition of blocking (the aim is to make sure that the occurrence of an unblocked driving variable indicates the infeasibility of the original system "$w - Mz = q$, $w \geqq 0$, $z \geqq 0$" through an inconsistent equation of the form (4.16)). Let $\alpha <$ minimum $\{q_i : i = 1 \text{ to } n\}$. We impose a lower bound of α on all negative variables. A negative basic variable can then block the driving variable by decreasing to its lower bound α. When this happens, the blocking negative basic variable is replaced from the basic vector by the driving variable, and made into a nonbasic variable at its lower bound α. Once any variable attains a nonnegative value its lower bound is immediately changed to zero. With this modification, each nonbasic variable either has value 0 or α. A basic solution is nondegenerate if each basic variable has value different from 0 or α in the solution. Since nonbasic variables can have nonzero values, the basic values may not be equal to the updated right hand side constant vector $\overline{q}$, so we have to maintain the basic values separately in a column called $\overline{b}$.

At any stage of this method, if $\hat{q}$, $\hat{b}$, $-\hat{m}_{ij}$ denote the updated right hand side constants vector, updated basic values vector, and the updated entries in the nonbasic columns respectively, then $\hat{b}_i = \hat{q}_i + \Sigma(\alpha \hat{m}_{ij}$: over j such that the corresponding variable is nonbasic at its lower bound $\alpha)$. If at this stage the driving column (the updated column of the driving variable) is $(-\hat{m}_{1s}, \ldots, -\hat{m}_{ns})^T$, and the distinguished variable is the basic variable in the rth row, it can be shown that $\hat{m}_{rs} \geqq 0$ using the facts that the PPTs of a PSD matrix are PSD, and that the principal subdeterminants of a PSD matrix are $\geqq 0$ (similar to the proof of the corresponding statement that

$\overline{\overline{m}}_{rr} > 0$ under Case 1). Compute $\theta =$ minimum $\{(-\hat{b}_r/\hat{m}_{rs})$, if $\hat{m}_{rs} \neq 0$; $(-\hat{b}_i/\hat{m}_{is})$, for all i such that $-\hat{b}_i \geqq 0$ and $\hat{m}_{is} < 0$; $(\alpha - \hat{b}_i)/\hat{m}_{is}$, for all i such that $\hat{b}_i < 0$ and $\hat{m}_{is} < 0\}$. The blocking variable is the ith basic variable corresponding to the i that attains the minimum here. Ties for the blocking variable should be resolved using the lexico minimum ratio test in place of the usual minimum ratio test as described above. If a blocking variable exists, the pivot step replaces the blocking variable in the basic vector by the driving variable. In the new basic solution obtained after the pivot step the blocking variable that just left the basic vector is zero if it was the distinguished variable or a nonnegative basic variable, or α if it was a negative valued basic variable that decreased to its lower bound. The old driving variable which is now the new rth basic variable, has a value of θ in the basic solution. The new value of the ith basic variable is $\hat{b}_i + \theta\hat{m}_{is}$ for $i \neq r$. All other variables (nonbasics) continue to have the same value in the basic solution as before. If the distinguished variable is still basic, the procedure is continued by choosing the new driving variable to be the complement of the blocking variable that just dropped from the basic vector. As before, the procedure does not allow any nonnegative variable to become negative. It can be verified that each iteration of the method results in an increase (or lexico increase) of the sum of the distinguished variable and its complement. The major cycle terminates when the distinguished variable reaches the value zero and drops out of the basic vector, leading to a complementary basic vector.

To choose the distinguished variable at the beginning of a major cycle, we look for a basic variable, say the rth, whose value in the current basic solution, $\overline{b}_r < 0$ (even though the current updated $\overline{q}_r$ may be $\geqq 0$). However, in this case it is possible that no such basic variable exists. This happens when we reach a complementary basic vector with nonnegative values for all the basic variables in the current basic solution. If all the nonbasic variables are zero in this solution, the present complementary basic vector is feasible to the original LCP (q, M) and we terminate. On the other hand, if there are some nonbasic variables which are at their lower bound α in the current solution, check whether the current updated right hand side constants vector $\overline{q}$ is $\geqq 0$. If so, set all the nonbasic variables to zero, this changes the basic values to $\overline{q}$, and since $\overline{q} \geqq 0$, the present complementary basic vector is feasible to the original LCP (q, M) and we terminate. However, if $\overline{q} \ngeqq 0$ in such a situation, select one of the negative nonbasic variables (with value $= \alpha$ in the present basic solution) as the distinguished variable. In the first step of the ensuing major cycle, that nonbasic distinguished variable is itself the driving variable. If it is blocked, it becomes a basic variable after the first pivot step, and the major cycle continues until this distinguished variable increases to zero. However, a major cycle like this in which the nonbasic distinguished variable is the driving variable may consist of one step without any pivots if this driving variable can increase all the way from α to zero without making any nonnegative basic variable negative.

If we have a complementary basic vector in which the driving variable is unblocked, it cannot be the distinguished variable (since a distinguished driving variable must be

a negative nonbasic variable which will not be increase beyond zero). So an unblocked driving variable when the present basic vector is complementary must be the complement of a negative basic variable upon which its increase has no effect. Being unblocked, the updated column of the driving variable must be $\leqq 0$, and this implies infeasibility of the original LCP as discussed earlier.

The pivot element in any almost complementary basic vector is always positive by the rules under which the method is operated. The pivot element is only negative in this method when the dropping basic variable is the distinguished variable, which signals the end of a major cycle.

Suppose the driving variable is unblocked when the present basic vector is almost complementary. When this happens, the distinguished variable must be basic. Suppose it is the rth. Its complement must also be basic. Suppose it is the pth basic variable. Let the updated column of the driving variable be $(-\hat{m}_{1s}, \ldots, -\hat{m}_{ns})^T$. Since the distinguished variable is not blocking, we must have $\hat{m}_{rs} = 0$. Also we must have $-\hat{m}_{is} \leqq 0$, as otherwise some basic variable would block. It can be verified that in this case $-\hat{m}_{ps} < 0$. Pivoting with $-\hat{m}_{ps}$ as the pivot element restores complementarity and it can be verified that after this pivot step, it is possible to conlcude that the original LCP is infeasible.

4.4 References

4.1 R. H. Bartels, G. H. Golub and M. A. Saunders, "Numerical Techniques in Mathematical Programming", in: J. B. Rosen, O. L. Mangasarian and K. Ritter (editors), *Nonlinear Programming*, Academic Press, New York, 1970, 123–176.

4.2 P. T. Boggs, "The Solutions of Nonlinear Systems of Equations by A-Stable Integration Techniques", *SIAM Journal of Numerical Analysis*, 8 (1971) 767–785.

4.3 K. M. Brown, "A Quadratically Convergent Newton-Like Method Based Upon Gaussian Elimination", *SIAM Journal of Numerical Analysis*, 6 (1979) 560–569.

4.4 R. W. Cottle, "Nonlinear Programs With Positively Bounded Jacobians", *SIAM Journal of Applied Mathematics*, 14 (1966) 147–158.

4.5 R. W. Cottle, "The Principal Pivoting Method of Quadratic Programming", in: G. B. Dantzig and A. F. Veinott, Jr. (editors) *Mathematics of Decision Sciences*, AMS, Rhode Island, 1968, 144–162.

4.6 G. B. Dantzig and R. W. Cottle, "Positive (Semi-)Definite Programing", in: J. Abadie (editor), *Nonlinear Programming*, North-Holland, Amsterdam, The Netherlands, 1967, 55–73.

4.7 R. L. Graves, "A Principal Pivoting Simplex Algorithm for Linear and Quadratic Programming", *Operations Research*, 15 (1967) 482–494.

4.8 G. J. Habetler and M. M. Kostreva, "On a Direct Algorithm for Nonlinear Complementarity Problems", *SIAM Journal on Control and Optimization,* 16 (1978) 504–511.

4.9 R. Helgason, J. Kennington and H. Lall, "A Polynomially Bounded Algorithm for a Single Constrained Quadratic Program", *Mathematical Programming*, 18 (1980) 338–343.

4.10 E. L. Keller, "The General Quadratic Optimization Problem", *Mathematical Programming*, 5, (1973) 311–337.

4.11 M. M. Kostreva, "Direct Algorithms for Complementarity Problems", Ph. D. Thesis, Rensselaer Polytechnic Institute, Troy, New York, 1976.

4.12 J. Moré, "Classes of Functions and Feasibility Conditions in Nonlinear Complementarity Problems", *Mathematical Programming*, 6 (1974) 327–338.

4.13 J. Moré and W. Rheinboldt, "On P- and S-Functions and Related Classes of n-dimensional Nonlinear Mappings", *Linear Algebra and Its Applications*, 6 (1973) 45–68.

4.14 K. G. Murty, "Note on a Bard-Type Scheme for Solving the Complementarity Problem", *Opsearch*, 11 (1974) 123–130.

4.15 T. D. Parsons, "Applications of Principal Pivoting", in: H. W. Kuhn (editor), *Proceedings of the Princeton Symposium on Mathematical Programming*, Princeton University Press, Princeton, New Jersey, 1970, 567–581.

4.16 C. Van de Panne and A. Whinston, "The Symmetric Formulation of the Simplex Method for Quadratic Programming", *Econometrica*, 37 (1969) 507–527.

4.17 L. T. Watson, "An Algorithm for the Linear Complementarity Problem", *International Journal of Computer Mathematics*, 6 (1978) 319–325.

4.18 P. Wolfe, "The Simplex Method for Quadratic Programming", *Econometrica,* 27 (1959) 382–398.

Chapter 5

THE PARAMETRIC LINEAR COMPLEMENTARITY PROBLEM

Let M be a given square matrix of order n and let b, b^* be given column vectors in $\mathbf{R}^n$. Let $q(\lambda) = b + \lambda b^*$. Assuming that $b^* \neq 0$, $q(\lambda)$ traces a straight line in $\mathbf{R}^n$, $\mathbf{L} = \{x : x = q(\lambda), \text{ for some } \lambda\}$, as λ takes all real values. We consider the following parametric LCP: find w, z satisfying

$$\begin{aligned} w - Mz &= q(\lambda) = b + \lambda b^* \\ w &\geqq 0,\ z \geqq 0 \\ w^T z &= 0 \end{aligned} \tag{5.1}$$

as functions of λ, for each value of λ in some specified interval. Here we discuss an algorithm developed in [5.12] by K. G. Murty for obtaining a solution of this parametric LCP as a function of λ. This algorithm is most useful when M is a P-matrix. This algorithm solves the LCP $(q(\lambda), M)$ for some fixed value of λ by any method (such as the complementary pivot method, or the principal pivoting methods), and then obtains solutions for the parametric LCP for all values of λ using only a series of single principal pivot steps.

The Algorithm

Step 1: Choose a value λ_0, and fix λ at λ_0 (λ_0 could be equal to zero), and solve the LCP $(q(\lambda_0), M)$ by any one of the algorithms discussed earlier, and obtain a complementary feasible basic vector for it. With this complementary feasible basic vector for (5.1) when $\lambda = \lambda_0$, go to Step 2.

Step 2: Determine the range of values of λ for which the present complementary basic vector remains feasible. The procedure for doing this is the same as in parametric

right hand side LP, and it is as follows: Let $(y_1, \ldots, y_n)$, where $y_j \in \{w_j, z_j\}$ for each $j = 1$ to n, be the present complementary basic vector. Let $\beta = (\beta_{ij})$ be the inverse of the present complementary basis. Let $\overline{b}$, $\overline{b}^*$ be the present updated right hand side constants vectors, that is $\overline{b} = \beta b$, $\overline{b}^* = \beta b^*$. Compute $\underline{\lambda}$, $\overline{\lambda}$; the **lower and upper characteristic values associated with the present complementary basic vector**, from the following.

$$\begin{aligned}
\underline{\lambda} &= -\infty, \text{ if } \overline{b}_i^* \geqq 0 \text{ for all } i \\
&= \text{Maximum } \{-\overline{b}_i/\overline{b}_i^* : i \text{ such that } \overline{b}_i^* > 0\}, \text{ otherwise} \\
\overline{\lambda} &= +\infty, \text{ if } \overline{b}_i^* \geqq 0 \text{ for all } i \\
&= \text{Minimum } \{-\overline{b}_i/\overline{b}_i^* : i \text{ such that } \overline{b}_i^* < 0\}, \text{ otherwise}\,.
\end{aligned} \tag{5.2}$$

Since the present complementary basic vector is feasible for (5.1) for at least one value of λ, we will have $\underline{\lambda} \leqq \overline{\lambda}$, and for all values of λ in the closed interval $\underline{\lambda} \leqq \lambda \leqq \overline{\lambda}$, the present complementary basic vector remains feasible, and hence the solution

$$\begin{gathered}
\text{Present } i\text{th basic variable } y_i = \overline{b}_i + \lambda \overline{b}_i^*, i = 1 \text{ to } n \\
\text{Complement of } y_i, t_i = 0, i = 1 \text{ to } n
\end{gathered} \tag{5.3}$$

is a solution of the parametric LCP $(q(\lambda), M)$. Go to Step 3 or 4 if it is required to find the solutions of the parametric LCP $(q(\lambda), M)$ for values of $\lambda > \overline{\lambda}$, or for values of $\lambda < \underline{\lambda}$ respectively.

Step 3: We come to this step when we have a complementary basic vector, $y = (y_1, \ldots, y_n)$ say, for which the upper characteristic value is $\overline{\lambda}$, and it is required to find solutions of the parametric LCP $(q(\lambda), M)$, for values of $\lambda > \overline{\lambda}$. Let $\overline{b}$, $\overline{b}^*$ be the present updated right hand side constant vectors. Find out $\mathbf{J} = \{i : i \text{ ties for the minimum in (5.2) for determining } \overline{\lambda}\}$, $r = \text{maximum } \{i : i \in \mathbf{J}\}$. So $\overline{b}_r^* < 0$ and $-\overline{b}_r/\overline{b}_r^* = \overline{\lambda}$. The value of the rth basic variable y_r, in the solution in (5.3) is zero when $\lambda = \overline{\lambda}$, and it becomes negative when $\lambda > \overline{\lambda}$. Let t_r be the complement of y_r and let $\overline{A}_{.r} = (\overline{a}_{1r}, \ldots, \overline{a}_{nr})^T$ be its updated column vector. If $\overline{a}_{rr} < 0$, perform a single principal pivot step in position r in y leading to the complementary basic vector $u = (y_1, \ldots, y_{r-1}, t_r, y_{r+1}, \ldots, y_n)$. Both y and u have the same BFS when $\lambda = \overline{\lambda}$ (since $y_r = 0$ when $\lambda = \overline{\lambda}$ in the solution in (5.3)). u is a complementary feasible basic vector for (5.1) when $\lambda = \overline{\lambda}$. The value of t_r in the basic solution of (5.1) with respect to u is $(\overline{b}_r/\overline{a}_{rr}) + \lambda(\overline{b}^*/\overline{a}_{rr})$, this quantity is 0 when $\lambda = \overline{\lambda}$, and since $\overline{a}_{rr} < 0$, $\overline{b}_r^* < 0$, we verify that this quantity is positive when $\lambda > \overline{\lambda}$. From this it can be verified that the lower characteristic value for u is $\overline{\lambda}$ = upper characteristic value for y. With u, go back to Step 2.

If $\overline{a}_{rr} \geqq 0$, either the single principal pivot step in position r cannot be carried out (when $\overline{a}_{rr} = 0$); or even after it is carried out, the new rth basic variable continues to be negative when $\lambda > \overline{\lambda}$ in the new basic solution (which happens when $\overline{a}_{rr} > 0$). Thus in this case, the algorithm is unable to solve the parametric LCP $(q, (\lambda), M)$ for $\lambda > \overline{\lambda}$,

it is even unable to determine whether there exists a solution to the LCP $(q, (\lambda), M)$ or not when $\lambda > \overline{\lambda}$.

Step 4: We come to this step when we have a complementary basic vector, $y = (y_1, \ldots, y_n)$ say, for which the lower characteristic value is $\underline{\lambda}$, and it is required to find solutions of the parametric LCP $(q, (\lambda), M)$ for values of $\lambda < \underline{\lambda}$. Let $\mathbf{J} = \{i : i \text{ ties for the maximum in (5.2) for determining } \underline{\lambda}\}$, $r =$ maximum $\{i : i \in \mathbf{J}\}$. Let t_r be the complement of y_r and let $\overline{A}_{.r} = (\overline{a}_{1r}, \ldots, \overline{a}_{nr})^T$ be its updated column vector. If $\overline{a}_{rr} < 0$, perform a single principal pivot step in position r in y. This leads to the next complementary feasible basic vector for which $\underline{\lambda}$ is the upper characteristic value, continue with it in the same way. If $\overline{a}_{rr} \geqq 0$, this algorithm is unable to solve, or even determine whether a solution exists for the parametric LCP $(q, (\lambda), M)$ when $\lambda < \underline{\lambda}$.

Example 5.1

Consider the parametric LCP $(q, (\lambda) = b + \lambda b^*, M)$, for which the original tableau is

w_1	w_2	w_3	z_1	z_2	z_3	b	b^*
1	0	0	−1	0	0	8	−1
0	1	0	−2	−1	0	4	−1
0	0	1	−2	−2	−1	2	−1

When $\lambda = 0$, (w_1, w_2, w_3) is a complementary feasible basic vector for this problem. The inverse tableau corresponding to this is:

First Inverse Tableau

Basic Variable	Inverse of the Complementary Basis			$\overline{b}$	$\overline{b}^*$	$-\overline{b}_i/\overline{b}^*_i$ for i such that $\overline{b}^*_i < 0$	$-\overline{b}_i/\overline{b}^*_i$ for i such that $\overline{b}^*_i > 0$	Range of Feasibility	Pivot Column z_3
w_1	1	0	0	8	−1	8		$-\infty < \lambda \leqq 2$	0
w_2	0	1	0	4	−1	4			0
w_3	0	0	1	2	−1	2			$\boxed{-1}$

So in the range $-\infty < \lambda \leqq 2$, $(w = (8 - \lambda, 4 - \lambda, 2 - \lambda)^T, z = 0)$ is a solution of this parametric LCP. To find out solutions of this parametric LCP when $\lambda > 2$, we have to make a single principal pivot step in position 3.

The updated column vector of z_3 is $\overline{A}_{.3} = (0, 0, -1)$. $\overline{a}_{33} = -1$, and hence we can continue. The pivot column is already entered by the side of the first inverse tableau. Performing the pivot leads to the next inverse tableau.

Second Inverse Tableau

Basic Variable	Inverse of the Complementary Basis			$\overline{b}$	$\overline{b}^*$	$-\overline{b}_i/\overline{b}_i^*$ for i such that $\overline{b}_i^* < 0$	$-\overline{b}_i/\overline{b}_i^*$ for i such that $\overline{b}_i^* > 0$	Range of Feasibility
w_1	1	0	0	8	−1	8		$2 < \lambda \leqq 4$
w_2	0	1	0	4	−1	4		
z_3	0	0	−1	−2	1		2	
						$\overline{\lambda} = 4$	$= \underline{\lambda} = 2$	

So in the range $2 \leqq \lambda \leqq 4$, the solution $(w_1, w_2, z_3) = (8 - \lambda, 4 - \lambda, -2 + \lambda)$, $(z_1, z_2, w_3) = (0, 0, 0)$ is a solution of this parametric LCP $(q(\lambda), M)$. Continuing in the same way, we get the following solutions for this problem summarized in the table below.

Optimality Range	Complementary Feasible Basic Vector	Complementary Solution $(w^T; z^T)$
$-\infty < \lambda < 2$	(w_1, w_2, w_3)	$(8 - \lambda, 4 - \lambda, 2 - \lambda; 0, 0, 0)$
$2 \leqq \lambda \leqq 4$	(w_1, w_2, z_3)	$(8 - \lambda, 4 - \lambda, 0; 0, 0, -2 + \lambda)$
$4 \leqq \lambda \leqq 6$	(w_1, z_2, z_3)	$(8 - \lambda, 0, 0; 0, -4 + \lambda, 6 - \lambda)$
$6 \leqq \lambda \leqq 8$	(w_1, z_2, w_3)	$(8 - \lambda, 0, -6 + \lambda; 0, -4 + \lambda, 0)$
$8 \leqq \lambda \leqq 10$	(z_1, z_2, w_3)	$(0, 0, -6 + \lambda; -8 + \lambda, -4 + \lambda, 0)$
$10 \leqq \lambda \leqq 12$	(z_1, z_2, z_3)	$(0, 0, 0; -8 + \lambda, 12 - \lambda, -10 + \lambda)$
$12 \leqq \lambda \leqq 14$	(z_1, w_2, z_3)	$(0, -12 + \lambda, 0; -8 + \lambda, 0, 14 - \lambda)$
$14 \leqq \lambda$	(z_1, w_2, w_3)	$(0, -12 + \lambda, -14 + \lambda; -8 + \lambda, 0, 0)$

Example 5.2

Consider the parametric LCP $(q(\lambda) = b + \lambda b^*, M)$ for which the original tableau is:

w_1	w_2	w_3	w_4	z_1	z_2	z_3	z_4	b	b^*
1	0	0	0	−1	1	1	1	3	−2
0	1	0	0	1	−1	1	1	5	−4
0	0	1	0	−1	−1	−2	0	−9	5
0	0	0	1	−1	−1	0	−2	−5	3

Putting $\lambda = 0$, we verify that this LCP is the same as the one solved in Example 2.8. The complementary feasible basic vector obtained for this problem (when $\lambda = 0$) in Example 2.8 is (z_1, z_2, z_3, z_4). The inverse tableau corresponding to (z_1, z_2, z_3, z_4) is

Basic Variable	Inverse of the Complementary Basis	$\overline{b}$	$\overline{b}^*$	$-\overline{b}_i/\overline{b}_i^*$ for i such that $\overline{b}_i^* < 0$	$-\overline{b}_i/\overline{b}_i^*$ for i such that $\overline{b}_i^* > 0$	Range of Feasibility
z_1	−1/2 0 −1/4 −1/4	2	−1	2		
z_2	0 −1/2 −1/4 −1/4	1	0			$\lambda \leqq 1$
z_3	1/4 1/4 −1/4 1/4	3	−2	3/2		
z_4	1/4 1/4 1/4 1/4	1	−1	1		
				Minimum $= \overline{\lambda} = 1$	Maximum $= \underline{\lambda} = -\infty$	

So when $\lambda \leqq 1$, the solution $(w = (w_1, w_2, w_3, w_4) = 0,\ z = (z_1, z_2, z_3, z_4) = (2 - \lambda, 1, 3 - 2\lambda, 1 - \lambda))$ is a solution of this parametric LCP $(q(\lambda), M)$. To look for solutions when $\lambda > 1$, we have to make a single principal pivot step in position 4. The updated column vector of w_4 is $\overline{A}_{.4} = (-1/4, -1/4, 1/4, 1/4)$. So $\overline{a}_{44} = 1/4 > 0$. Since $\overline{a}_{44}$ is strictly positive, the algorithm discussed above is unable to process this parametric LCP $(q(\lambda), M)$ when $\lambda > 1$.

Theorem 5.1 *Let M be a given P-matrix of order n. Consider the parametric LCP $(q(\lambda), M)$. The algorithm discussed above finds solutions of this parametric LCP for all real values of λ in a finite number of pivot steps. Also, for each λ, the solution obtained is the unique solution of this parametric LCP for that value of λ.*

Proof. In the notation of the algorithm, let $y = (y_1, \ldots, y_n)$ be the complementary basic vector in Step 2 at some stage of the algorithm, for which the range of feasibility

is $\underline{\lambda} \leqq \lambda \leqq \overline{\lambda}$. In order to find out solutions for $\lambda > \overline{\lambda}$, suppose we have to make a single principal pivot step in position r. Let t_r be the complement of y_r and let $\overline{A}_{.r}$ be the updated column vector of t_r. By Corollary 3.5 the method can continue. Let the BFS with respect to y be the one given in (5.3). Let $\mathbf{T} = \{i : i \text{ such that } \overline{b}_i + \overline{\lambda}\overline{b}_i^* = 0\}$. Clearly $r \in \mathbf{T}$ and in fact the set $\mathbf{J}$ defined in Step 3 at this stage satisfies $\mathbf{J} \subset \mathbf{T}$. As long as λ remains fixed at $\overline{\lambda}$, any principal pivot steps performed on positions in $\mathbf{T}$ will not change the basic solution (because when the basic variable in the pivot row is 0 in the basic solution, the pivot step is a degenerate pivot step that leaves the basic solution unchanged). Let $u = (u_1, \ldots, u_n)$ be any complementary basic vector satisfying the property that $u_i = y_i$ for $i \notin \mathbf{T}$, $u_i = y_i$ or its complement for $i \in \mathbf{T}$. Suppose the updated right hand side constant vectors with respect to u are $\hat{b}$, $\hat{b}^*$. By the above argument, the basic solution of (5.1) with respect to u at $\lambda = \overline{\lambda}$ is

$$u_i = \hat{b}_i + \overline{\lambda}\hat{b}_i^* = \overline{b}_i + \overline{\lambda}\,\overline{b}_i^*, \;\; i = 1 \text{ to } n$$
$$(\text{Complement of } u_i) = 0, \;\; i = 1 \text{ to } n.$$

So $\hat{b}_i + \overline{\lambda}\hat{b}_i = 0$ for $i \in \mathbf{T}$ and > 0 for $i \notin \mathbf{T}$. So the upper characteristic point associated with u is $> \overline{\lambda}$ iff $\hat{b}_i^* \geqq 0$ for all $i \in \mathbf{T}$. Thus, if $\mathbf{T}$ is a singleton set, the pivot step carried out in Step 3 at this stage is guaranteed to produce a complementary feasible basic vector for which the upper characteristic value is $> \overline{\lambda}$. If $\mathbf{T}$ has 2 or more elements, let $\omega = (w_i, i \in \mathbf{T})$, $\xi = (z_i, i \in \mathbf{T})$, $\mathcal{M}$ the principal submatrix of M corresponding to the subset $\mathbf{T}$, and $\gamma = (\overline{b}_i^*, i \in \mathbf{T})$. Consider the LCP $(\gamma, \mathcal{M})$ in the variables (ω, ξ). Since $\mathcal{M}$ is a P-matrix, by Theorem 4.1, the LCP $(\gamma, \mathcal{M})$ can be solved by Principal Pivoting Method I in a finite number of pivot steps without cycling, starting with the complementary basic vector $(y_i, i \in \mathbf{T})$ until a complementary basic vector is obtained for it, with respect to which the updated γ is $\geqq 0$. The choice of the pivot row r in Step 3 of the parametric algorithm implies that when it is continued from the canonical tableau of (5.1) with respect to y, keeping $\lambda = \overline{\lambda}$, it will go through exactly the same sequence of pivotal exchanges as in the LCP $(\gamma, \mathcal{M})$, when it is solved by Principal Pivoting Method I, until we obtain a complementary feasible basic vector, $u = (u_1, \ldots, u_n)$ say, satisfying the property that the updated b_i^* with respect to u is $\geqq 0$ for each $i \in \mathbf{T}$. By the above argument the upper characteristic value of u is $> \overline{\lambda}$, and hence when we reach the basic vector u, we are able to strictly increase the value of λ beyond $\overline{\lambda}$. Also, once we cross the interval of feasibility of a complementary basic vector in this parametric algorithm, we will never encounter this basic vector again. We can apply the same argument in Step 4 for decreasing λ below $\underline{\lambda}$. Continuing in this way, since there are only $2^{\mathbf{n}}$ complementary basic vectors, these arguments imply that after at most a finite number (less than $2^{\mathbf{n}}$) of pivot steps, we will obtain solutions of the parametric LCP $(q(\lambda), M)$ for all λ.

The fact that the solution obtained is the unique solution for each λ, follows from Theorem 3.13.

□

When there are ties for the i that attains the minimum in (5.2) of Step 3 and the pivot row is chosen among $i \in \mathbf{J}$ arbitrarily (instead of choosing it as the bottommost

as mentioned in Step 3), cycling can occur at this value of $\lambda = \overline{\lambda}$, as shown in the following example due to A. Gana [5.6]. He considers the parametric LCP with the following data

$$M = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 2 \\ 2 & 0 & 1 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad b^* = \begin{pmatrix} -1 \\ -1 \\ -1 \end{pmatrix}.$$

Starting with the complementary feasible basic vector (w_1, w_2, w_3) when $\lambda = 0$, we want to solve this problem for all $\lambda \geqq 0$. Here is a sequence of complementary basic vectors obtained when the pivot row in Step 3 is chosen among $i \in \mathbf{J}$ arbitrarily. Pivot elements are in a box.

Basic Variables	w_1	w_2	w_3	z_1	z_2	z_3	b	b^*	Feasibility Interval
w_1	1	0	0	$\boxed{-1}$	-2	0	1	-1	
w_2	0	1	0	0	-1	-2	1	-1	$0 \leqq \lambda \leqq 1$
w_3	0	0	1	-2	0	-1	1	-1	
z_1	-1	0	0	1	2	0	-1	1	
w_2	0	1	0	0	$\boxed{-1}$	-2	1	-1	$1 \leqq \lambda \leqq 1$
w_3	-2	0	1	0	4	-1	-1	1	
z_1	$\boxed{-1}$	2	0	1	0	-4	1	-1	
z_2	0	-1	0	0	1	2	-1	1	$1 \leqq \lambda \leqq 1$
w_3	-2	4	1	0	0	-9	3	-3	
w_1	1	-2	0	-1	0	4	-1	1	
z_2	0	-1	0	0	1	2	-1	1	$1 \leqq \lambda \leqq 1$
w_3	0	0	1	-2	0	$\boxed{-1}$	1	-1	
w_1	1	-2	4	-9	0	0	3	-3	
z_2	0	$\boxed{-1}$	2	-4	1	0	1	-1	$1 \leqq \lambda \leqq 1$
z_3	0	0	-1	2	0	1	-1	1	
w_1	1	0	0	$\boxed{-1}$	-2	0	1	-1	
w_2	0	1	-2	4	-1	0	-1	1	$1 \leqq \lambda \leqq 1$
z_3	0	0	-1	2	0	1	-1	1	

Basic Variables	w_1	w_2	w_3	z_1	z_2	z_3	b	b^*	Feasibility Interval
z_1	-1	0	0	1	2	0	-1	1	
w_2	4	1	-2	0	-9	0	3	-3	$1 \leqq \lambda \leqq 1$
z_3	2	0	$\boxed{-1}$	0	4	1	1	-1	
z_1	-1	0	0	1	2	0	-1	1	
w_2	0	1	0	0	-1	-2	1	-1	$1 \leqq \lambda \leqq 1$
w_3	-2	0	1	0	4	-1	-1	1	

The complementary basic vector (z_1, w_2, w_3) repeated at $\lambda = 1$, and hence cycling has occurred, and the execution can go through this cycle repeatedly without ever being able to increase λ beyond 1. Theorem 5.1 indicates that if the pivot row is chosen as mentioned in Steps 3, 4 of the parametric algorithm, this cycling cannot occur.

Geometric Interpretation

Let M be a given square matrix of order n. Consider the parametric LCP $(q(\lambda) = b+\lambda b^*, M)$. In the process of solving this problem by the parametric LCP algorithm discussed above, let $y = (y_1, \ldots, y_n)$, where $y_j \in \{w_j, z_j\}$ for each $j = 1$ to n, be a complementary basic vector obtained in some stage. Let $D_{.j}$ be the column vector associated with y_j in (5.1) for $j = 1$ to n. Let $[\underline{\lambda}, \overline{\lambda}]$ be the interval of feasibility of y. To find solutions for the parametric LCP $(q(\lambda), M)$ when $\lambda > \overline{\lambda}$, suppose we have to make a principal pivot step in position r. Let t_r be the complement of y_r and let $A_{.r}$ be the column associated with t_r in (5.1). So $A_{.r}$ is the complement of $D_{.r}$. Since the value of y_r in the solution in 5.1 is zero when $\lambda = \overline{\lambda}$, we have $q(\overline{\lambda}) \in \text{Pos}\{D_{.1}, \ldots, D_{.r-1}, D_{.r+1}, \ldots, D_{.n}\}$. Thus the portion of the straight line $\mathbf{L}$ in (5.1) corresponding to $\underline{\lambda} \leqq \lambda \leqq \overline{\lambda}$ lies in the complementary cone $\mathbf{K}_1 = \text{Pos}\{D_{.1}, \ldots, D_{.n}\}$, and as λ increases through $\overline{\lambda}$, it leaves the cone $\mathbf{K}_1$ through its facet $\mathbf{F} = \text{Pos}\{D_{.1}, \ldots, D_{.r-1}, D_{.r+1}, \ldots, D_{.n}\}$. Let $\mathbf{H}$ denote the hyperplane in $\mathbf{R}^n$ which is the linear hull of $\{D_{.1}, \ldots, D_{.r-1}, D_{.r+1}, \ldots, D_{.n}\}$. Let $\overline{A}_{.r} = (\overline{a}_{1r}, \ldots, \overline{a}_{nr})^T$ be the updated column associated with t_r. By Theorem 3.16, the hyperplane $\mathbf{H}$ strictly separates $D_{.r}$ and $A_{.r}$ iff $\overline{a}_{rr} < 0$. If $\overline{a}_{rr} < 0$, $q(\overline{\lambda})$ is on the common facet $\mathbf{F}$ of the complementary cones $\mathbf{K}_1$ and $\mathbf{K}_2 = \text{Pos}\{D_{.1}, \ldots, D_{.r-1}, A_{.r}, D_{.r+1}, \ldots, D_{.n}\}$. See Figure 5.1. As λ increases beyond $\overline{\lambda}$, the line $\mathbf{L}$ leaves the complementary cone $\mathbf{K}_1$ and enters the complementary cone $\mathbf{K}_2$ through their common facet $\mathbf{F}$.

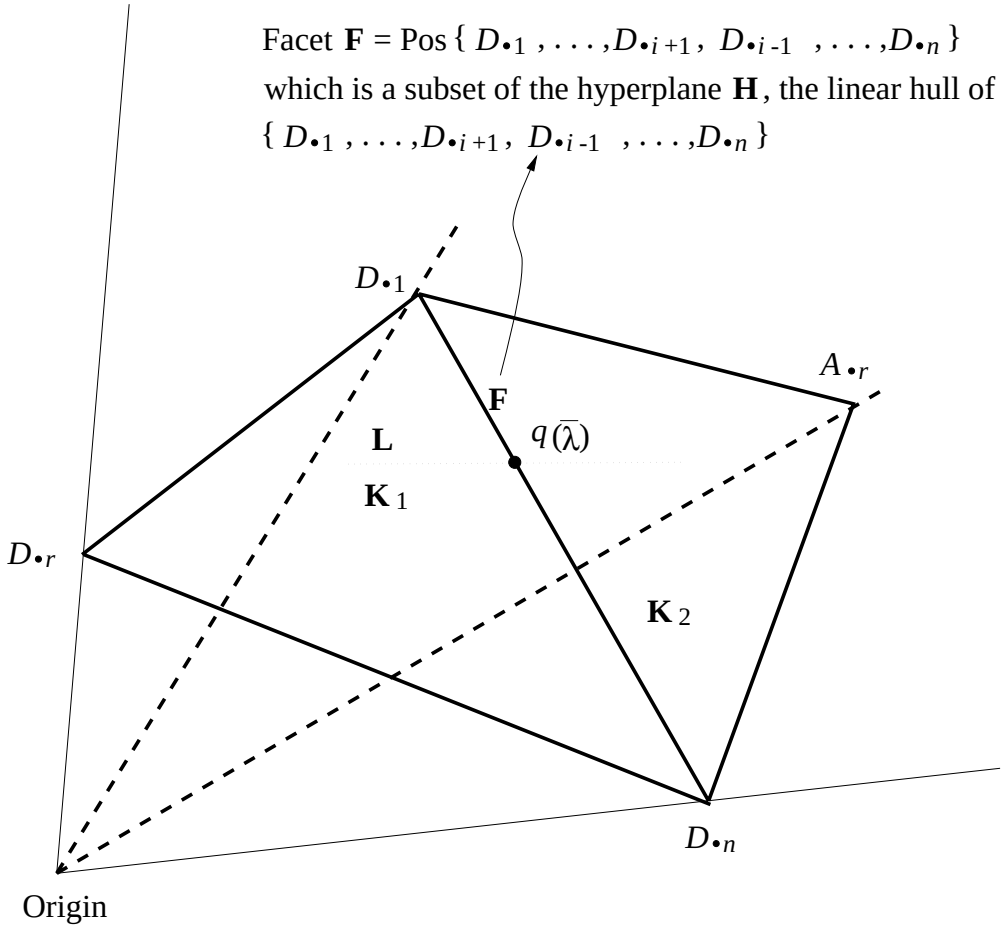


Figure 5.1 Situation when $\overline{a}_{rr} < 0$. As λ increases through $\overline{\lambda}$, the point $q(\lambda)$ travels along the straight line $\mathbf{L}$, leaves the complementary cone $\mathbf{K}_1$ and enters the complementary cone $\mathbf{K}_2$, through their common facet $\mathbf{F}$.

If M is a P-matrix, by the strict separation property discussed in Section 3.3, this situation occurs whenever Step 3 or 4 is carried out in the parametric LCP algorithm, and the algorithm finds the solutions of the parametric LCP for all values of the parameter λ.

If $\overline{a}_{rr} = 0$, $A_{.r}$ lies on the hyperplane $\mathbf{H}$ itself. If $\overline{a}_{rr} > 0$, $A_{.r}$ lies on the same side of the hyperplane $\mathbf{H}$ as $D_{.r}$. In either of these cases, as λ increases through $\overline{\lambda}$, the line $\mathbf{L}$ leaves both the complementary cones $\mathbf{K}_1$ and $\mathbf{K}_2$ and $(y_1, \ldots, y_{r-1}, t_r, y_{r+1}, \ldots, y_n)$ is not a complementary feasible basic vector for the parametric LCP $(q(\lambda), M)$ when $\lambda > \overline{\lambda}$. Hence if $\overline{a}_{rr} \geqq 0$, the parametric LCP algorithm is unable to find solutions of the parametric LCP $(q(\lambda), M)$ when λ increases beyond $\overline{\lambda}$.

Hence, geometrically, the parametric LCP algorithm discussed above can be interpreted as a walk along the straight line $\mathbf{L}$ crossing from one complementary cone in $\mathcal{C}(M)$ to an adjacent complement cone through their common facet.

Exercises

5.1 Let M be a given PSD matrix of order n, which is not PD. Discuss an approach for solving the parametric LCP $(q(\lambda) = b + \lambda b^*, M)$ for all values of λ for which it has a solution, and determining the range of values of λ for which it has no solution, based on the Graves' principal pivoting method of Section 4.2.

5.2 Suppose M is a copositive plus matrix and not a P-matrix. Discuss an approach for processing the parametric LCP (5.1) in this case, by the algorithm discussed above, using the complementary pivot algorithm to extend the value of λ whenever the pivot element in the parametric algorithm turns out to be nonnegative. Also prove that in this case, the set of all values of λ for which the parametric LCP (5.1) has a solution, is an interval.

5.1 PARAMETRIC CONVEX QUADRATIC PROGRAMMING

Here we consider a problem of the following form:

$$\begin{array}{lrl} \text{minimize} & Q_\lambda(x) = & (c + \lambda c^*)x + \frac{1}{2}x^T D x \\ \text{subject to} & Ax & \geqq b + \lambda b^* \\ & x & \geqq 0 \end{array} \tag{5.4}$$

where D is a symmetric PSD matrix of order n, and λ is a real valued parameter. The parameter λ in the right hand side constants vector in the constraints, and the linear part of the objective function, is the same. If $b^* = 0$, or $c^* = 0$, we get the special case of the problem in which the parameter appears in only the right hand side constants vector, or the linear part of the objective function, respectively. It is required to find an optimum solution of this problem, treating λ as a parameter, for all values of λ.

By the results in Chapter 1, this problem is equivalent to a parametric LCP $(q + \lambda q^*, M)$ where M is a PSD matrix. For the problem above, the data in the parametric LCP is given by

$$M = \begin{pmatrix} D & -A^T \\ A & 0 \end{pmatrix}, \quad q = \begin{pmatrix} c^T \\ -b \end{pmatrix} \quad q^* = \begin{pmatrix} (c^*)^T \\ -b^* \end{pmatrix}. \tag{5.5}$$

We now discuss an algorithm for solving problems of this type. In preparing this section, I benefitted a lot from discussions with R. Saigal.

Algorithm for Parametric LCP $(q + \lambda q^*, M)$ When M is PSD

Initialization

Find a value for the parameter λ for which the system

$$\begin{aligned} w - Mz &= q + \lambda q^* \\ w,\ z &\geqq 0 \end{aligned} \tag{5.6}$$

has a feasible solution. Since M is PSD, by the results in Chapter 2, the LCP $(q + \lambda q^*, M)$ has a solution iff (5.6) has a feasible solution for that λ. Phase I of the parametric right hand side simplex method can be used to find a feasible solution for (5.6) (see Section 8.8 of [2.26]). When M, q, q^* are given by (5.5), if (5.6) is infeasible for a value of λ, by the results in Chapter 2, (5.4) does not have an optium solution for that λ (it is either infeasible, or $Q_\lambda(x)$ is unbounded below on the set of feasible solutions for it).

If there exists no value for λ for which (5.6) has a feasible solution, the parametric LCP $(q + \lambda q^*, M)$ does not have a solution for any λ, terminate. Otherwise, let λ_0 be a value of λ, for which (5.6) has a feasible solution (the parametric right hand side simplex algorithm, see Section 8.6 of [2.26], can in fact be used to determine the interval of values of λ for which (5.6) is feasible).

Now, find a complementary feasible basis for the LCP $(q + \lambda_0 q^*, M)$ with λ fixed equal to λ_0. The complementary pivot algorithm of Section 2.2 can be used for finding this. Since (5.6) is feasible when $\lambda = \lambda_0$ and M is PSD, by the results in Chapter 2, the complementary pivot algorithm applied on the LCP $(q + \lambda_0 q^*, M)$ will terminate with a complementary feasible basic vector for it, in a finite number of pivot steps, if the lexicographic minimum ratio rule is used to determine the dropping variable in each step. Let the complementary feasible basic vector be $y = (y_1, \ldots, y_n)$, (where $y_j \in \{w_j, z_j\}$ for each $j = 1$ to n), associated with the complementary basis B. Let $\overline{q}$, $\overline{q}^*$ be the updated right hand side constants vectors ($\overline{q} = B^{-1}q, \overline{q}^* = B^{-1}q^*$). Let

$$\begin{aligned} \lambda_1 &= -\infty, \text{ if } \overline{q}^* \leqq 0 \\ &= \text{ Maximum } \{-\overline{q}_i/\overline{q}_i^* : i \text{ such that } \overline{q}_i^* > 0\}, \text{ otherwise,} \\ \lambda_2 &= +\infty, \text{ if } \overline{q}^* \geqq 0 \\ &= \text{ Minimum } \{-\overline{q}_i/\overline{q}_i^* : i \text{ such that } \overline{q}_i^* < 0\}, \text{ otherwise.} \end{aligned}$$

Then, for all $\lambda_1 \leqq \lambda \leqq \lambda_2$, y is a complementary feasible basic vector for the parametric LCP $(q+\lambda q^*, M)$. This interval is nonempty since λ_0 is contained in it. In this interval, a complementary feasible solution for the parametric LCP is

$$\text{complement of } y \text{ is } 0$$

$$y = \overline{q} + \lambda \overline{q}^*.$$

Procedure to Increase the Value of λ

Suppose we have a complementary basic vector $y = (y_1, \ldots, y_n)$, where $y_j \in \{w_j, z_j\}$ for each $j = 1$ to n, corresponding to the complementary basis B, for which the upper characteristic value is $\overline{\lambda}$, which is finite. Here we discuss how to proceed to find complementary solutions of the parametric LCP when $\lambda > \overline{\lambda}$. Assume that B is lexico feasible for $\lambda = \overline{\lambda}$. Let $\beta = B^{-1}$, $\overline{q} = B^{-1}q$, $\overline{q}^* = B^{-1}q^*$. Then $(\overline{q}_i + \overline{\lambda}\overline{q}_i^*, \beta_i.) \succ 0$ for all $i = 1$ to n. Determine the i which attains the lexico minimum $\{-(\overline{q}_i, \beta_i.)/(\overline{q}_i^*) : i$ such that $\overline{q}_i^* < 0\}$, and suppose it is p. Let the complement of the variable y_p be t_p. Suppose the updated column vector of t_p in the canonical tableau for

w	z	
I	$-M$	$q + \lambda q^*$

(5.7)

with respect to the complementary basic vector y be $(\overline{a}_{1p}, \ldots, \overline{a}_{np})^T$. Since M is PSD, by the results in Chapter 3, $\overline{a}_{pp} \geqq 0$.

If $\overline{a}_{pp} < 0$, performing a single principal pivot step in position p in the present complementary basic vector y, leads to a new complementary basic vector which will be feasible for some values of $\lambda > \overline{\lambda}$ under nondegeneracy. We repeat this whole process with that complementary basic vector.

If $\overline{a}_{pp} = 0$, to increase the value of λ beyond $\overline{\lambda}$, we enter into a special complementary pivot phase described below.

The Complementary Pivot Phase to Increase the Value of λ

We enter this phase when we obtain a complementary basic vector $y = (y_1, \ldots, y_n)$, where $y_j \in \{w_j, z_j\}$ for each $j = 1$ to n, with finite upper characteristic value $\overline{\lambda}$, and $\overline{a}_{pp} = 0$, as discussed above.

In the present canonical tableau, transfer the column of the parameter λ from the right hand side to the left hand side, and treat λ now as a variable. This leads to

Basic Variables	y	t	λ	
y	I	$\ldots$	$-\overline{q}^*$	$\overline{q}$

(5.8)

In this tableau, perform a pivot step in the column of λ, with row p as the pivot row, this is possible since $-\overline{q}_p^* > 0$. This leads to the following tableau.

Tableau 5.1

Basic Variables	$y_1 \ldots y_{p-1}$ λ $y_{p+1} \ldots y_n$	y_p $t_1 \ldots t_n$	
y_1			
$\vdots$			
y_{p-1}	I	$\ldots$	$\tilde{q}$
λ			
y_{p+1}			
$\vdots$			
y_n			

Tableau 5.1 is the canonical tableau with respect to the basic vector $(y_1, \ldots, y_{p-1}, \lambda, y_{p+1}, \ldots, y_n)$. As defined in Chapter 2, this is an ACBV (almost complementary basic vector). Here, λ plays the same role as the artificial variable z_0 in Chapter 2. There is one difference. In Chapter 2, z_0 was a nonnegative artificial variable, here λ is a variable which is a natural parameter, and it can take either negative or positive values.

From the manner in which Tableau 5.1 is obtained, it is clear that the value of λ in the basic solution corresponding to Tableau 5.1 is $\tilde{q}_p = -\overline{q}_p/\overline{q}_p^* = \overline{\lambda}$. Treat Tableau 5.1 as the original tableau for this phase. The word **basis** in this phase refers to the matrix of columns from Tableau 5.1, corresponding to the basic variables in any basic vector for Tableau 5.1. This phase requires moving among ACBVs in which λ will always be the pth basic variable. Let B be the basis corresponding to such an ACBV, and let $\hat{q} = B^{-1}\tilde{q}$, $\beta = B^{-1}$. This ACBV is said to be **feasible for this phase** if $\hat{q}_i \geqq 0$ for all $i \neq p$ and **lexico feasible for this phase** if $(\hat{q}_i, \beta_{i.}) \succ 0$ for all $i \neq p$. Let B be such a basis, let $\hat{q} = B^{-1}\tilde{q}$, $\beta = B^{-1}$ and suppose it is required to bring the column of a nonbasic variable, say x_s, into the basis B. Let $(\hat{a}_{1s}, \ldots, \hat{a}_{ns})^T$ be the updated column of x_s (it is, B^{-1} (column of x_s in Tableau 5.1)). The **lexico minimum ratio test for this phase** determines the dropping variable to be the rth basic variable, where r is the i which attains the lexico minimum $\{(\hat{q}_i, \beta_{i.})/\hat{a}_{is} : i$ such that $i \in \{1, \ldots, p-1, p+1, \ldots, n\}$ and $\hat{a}_{is} > 0\}$. The **minimum ratio** for this pivot step, is defined to be $(\hat{q}_r/\hat{a}_{rs})$, it is always $\geqq 0$. The initial ACBV in Tableau 5.1 is lexico feasible in the sense defined here, and all the ACBVs obtained during this phase will have the same property.

Now, bring the variable t_p into the initial ACBV $(y_1, \ldots, y_{p-1}, \lambda, y_{p+1}, \ldots, y_n)$, determining the dropping variable by the lexico minimum ratio test as discussed above. Continue this phase using the **complementary pivot rule**, that is, the entering variable in any step, is always the complement of the dropping basic variable in the previous step. We prove below that the value of λ in the basic solution keeps on increasing in this phase.

At some stage, let $(\xi_1, \ldots, \xi_{p-1}, \lambda, \xi_{p+1}, \ldots, \xi_n)$ be the ACBV with the values of the basic variables in the corresponding BFS to be $\hat{q} = (\hat{q}_1, \ldots, \hat{q}_n)^T$. So the value

of λ in this solution is $\hat{q}_p$. Let v_s denote the entering variable into this ACBV, as determined by the complementary pivot rule. Let $(\hat{a}_{1s}, \ldots, \hat{a}_{ns})^T$ be the pivot column (updated column of v_s), and let θ denote the minimum ratio, as defined above, for this step. We prove below that $\hat{a}_{ps} \leqq 0$. The solution

$$\begin{aligned} \xi_i &= \hat{q}_i - \alpha \hat{a}_{is}, i \in \{1, \ldots, p-1, p+1, \ldots, n\} \\ v_s &= \alpha \\ \text{all other variables} &= 0 \end{aligned}$$

is a complementary feasible solution of the original parametric LCP when $\lambda = \hat{q}_p - \alpha \hat{a}_{ps}$, for $0 \leqq \alpha \leqq \theta$. As the value of λ keeps on increasing during this phase, this process keeps getting solutions of the original parametric LCP for higher and higher values of λ, as the phase progesses.

This phase only terminates when an ACBV, say, $(\eta_1, \ldots, \eta_{p-1}, \lambda, \eta_{p+1}, \ldots, \eta_n)$ is reached satisfying the property that if ν denotes the entering variable into this ACBV, as determined by the complementary pivot rule, and $(a_1^*, \ldots, a_n^*)^T$ is the pivot column (updated column of ν), then $a_i^* \leqq 0$ for all $i \in \{1, \ldots, p-1, p+1, \ldots, n\}$. This is similar to ray termination of Chapter 2. Let $q^* = (q_1^*, \ldots, q_n^*)^T$ be the present updated right hand side constants vector. If $a_p^* < 0$, then the solution

$$\begin{aligned} \eta_i &= q_i^* - \alpha a_i^*, i \in \{1, \ldots, p-1, p+1, \ldots, n\} \\ \nu &= \alpha \\ \text{all other variables} &= 0 \end{aligned}$$

is a complementary solution of the original parametric LCP when $\lambda = q_p^* - \alpha a_p^*$, for all $\alpha \geqq 0$. In this case, this solution therefore, provides the solution of the parametric LCP for all $\lambda \geqq q_p^*$, terminate.

If $a_p^* = 0$ when this termination occurs, the original parametric LCP is infeasible whenever $\lambda > q_p^*$ (this fact is proved below), terminate.

Procedure to Decrease the Value of λ

Suppose we have a complementary basic vector $y = (y_1, \ldots, y_n)$, for which the lower characteristic value is $\underline{\lambda}$, finite. Let $\beta = B^{-1}$ be the inverse of the complementary basis corresponding to y, and $\overline{q} = \beta q$, $\overline{q}^* = \beta q^*$. Assuming that y is lexico feasible for $\lambda = \underline{\lambda}$, we have $(\overline{q}_i + \underline{\lambda} \overline{q}_i^*, \beta_i.) \succ 0$ for all i. Determine the i that attains the lexico maximum $\{-(\overline{q}_i, \beta_i.)/(\overline{q}_i^*) : i \text{ such that } \overline{q}_i^* > 0\}$, and suppose it is p. Let the updated column of the complement of y_p in the canonical tableau of (5.7) with respect to y be $(\overline{a}_{1p}, \ldots, \overline{a}_{np})^T$. If $\overline{a}_{pp} < 0$, perform a single principal pivot step in position p in the present complementary basic vector y, and continue in the same way. If $\overline{a}_{pp} = 0$, to decrease λ below $\underline{\lambda}$, enter into a special complementary pivot phase. This phase begins with performing a pivot step in the column of λ in (5.8) with row p as the pivot row, to transform the column of λ in (5.8) into $-I_{.p}$ (the usual pivot step would transform

the column of λ in (5.8) into $+I_{.p}$), leading to an ACBV as before. Except for this change, the complementary pivot procedure is carried out exactly as before. In all the canonical tableaus obtained in this phase, λ remains the pth basic variable, with its updated column as $-I_{.p}$. The value of λ keeps on decreasing as this phase progresses, and termination occurs when ray termination, as described earlier, occurs. During this procedure, the complementary solutions of the original parametric LCP for different values of λ are obtained using the same procedure as discussed earlier, from the basic solution of the system in Tableau 5.1 corresponding to the ACBV at each stage.

Proof of the Algorithm

Here we prove the claims made during the complementary pivot phase for increasing the value of λ.

Theorem 5.2 *Let $(\xi_1, \ldots, \xi_{p-1}, \lambda, \xi_{p+1}, \ldots, \xi_n)$ be an ACBV obtained during this phase. Let $\hat{q} = (\hat{q}_1, \ldots, \hat{q}_n)^T$ be the updated right hand side constants vector with respect to this ACBV. Let v_s denote the entering variable into this ACBV as determined by the complementary pivot rule. Let $(\hat{a}_{1s}, \ldots, \hat{a}_{1n})^T$ be the updated column of v_s. Then $\hat{a}_{ps} \leqq 0$, and the value of λ increases or remains unchanged when v_s enters this ACBV.*

Proof. We will first prove that $\hat{a}_{ps} \leqq 0$. The first ACBV in this phase was $(y_1, \ldots, y_{p-1}, \lambda, y_{p+1}, \ldots, y_n)$ and the entering variable into it is t_p. From the manner in which this phase was initiated, we know that the updated column of t_p in the canonical tableau of (5.7) with respect to y, $(\overline{a}_{1p}, \ldots, \overline{a}_{np})$, has its pth entry $\overline{a}_{pp} = 0$. Thus the pth entry in the column of t_p in Tableau 5.1 is also zero, and when t_p enters the ACBV in Tableau 5.1, no change occurs in its row p, which verifies the statement of this theorem for the initial ACBV in this phase. We will now show that it holds in all subsequent ACBVs obtained in this phase too.

Let $(\zeta_1, \ldots, \zeta_n)$ denote the ACBV just before the current ACBV $(\xi_1, \ldots, \xi_{p-1}, \lambda, \xi_{p+1}, \ldots, \xi_n)$. Suppose the statement of the theorem holds true in all steps of this phase until the ACBV ζ. We will now prove that this implies that the statement of this theorem must also hold for the complementary pivot step of bringing v_s into this ACBV $(\xi_1, \ldots, \xi_{p-1}, \lambda, \xi_{p+1}, \ldots, \xi_n)$.

Let u_s denote the complement of v_s. Since v_s is the entering variable chosen by the complementary pivot rule, u_s must have just dropped out of the basic vector ζ leading to the present basic vector $(\xi_1, \ldots, \xi_{p-1}, \lambda, \xi_{p+1}, \ldots, \xi_n)$. Let u_r denote the entering variable into the ACBV ζ that replaced u_s from it. Suppose the pivot row for entering u_r into ζ was row p' (so, u_s must have been the p'th basic variable in ζ). Let the updated entries in the canonical tableau of Tableau 5.1 with respect to the ACBV ζ, in rows p and p' be as given below.

Variable $\rightarrow$	ζ_1	$\ldots$	$\zeta_{p'} = u_s$	$\ldots$	$\zeta_p = \lambda$	ζ_{p+1}	$\ldots$	ζ_n	u_r	v_s	$\ldots$
row p'	0	$\ldots$	1	$\ldots$	0	0	$\ldots$	0	δ_1	δ_2	$\ldots$
row p	0	$\ldots$	0	$\ldots$	1	0	$\ldots$	0	δ_3	δ_4	$\ldots$

$\zeta = (\zeta_1, \ldots, \zeta_n)$ is an ACBV with $\zeta_p = \lambda$, and u_r is the entering variable into it chosen by the complementary pivot rule. These facts imply that $(\zeta_1, \ldots, \zeta_{p-1}, u_r, \zeta_{p+1}, \ldots, \zeta_n)$ becomes a complementary basic vector when the variables are properly ordered. It cannot be a basic vector unless $\delta_3 \neq 0$. So, $\delta_3 \neq 0$. Also since the statement of the theorem holds for the ACBV ζ, we have $\delta_3 \leqq 0$, so $\delta_3 < 0$. Also, since u_r is the entering variable into the ACBV ζ and row p' is the pivot row for this pivot step, we must have $\delta_1 > 0$. The pivot step in the column of u_r with δ_1 as the pivot element, transforms δ_4 into $\delta_4 - \frac{\delta_3 \delta_2}{\delta_1}$, by definition this is $\hat{a}_{ps}$, and we want to show that this is $\leqq 0$. As mentioned earlier, $(\zeta_1, \ldots, \zeta_{p-1}, u_r, \zeta_{p+1}, \ldots, \zeta_n)$ is a permutation of a complementary basic vector. So in the canonical tableau with respect to ζ, if we perform a pivot step in the column of u_r, with δ_3 as the pivot element (row p as the pivot row) and rearrange the rows and columns properly, we get the canonical tableau with respect to a complementary basic vector. This pivot step transforms the element δ_2 in the column of v_s into $\delta_2 - \frac{\delta_4 \delta_1}{\delta_3}$, this will be the entry in the column of v_s in row p', which is the row in which u_s is the basic variable. M is PSD, by the results of Chapter 3 every PPT of a PSD matrix is PSD, and by the results in Chapter 1 every diagonal entry in a PSD matrix is $\geqq 0$, these facts imply that this element $\delta_2 - \frac{\delta_4 \delta_1}{\delta_3} \leqq 0$. This, and $\delta_3 < 0$, $\delta_1 > 0$ established earlier imply that $\hat{a}_{ps} = \delta_4 - \frac{\delta_3 \delta_2}{\delta_1} \leqq 0$.

In all the pivot steps in this phase, the pivot elements are > 0, and all the updated right hand side constants with the possible exception of the pth, stay $\geqq 0$. These facts, and the fact that $\hat{a}_{ps} \leqq 0$ imply that when v_s enters the ACBV $(\xi_1, \ldots, \xi_{p-1}, \lambda, \xi_{p+1}, \ldots, \xi_n)$, the value of λ, the pth basic variable, either increases or stays the same (but never decreases).

Thus if the statement of the theorem holds for the ACBV ζ, it must hold for the ACBV $(\xi_1, \ldots, \xi_{p-1}, \lambda, \xi_{p+1}, \ldots, \xi_n)$ following it. We have already established the theorem for the initial ACBV in this phase. Hence, by induction, the theorem holds in all ACBVs obtained during this phase.

□

So, the value of λ, the pth basic variable in the ACBV, increases during this phase. From the arguments used in Chapter 2, it is clear that the adjacency path of ACBVs in this phase continues unambiguously, and no ACBV can reappear. Since there are only a finite number of ACBVs, these facts imply that this phase must terminate with the special type of ray termination discussed here, after at most a finite number of steps.

We will now prove the claims made when ray termination occurs in this phase.

Theorem 5.3 *Let $(\eta_1, \ldots, \eta_{p-1}, \lambda, \eta_{p+1}, \ldots, \eta_n)$ be the terminal ACBV in the complementary pivot phase to increase λ. Let ν denote the entering variable into this ACBV chosen by the complementary pivot rule, and let $(a_1^*, \ldots, a_n^*)^T$ be the updated column of ν with respect to this ACBV. Let $q^* = (q_1^*, \ldots, q_n^*)^T$ be the updated right hand side constants vector with respect to this terminal ACBV in this phase. If $a_p^* = 0$, the original parametric LCP has no solution when $\lambda > q_p^*$.*

Proof. Let the complement of ν be u'. By the arguments used earlier, $(\eta_1, \ldots, \eta_{p-1}, u', \eta_{p+1}, \ldots, \eta_n)$ must be a permutation of a complementary basic vector. So, performing a pivot step in the canonical tableau with respect to the ACBV $(\eta_1, \ldots, \eta_{p-1}, \lambda, \eta_{p+1}, \ldots, \eta_n)$ with u' as the entering variable and row p as the pivot row, leads to a canonical tableau with respect to a complementary basic vector, with some rows and columns rearranged. Since $a_p^* = 0$, this pivot step would not alter the column vector of ν, and hence it remains as $(a_1^*, \ldots, a_n^*)^T \leqq 0$ with $a_p^* = 0$. M is PSD, and every PPT of a PSD matrix is PSD. These facts together with Result 1.6 imply that the updated row corresponding to u' in the canonical tableau (5.7) with respect to the complementary basic vector which is a permutation of $(\eta_1, \ldots, \eta_{p-1}, u', \eta_{p+1}, \ldots, \eta_n)$, has all nonnegative entries on the left hand side. When $\lambda > q_p^*$, the updated right hand side constant in this row will be < 0. This implies that the system (5.7) cannot have a nonnegative solution when $\lambda > q_p^*$, that is, that the original parametric LCP has no solution when $\lambda > q_p^*$.

□

5.2 Exercises

5.3 Let M, q, q^*, a be given matrices of orders $n \times n$, $n \times 1$, $n \times 1$, $n \times 1$ respectively. Assume that M is a P-matrix. Let $(w(\lambda), z(\lambda))$ be the solution of the parametric LCP $(q + \lambda q^*, M)$ as a function of λ. Let $\overline{\lambda}$ = maximum $\{\lambda : z(\lambda) \leqq a\}$. Also, let $\hat{\lambda}$ = maximum $\{\lambda : z(\alpha) \leqq a$, for all α satisfying $0 \leqq \alpha \leqq \lambda\}$. Discuss an efficient algorithm for find $\overline{\lambda}$, given M, q, q^*, a. Also, derive necessary and sufficient conditions on this data for $\hat{\lambda} = \overline{\lambda}$ to hold. (I. Kaneko [5.9] and O. De Donato and G. Maier [1.4]).

5.4 Let

$$M = \begin{pmatrix} 2 & 1 & -1 \\ -1 & 3 & 0 \\ 0 & 1 & 4 \end{pmatrix}, \qquad q(\lambda) = \begin{pmatrix} 1 & - & \lambda \\ 2 & + & \lambda \\ 3 & & -2\lambda \end{pmatrix}.$$

Solve the parametric LCP $(q(\lambda), M)$ for all real values of λ.

5.5 Let $q = (-1, -2, -3)^T$ and M be the matrix given in Exercise 5.4. Solve the LCP (q, M) by Principal Pivoting Method I.

5.6 Prove that the value of z_0 (artificial variable) is strictly monotone decreasing in the complementary pivot method when applied on the LCP (q, M) associated with a P-matrix.

Prove that the same thing is true when the LCP (q, M) is one corresponding to an LP.

Is it also true when the LCP is one corresponding to a convex quadratic program in which the matrix D is PSD and not PD?

5.7 Consider the following problem

$$\begin{aligned} \text{minimize } z(x) &= cx + \alpha\sqrt{(1/2)(x^T Dx)} \\ \text{subject to } Ax &\geqq b \\ x &\geqq 0 \end{aligned}$$

where D is a square symmetric PD matrix of order n, $\alpha > 0$, A is an $m \times n$ matrix and $b \in \mathbf{R}^m$. Let $\mathbf{K}$ denote the set of feasible solutions of this problem.

i) Show that $z(x)$ is a convex function which is a homogeneous function of degree 1.

ii) If $\alpha < \sqrt{2cD^{-1}c^T}$, prove that every optimum solution of this problem must be a boundary point of $\mathbf{K}$.

iii) If $0 \notin \mathbf{K}$ and if the problem has an optimum solution, prove that there exists a boundary point of $\mathbf{K}$ which is an optimum solution of the problem.

iv) Develop an efficient procedure for solving this problem.

v) Solve the problem

$$\begin{aligned} \text{minimize } & -x_1 - x_2 + \sqrt{(x_1^2 + x_2^2)/2} \\ \text{subject to } & -x_1 - 3x_2 \geqq -14 \\ & -x_1 + x_2 \geqq -2 \\ & x_1, x_2 \geqq 0 \end{aligned}$$

using the method developed in (iv). (C. Sodini [5.15]).

5.8 Consider the following problem

$$\begin{aligned} \text{minimize } f(x) &= (c_0 + cx + (1/2)x^T Dx)/(d_0 + dx)^p \\ \text{subject to } Ax &\geqq b \\ x &\geqq 0 \end{aligned}$$

where D is a square symmetric PD matrix of order n, p is 1 or 2 and $d_0 + dx > 0$ over $x \in \mathbf{K} = \{x : Ax \geqq b, x \geqq 0\}$. Develop an approach for solving this problem. (S. Schaible [5.14]; A. Cambini, L. Martein and C. Sodini [5.4]).

5.9 Consider the following problem

$$\begin{aligned} \text{minimize } Q_\alpha(x) &= cx + \frac{\alpha}{2}x^T Dx \\ \text{subject to } Ax &\geqq b \\ x &\geqq 0 \end{aligned}$$

where D is a PSD matrix of order n, and α is a nonnegative parameter. It is required to solve this problem for all $\alpha \geqq 0$. Formulate this problem as a parametric LCP of the form $(q + \lambda q^*, M)$, $\lambda \geqq 0$, and discuss how to solve it.

Note 5.1 This problem arises in the study of portfolio models. The linear function $(-cx)$ may represent the expected yield, and the quadratic term $\frac{1}{2}x^T Dx$ may be the variance of the yield (the variance measures the extent of random fluctuation in the actual yield from the expected). $Q_\alpha(x)$ is a positive weighted combination of the two objectives which are to be minimized in this model.

5.10 If q is nondegenerate in the LCP (q, M) (i. e., if every solution (w, z) to the system of equations, $w - Mz = q$, makes at least n variables nonzero), prove that the number of solutions of the LCP (q, M) is finite.

5.11 Let $\mathcal{C}_1$ be the set of solutions of

$$\begin{aligned} w - Mz &= q \\ w, z &\geqq 0 \\ w_j z_j &= 0, \ \ j = 2 \text{ to } n. \end{aligned}$$

Prove that $\mathcal{C}_1$ is the union of disjoint paths in $\mathbf{R}^n$.

5.12 Consider the LCP (q, M). Define $\mathbf{S}(q) = \{(w, z) : (w, z)$ is a solution of the LCP $(q, M)\}$. Prove that if there exists a $q \in \mathbf{R}^n$ such that $\mathbf{S}(q)$ is a nonempty unbounded set, then $\mathbf{S}(0)$ contains a nonzero point, that is, the LCP $(0, M)$ has a nonzero solution.

5.3 References

5.1 R. Benveniste, "One Way to Solve the Parametric Quadratic Programming Problem", *Mathematical Programming*, 16 (1979) 63–80.

5.2 M. J. Best, "An Algorithm for the Solution of the Parametric Quadratic Programming Problem", CORR Report 82-24, Faculty of Mathematics, University of Waterloo, July 1982.

5.3 J. C. G. Boot, "On Sensitivity Analysis in Convex Quadratic Programming Problems", *Operations Research*, 11 (1963) 771–786.

5.4 A. Cambini, L. Martein and C. Sodini, "An Algorithm for two particular nonlinear fractional programs", Technical Report, Dipartimento di Matematica, Universita di Pisa, Italy, 1983.

5.5 R. W. Cottle, "Monotone Solutions of the Parametric Linear Complementary Problem", *Mathematical Programming*, 3 (1972) 210–224.

5.6 A. Gana, "Studies in the Complementarity Problem", Ph. D. Dissertation, Department of Industrial and Operations Engineering, The University of Michigan, Ann Arbor, Michigan 48109, USA, 1982.

5.7 M. D. Grigoriadis and K. Ritter, "A Parametric Method for Semidefinite Quadratic Programs", *SIAM Journal on Control*, 7 (1969) 559–577.

5.8 I. Kaneko, "Isotone Solutions of Parametric Linear Complementarity Problems", *Mathematical Programming*, 12 (1977) 48–59.

5.9 I. Kaneko, "A Maximization Problem Related to Parametric Linear Complementarity", *SIAM Journal on Control and Optimization*, 16 (1978) 41–55.

5.10 I. Kaneko, "A Parametric Linear Complementarity Problem Involving Derivatives", *Mathematical Programming*, 15 (1978) 146–154.

5.11 N. Meggido, "On Monotonicity in Parametric Linear Complementarity Problems", *Mathematical Programming*, 12 (1977) 60–66.

5.12 K. G. Murty, "On the Parametric Complementarity Problem", Engineering Summer Conference Notes, The University of Michigan, 1971.

5.13 J. S. Pang and R. Chandrasekaran, "Linear Complementarity Problems solvable by a Polynomially Bounded Pivoting Algorithm", *Mathematical Programming Study*, 25 (1985) 13–27.

5.14 S. Schaible, "Fractional Programming: Applications and Algorithms", *European Journal of Operations Research*, 7 (1981) 111–120.

5.15 C. Sodini, "Minimizing the sum of a linear function and the square root of a convex quadratic form", Technical Report, Dipartimento di Matematica, Universita di Pisa, Italy, 1983.

Chapter 6

COMPUTATIONAL COMPLEXITY OF COMPLEMENTARY PIVOT METHODS

In this Chapter, we discuss the worst case behavior of the computational growth requirements of the complementary and principal pivot methods for solving LCPs, as a function of n, the order, and the size of the LCP. These results are from K. G. Murty [6.5]. We construct a class of LCPs with integer data, one of order n for each $n \geqq 2$, and prove that the pivotal methods discussed in Chapters 2, 4, and 5 require $2^{\mathbf{n}} - 1$ or $2^{\mathbf{n}}$ pivot steps to solve the problem of order n in the class. The size of the n^{th} problem in this class, defined to be the total number of bits of storage needed to store all the data in the problem in binary form is $\leqq 4n^{\mathbf{2}} + 3n$. These results establish that in the worst case, the computational growth requirements of complementary pivot methods are not bounded above by any polynomial in the order or size of the LCP.

To study the worst case computational complexity of complementary pivot methods, we look at the following question: What is the maximum number of complementary cones through which a straight line in $\mathbf{R}^n$ can cut across? For a problem of order n, the answer turns out to be $2^{\mathbf{n}}$, that is, there may exist straight lines which cut across the interiors of every one of the $2^{\mathbf{n}}$ complementary cones.

Let $\widetilde{M}(n) = (\widetilde{m}_{ij})$ be the lower triangular matrix of order n, defined by $\widetilde{m}_{ij} = 1$ for $i = 1$ to n, $\widetilde{m}_{ij} = 0$ for all $j > i$, and $\widetilde{m}_{ij} = 2$ for all $j < i$. See (1.15), page 19. Since $\widetilde{M}(n)$ is lower triangular, all principal subdeterminants of $\widetilde{M}(n)$ are equal to 1, and hence $\widetilde{M}(n)$ is a P-matrix. Since $\widetilde{M}(n) + (\widetilde{M}(n))^T$ is a matrix all of whose entries are 2, it is singular, and clearly it is a PSD matrix. Hence $\widetilde{M}(n)$ is a P-matrix, PSD matrix (and hence a copositive plus matrix), but not a PD matrix. Let e_n be the column vector in $\mathbf{R}^n$ all of whose entries are equal to 1. Let:

$$\begin{aligned}
\tilde{q}(n) =&(2^{\mathbf{n}}, 2^{\mathbf{n-1}}, \ldots, 2)^T \\
\hat{q}(n) =&(-2^{\mathbf{n}}, -2^{\mathbf{n}} - 2^{\mathbf{n-1}}, -2^{\mathbf{n}} - 2^{\mathbf{n-1}} - 2^{\mathbf{n-2}}, \ldots, \\
&- 2^{\mathbf{n}} - 2^{\mathbf{n-1}} - \ldots - 2^{\mathbf{2}} - 2)^T \qquad (6.1) \\
a(s) =&2^{\mathbf{s}} - 1, \text{ for any } s \geqq 2 \\
\mathbf{L}(n) =&\{x : x = \tilde{q}(n) + \gamma(-e_n) : \gamma \text{ a real parameter}\}
\end{aligned}$$

Theorem 6.1 . *The straight line* $\mathbf{L}(n)$ *cuts across the interior of every one of the* $2^{\mathbf{n}}$ *complementary cones in the class* $\mathcal{C}(\widetilde{M}(n))$ *for any* $n \geqq 2$.

Proof. Consider the class of parametric LCPs $(\tilde{q}(n)+\gamma(-e_n), \widetilde{M}(n))$ for $n \geqq 2$, where γ is a real valued parameter. Consider the case $n = 2$ first. The following can be verified in this case :

Tableau 6.1

Complementary Cone corresponding to the Complementary Basic Vector	Portions of $\mathbf{L}(2)$ corresponding to values of the Parametr γ which lie in this Complementary Cone
(w_1, w_2)	$2 = \gamma$
(w_1, z_2)	$2 \leqq \gamma \leqq 4$
(z_1, z_2)	$4 \leqq \gamma \leqq 6$
(z_1, w_2)	$6 \leqq \gamma$

Also whenever γ is an interior point of one of these intervals, all the basic variables are strictly positive in the complementary BFS of $(\tilde{q}(2) + \gamma(-e_n), \widetilde{M}(2))$; and this implies that the point $\tilde{q}(2) + \gamma(-e_n)$ corresponding to that value of γ is in the interior of the corresponding complementary cone. Hence, the statement of this Theorem is true when $n = 2$. We now make an induction hypothesis.

Induction Hypothesis: The theorem is true for the LCP of order $n - 1$ in the class. Specifically, the complementary basic vectors for the parametric LCP $(\tilde{q}(n - 1) + \gamma(-e_n), \widetilde{M}(n-1))$ can be ordered as a sequence $v_0, v_1, \ldots, v_{a(n-1)}$, such that the complementary cone corresponding to the complementary basic vector v_r contains the portion of the straight line $\mathbf{L}(n-1)$ corresponding to $\gamma \leqq 2$ if $r = 0$; $2r \leqq \gamma \leqq 2(r+1)$, if $1 \leqq r \leqq 2^{\mathbf{n-1}} - 2$; and $2^{\mathbf{n}} - 2 \leqq \gamma$ if $r = 2^{\mathbf{n-1}} - 1$. Also the straight line $\mathbf{L}(n-1)$ cuts across the interior of each of these complementary cones.

Now consider the parametric LCP of order n in the class, namely $(\tilde{q}(n) + \gamma(-e_n), \widetilde{M}(n))$, the original tableau for which is Tableau 6.2

Tableau 6.2

w_1	w_2	...	w_n	z_1	z_2	...	z_n	
1	0	...	0	-1	0	...	0	$2^{\mathbf{n}} - \gamma$
0	1	...	0	-2	-1	...	0	$2^{\mathbf{n-1}} - \gamma$
$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
0	1	...	1	-2	-2	...	-1	$2 - \gamma$

The principal subproblem of this in the variables $(w_2, \ldots, w_n)$, $(z_2, \ldots, z_n)$, is the same as the parametric LCP of order $n-1$ in the class we are discussing, with the exception that the variables in it are called as $w_2, \ldots, w_n$; $z_2, \ldots, z_n$. By induction hypothesis, the complementary basic vectors of this principal subproblem can be ordered in a sequence as $v_0, v_1, \ldots, v_{a(n-1)}$, where $v_0 = (w_2, \ldots, w_n)$, $v_1 = (w_2, \ldots, w_{n-1}, z_n)$, etc. such that the complementary cone for this principal subproblem, corresponding to the complementary basic vector v_r, contains the portion of the straight line $\mathbf{L}(n-1)$ corresponding to $\gamma \leqq 2$ if $r = 0$; $2r \leqq \gamma \leqq 2(r+1)$ if $1 \leqq r \leqq 2^{\mathbf{n-1}} - 2$, and $\gamma \geqq 2^{\mathbf{n}} - 2$ if $r = 2^{\mathbf{n-1}} - 1$; and as long as γ is in the interior of one of these intervals, the corresponding point on $L(n-1)$ is in the interior of the corresponding complementary cone. Notice that in the original problem in Tableau 6.2, $q_1(\gamma) = 2^{\mathbf{n}} - \gamma$ remains nonnegative for all $\gamma \leqq 2^{\mathbf{n}}$ and strictly positive for all $\gamma < 2^{\mathbf{n}}$. This, together with the result for the principal subproblem, implies that the complementary cone corresponding to the complementary basic vector $V_r = (w_1, v_r)$ of the original problem (Tableau 6.2) contains the portion of the line $\mathbf{L}(n)$ corresponding to values of γ satisfying $\gamma \leqq 2$, if $r = 0$; $2r \leqq \gamma \leqq 2r + 2$, if $1 \leqq r \leqq -1 + 2^{\mathbf{n-1}} = a(n-1)$. It also implies that in each case, the straight line $\mathbf{L}(n)$ cuts across the interior of these complementary cones.

Now perform a single principal pivot step in Position 1 in the original problem in Tableau 6.2. This leads to Tableau 6.3

Tableau 6.3

w_1	w_2	...	w_n	z_1	z_2	...	z_n	q
-1	0	...	0	1	0	...	0	$\gamma - 2^{\mathbf{n}}$
-2	1	...	0	0	-1	...	0	$(-2^{\mathbf{n+1}} + \gamma) + 2^{\mathbf{n-1}}$
-2	0	...	0	0	-2	...	0	$(-2^{\mathbf{n+1}} + \gamma) + 2^{\mathbf{n-2}}$
$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
-2	0	...	1	0	-2	...	-1	$(-2^{\mathbf{n+1}} + \gamma) + 2$

Let $-\lambda = -2^{\mathbf{n+1}} + \gamma$ and treat λ as the new parameter. As γ increases from $2^{\mathbf{n}}$ to $2^{\mathbf{n+1}} - 2$, λ decreases from $2^{\mathbf{n}}$ to 2. As a function of λ, the vector of the right hand side constants in Tableau 6.3 is $(2^{\mathbf{n}} - \lambda, 2^{\mathbf{n-1}} - \lambda, \ldots, 2 - \lambda)^T$.

Now look at the principal subproblem of the parametric LCP in Tableau 6.3 in the variables $w_2, \ldots, w_n; z_2, \ldots, z_n$. This principal subproblem considered with λ as the parameter can be verified to be the same as the parametric LCP of order $n-1$ in the class we are discussing, with the exception that the variables in it are called as $w_2, \ldots, w_n; z_2, \ldots, z_n$, and the parameter is λ.

Using arguments similar to those as above on these problems, and translating everything to the original parameter γ again, we conclude that the complementary cone corresponding to the complementary basic vector $V_r = (z_1, v_{b(r)})$ of the original problem, where $b(r) = \mathbf{2^n} - r - 1$, contains the portion of the straight line $\mathbf{L}(n)$ corresponding to values of γ satisfying $2r \leqq \gamma \leqq 2r+2$, if $\mathbf{2^{n-1}} \leqq r \leqq \mathbf{2^n} - 2$; and $\gamma \geqq \mathbf{2^{n+1}} - 2$, if $r = \mathbf{2^n} - 1$.

Thus if $v_0, \ldots, v_{a(n-1)}$ is the ordered sequence of complementary basic vectors for the principal subproblem of the parametric LCP in Tableau 6.2 in the variables $w_2, \ldots, w_n; z_2, \ldots, z_n$; let the ordered sequence of complementary basic vectors for the parametric LCP in Tableau 6.2 be

$$\begin{aligned} V_0 =& (w_1, v_0), (w_1, v_1), \ldots, (w_1, v_{a(n-1)}); \\ & (z_1, v_{a(n-1)}), (z_1, v_{a(n-1)-1}), \ldots, (z_1, v_0) = V_{a(n)}. \end{aligned} \tag{6.2}$$

Then the induction hypothesis implies the result that the complementary cone corresponding to the complementary basic vector V_r contains the portion of the straight line $\mathbf{L}(n)$ corresponding to $\gamma \leqq 2$, if $r = 0$; $2r \leqq \gamma \leqq 2r+2$, if $1 \leqq r \leqq \mathbf{2^n} - 2$; $\gamma \geqq \mathbf{2^{n+1}} - 2$, if $r = \mathbf{2^n} - 1$. Also in each case, the straight line cuts across the interior of the complementary cone. Hence the induction hypothesis implies that the statement of Theorem 6.1 also holds for the parametric LCP of order n in the class we are discussing. The statement of Theorem 6.1 has already been verified to be true from $n = 2$. Hence it is true for all $n \geqq 2$.

□

6.1 Computational Complexity of the Parametric LCP Algorithm

Theorem 6.2 *Consider the class of parametric LCPs $(\tilde{q}(n) + \gamma(-e_n), \widetilde{M}(N))$, for $n \geqq 2$. The parametric LCP algorithm discussed in Chapter 5 requires $\mathbf{2^n}$ pivot steps to solve the nth problem in the class for all real values of the parameter γ.*

Proof. Let $V_0, V_1, \ldots, V_{a(n)}$ be the sequence of complementary basic vectors for the parametric LCP of order n in this class obtained in the proof of Theorem 6.1. From the proof of Theorem 6.1, we conlcude that the complementary basic vector V_r is feasible to the parametric LCP $(\tilde{q}(n) + \gamma(-e_n), \widetilde{M}(n))$ in the interval $\gamma \leqq 2$ if $r = 0$; $2r \leqq \gamma \leqq 2r+2$, if $1 \leqq r \leqq \mathbf{2^n} - 2$; $\gamma \geqq \mathbf{2^{n+1}} - 2$, if $r = \mathbf{2^n} - 1$. Hence, when the parametric LCP

algorithm is applied to solve $(\tilde{q}(n) + \gamma(-e_n), \widetilde{M}(n))$ for all values of the parameter γ, it terminates only after going through all the complementary basic vectors, $V_0, V_1, \ldots,$ $V_{a(n)}$; and thus requires $a(n) + 1 = 2^{\mathbf{n}}$ pivot steps.

□

Example 6.1

See Example 5.1 in Chapter 5. There the parametric LCP $(\tilde{q}(3) + \gamma(-e_3), \widetilde{M}(3))$ is solved for all values of the parameter γ (there the parameter is denoted by λ intead γ) using the parametric LCP algorithm and verify that it took $2^{\mathbf{3}} = 8$ pivot steps in all.

6.2 Geometric Interpretation of a Pivot Step in the Complementary Pivot Method

Let M be a given square matrix of order n, and q a column vector in $\mathbf{R}^n$. Consider the LCP (q, M). The original tableau for solving it by the complementary pivot method is (2.3) of Section 2.2.1.

Let $(y_1, \ldots, y_{r-1}, y_{r+1}, \ldots, y_n, z_0)$ be a basic vector obtained in the process of solving this LCP by the complementary pivot method where $y_j \in \{w_j, z_j\}$ for all j. Let $A_{.j}$ denote the column vector associated with y_j in (2.3) for each j. If $\bar{q} = (\bar{q}_1, \ldots, \bar{q}_n)^T$ is the update right hand constants vector in the canonical tableau of (2.3) with respect to the basic vector $(y_1, \ldots, y_{r-1}, y_{r+1}, \ldots, y_n, z_0)$, then $\bar{q} \geqq 0$ (since this basic vector must be a feasible basic vector) and we have

$$q = A_{.1}\bar{q}_1 + \ldots + A_{.r-1}\bar{q}_{r-1} + A_{.r+1}\bar{q}_r + \ldots + A_{.n}\bar{q}_{n-1} + (-e_n)\bar{q}_n \qquad (6.3)$$

$\bar{q}_n$ is the value of z_0 in the present BFS. If $\bar{q}_n = 0$, the present BFS is a complementary feasible solution and the method would terminate. So assume $\bar{q}_n > 0$ and denote it by the symbol $\tilde{z}_0$. Then (6.3) implies that $q + \tilde{z}_0 e_n \in \text{Pos}\{A_{.1}, \ldots, A_{.r-1}, A_{.r+1}, \ldots, A_{.n}\}$. The present left out complementary pair is (w_r, z_r), and one of the variables from this pair will be choosen as the entering variable at this stage, let us denote it by $y_r \in \{w_r, z_r\}$ and let $A_{.r}$ denote the column vector associated with y_r in (2.3). If y_r replaces z_0 from the basic vector in theis step, we get a complementary feasible basic vector at the end of this pivot step, and the method terminates. Suppose the dropping variable is not z_0, but some y_i for $i \in \{1, \ldots, r-1, r+1, \ldots, n\}$. Let $\hat{z}_0 > 0$ be the value of z_0 in the new BFS obtained after this pivot step. Then using the same arguments as before we conclude that $q + \hat{z}_0 e_n \in \text{Pos}\{A_{.1}, \ldots, A_{.i-1}, A_{.i+1}, \ldots, A_{.n}\}$.

Under these conditions, clearly $(y_1, \ldots, y_n)$ is itself a complementary basic vector, and let $\mathbf{K} = \text{Pos}(A_{.1}, \ldots, A_{.n})$ be the complementary cone associated with it. The net effect in this pivot step is therefore that of moving from the point $q + \tilde{z}_0 e_n$ contained

on the facet $\text{Pos}\{A_{.1}, \ldots, A_{.r-1}, A_{.r+1}, \ldots, A_{.n}\}$ of $\mathbf{K}$ to the point $q + \hat{z}_0 e_n$ on the facet $\text{Pos}\{A_{.1}, \ldots, A_{.i-1}, A_{.i+1}, \ldots, A_{.n}\}$ of $\mathbf{K}$, along the half-line $\{x : x = q + \lambda e_n,$ λ a nonnegative real number$\}$. See Figure 6.1. The complementary pivot method continues in this manner walking along the half-line $\{x : x = q + \lambda e_n, \lambda \geqq 0\}$ cutting across different complementary cones, until at some stage it enters a complementary cone containing the point q on this half-line.

We will now use this geometric interpretation, to establish the computational complexity of the complementary pivot method in the worst case.

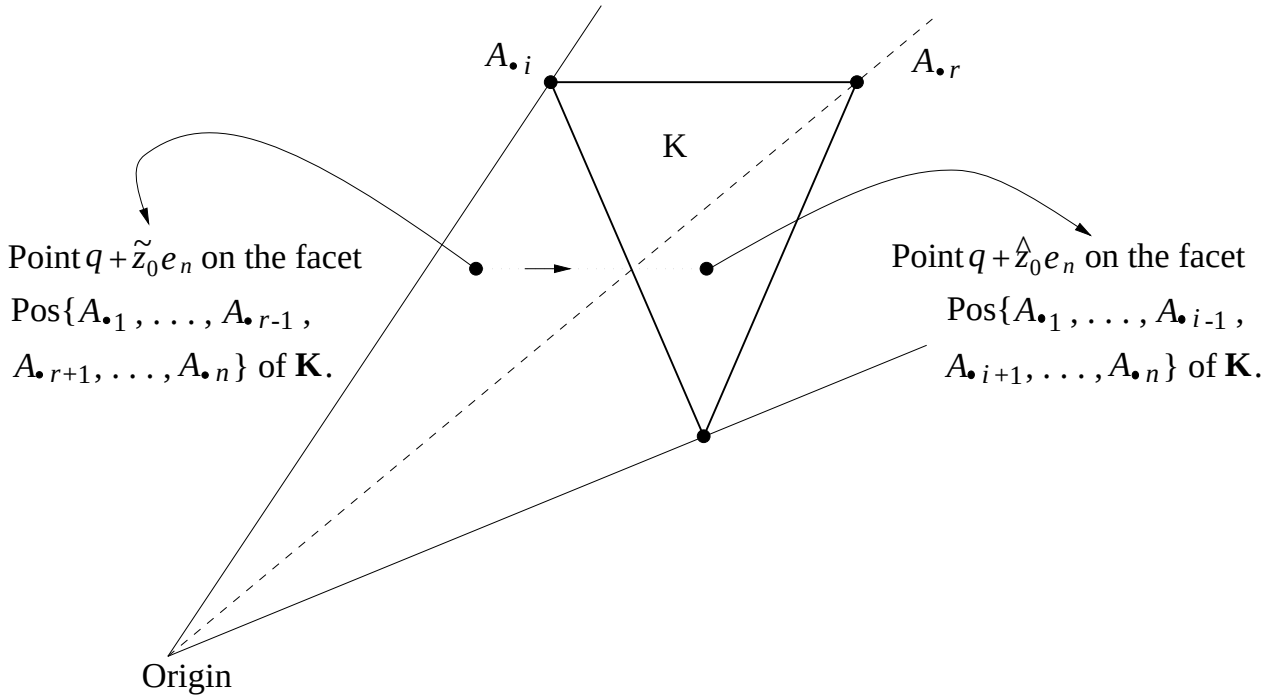


Figure 6.1 Geometric interpretation of a pivot step in the complementary pivot method as a walk from one facet of a complementary cone to another facet of the same cone along the half-line $\{x : x = q + \lambda e_n, \lambda \geqq 0\}$ as λ varies from $\tilde{z}_0$ to $\hat{z}_0$.

6.3 Computational Complexity of the Complementary Pivot Method

Theorem 6.3 *For any $n \geqq 2$, the complementary pivot method requires $2^{\mathbf{n}}$ pivot steps before termination when applied on the LCP $(\hat{q}(n), \widetilde{M}(n))$.*

Proof. Notice that $\tilde{q}(n) + \gamma(-e_n) = \hat{q}(n) + \lambda e_n$ where $\lambda = 2^{\mathbf{n+1}} - \gamma$. Hence the straight line $\{x : x + \hat{q}(n) + \lambda e_n, \lambda \text{ a real number}\}$ is the same as the line $\mathbf{L}(n)$ defined in equation (6.1) but in the reverse direction.

The original tableau for applying the complementary pivot method to solve the LCP $(\hat{q}(n), M(n))$ is shown in Tableau 6.4.

Tableau 6.4

w_1	w_2	...	w_n	z_1	z_2	...	z_n	z_0	q
1	0	...	0	-1	0	...	0	-1	$-2^{\mathbf{n}}$
0	1	...	0	-2	-1	...	0	-1	$-2^{\mathbf{n}} - 2^{\mathbf{n-1}}$
$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$
0	0	...	1	-2	-2	...	-1	-1	$-2^{\mathbf{n}} - 2^{\mathbf{n-1}} - \ldots - 2$

The initial basic vector obtained in the complementary pivot method is $(w_1, \ldots, w_{n-1}, z_0)$ and in the solution corresponding to this basic vector, the value of z_0 is $2^{\mathbf{n}} + 2^{\mathbf{n-1}} + \ldots + 2 = 2^{\mathbf{n+1}} - 2$. The entering variable into the basic vector at this initial stage is z_n.

Let $V_0, \ldots, V_{a(n)}$ be the ordering of the complementary basic vectors for this problem, obtained in the proof of Theorem 6.1. $V_0 = (w_1, \ldots, w_n)$, $V_1 = (w_1, \ldots, w_{n-1}, z_n)$, etc. Let $\mathbf{K}_r$ be the complementary cone corresponding to the complementary basic vector V_r for the LCP $(\hat{q}(n), \widetilde{M}(n))$. Using the geometric interpretation discussed above, the effect of the initial pivot step of bringing z_n into the basic vector $(w_1, \ldots, w_{n-1}, z_0)$ can be interpreted as a walk through the complementary cone $\mathbf{K}_1$, beginning with the point $\hat{q}(n) + \lambda_1 e_n$ (where $\lambda_1 = 2^{\mathbf{n+1}} - 2$) on the facet of $\mathbf{K}_1$ corresponding to the Pos cone of the columns of $w_1, \ldots, w_{n-2}, w_{n-1}$ in Tableau 6.4, to the point $\hat{q}(n) + \lambda_2 e_n$ (where $\lambda_2 = 2^{n+1} - 4$) on the facet of $\mathbf{K}_1$ corresponding to the Pos cone of the columns of $w_1, \ldots, w_{n-2}, z_n$ in Tableau 6.4 along the half-line $\{x : x = \hat{q}(n) + \lambda e_n, \lambda \geqq 0\}$. Here λ_1, λ_2 are the values of z_0 in the basic solution of Tableau 6.4 corresponding to the basic vectors $(w_1, \ldots, w_{n-1}, z_0)$ and $(w_1, \ldots, w_{n-2}, z_n, z_0)$ respectively. Thus the initial pivot step of introducing z_n into the basic vector $(w_1, \ldots, w_{n-1}, z_0)$ can be interpreted as the walk across the complementary cone $\mathbf{K}_1$, starting at the value $\lambda = \lambda_1$ to the value $\lambda = \lambda_2$ along the half-line $\{x : x = \hat{q}(n) + \lambda e_n, \lambda \geq 0\}$. Similarly the r^{th} pivot step performed during the complementary pivot method applied on this problem, can be interpreted as the walk through the complementary cone $\mathbf{K}_r$ along the half-line

$\{x : x = \hat{q}(n) + \lambda e_n, \lambda \geqq 0\}$, for $r \geqq 2$. Since the straight line $\{x : x = \hat{q}(n) + \lambda e_n,$ λ a real number$\}$ is the same as the line $\mathbf{L}(n)$ defined in equation (6.1), from the results in the proof of Theorem 6.1 and the geometric interpretation of the pivot steps in the complementary pivot method discussed above, we reach the following conclusions: the complementary pivot method starts with a value for z_0 of $\mathbf{2^{n+1} - 2}$ in the initial step. All pivot steps are nondegenerate and the value of z_0 decreases by 2 in every pivot step. Hence the method terminates when the value of z_0 becomes zero after the $(\mathbf{2^n - 1})$th pivot step. This last pivot step in the method corresponds to a walk into the complementary cone $\mathbf{K}_{a(n)}$ associated with the complementary basic vector $V_{a(n)} = (z_1, w_2, \ldots, w_n)$ along the half-line $\{x : x = \hat{q}(n) + \lambda e_n, \lambda \geqq 0\}$. Hence the terminal basic vector obtained in the complementary pivot method applied on this problem will be $(z_1, w_2, \ldots, w_n)$ and it can be verified that the solution of the LCP $(\hat{q}(n), \widetilde{M}(n))$ is $(w = (0, \mathbf{2^{n-1}}, \ldots, 2)$, $z = (\mathbf{2^n}, 0, \ldots, 0))$. Therefore counting the first pivot step in which the canonical tableau with respect to the initial basic vector $(w_1, \ldots, w_{n-1}, z_0)$ is obtained, the complementary pivot method requires $\mathbf{2^n}$ pivot steps for solving the LCP $(\hat{q}(n), \widetilde{M}(n))$, for any $n \geqq 2$.

□

Example 6.2

See Example 2.10 in Section 2.2.7 where the LCP $(\hat{q}(3), \widetilde{M}(3))$ of order 3 is solved by the complementary pivot method and verify that it required $\mathbf{2^3} = 8$ pivot steps before termination.

6.4 Computational Complexity of Principal Pivoting Method I

Theorem 6.4 *Principal pivoting Method I requires* $\mathbf{2^n - 1}$ *pivot steps before termination, when applied on the LCP* $(-e_n, \widetilde{M}(n))$, *for any* $n \geqq 2$.

Proof. Proof is by induction on n. The original tableau for this problem is shown in Tableau 6.5

Tableau 6.5

w_1	w_2	...	w_n	z_1	z_2	...	z_n	q
1	0	...	0	−1	0	...	0	−1
0	1	...	0	−2	−1	...	0	−1
⋮	⋮		⋮	⋮	⋮		⋮	⋮
0	0	...	1	−2	−2	...	−1	−1

It can be verified that $(z_1, w_2, \ldots, w_n)$ is a complementary feasible basis for Tableau 6.5 and the solution of this LCP is $(w_1, \ldots, w_n) = (0, 1, \ldots, 1)$, $(z_1, \ldots, z_n) = (1, 0, \ldots, 0)$.

In Example 4.1 of Section 4.1 the LCP $(-e_3, \widetilde{M}(3))$ was solved by Principal Pivoting Method I, and it required $2^3 - 1 = 7$ pivot steps, thus verifying the theorem for $n = 3$. The theorem can also be verified to be true when $n = 2$. We now set up an induction hypothesis.

Induction Hypothesis: When applied on the LCP $(-e_{n-1}, \widetilde{M}(n-1))$, the Principal Pivoting Method I requires $2^{n-1} - 1$ pivot steps before termination.

We will now prove that the induction hypothesis implies that the Principal Pivoting Method I requires $2^n - 1$ pivot steps before termination when apllied on the LCP $(-e_n, \widetilde{M}(n))$ of order n.

When it is applied on the LCP in Tableau 6.5 the initial basic vector in Principal Pivoting Method I is $(w_1, \ldots, w_n)$. The entering variable into this initial complementary basic vector is z_n. Since $\widetilde{M}(n)$ is a P-matrix, by the results in Section 4.1, the method terminates when all the updated right hand side constants become nonnegative.

By the pivot row choice rule used in Principal Pivoting Method I, the question of using Row 1 in Tableau 6.5 as the pivot row does not arise until a complementary basic vector satisfying the property that the entries in Rows 2 to n of the updated right hand side constant vectors corresponding to it are all nonnegative, is reached. So until such a complementary basic vector is reached, the pivot steps choosen are exactly those that will be choosen in solving the principal subproblem of Tableau 6.5 in the variables $(w_2, \ldots, w_n)$; $(z_2, \ldots, z_n)$. This principal subproblem is actually the LCP $(-e_{n-1}, \widetilde{M}(n-1))$ of order $n-1$, with the exception that the variables in it are called $(w_2, \ldots, w_n)$; $(z_2, \ldots, z_n)$. By the induction hypothesis, to solve this principal subproblem, Principal Pivoting Method I takes $2^{n-1} - 1$ pivot steps. By the results discussed above $(z_2, w_3, \ldots, w_n)$ is the unique complementary feasible basic vector for this principal subproblem.

Hence when Principal Pivoting Method I is applied on Tableau 6.5, after $2^{n-1} - 1$ pivot steps it reaches the complementary basic vector $(w_1, z_2, , w_3, \ldots, w_n)$. The canonical tableau of Tableau 6.5 corresponding to the complementary basic vector $(w_1, z_2, w_3, \ldots, w_n)$ is given in Tableau 6.6.

Tableau 6.6 Canonical Tableau after $2^{n-1} - 1$ Pivot Steps are carried out, beginning with Tableau 6.5 in Principal Pivoting Method I.

Basic Variable	w_1	w_2	w_3	...	w_n	z_1	z_2	z_3	...	z_n	q
w_1	1	0	0	...	0	$\boxed{-1}$	0	0	...	0	-1
z_2	0	-1	0	...	0	2	1	0	...	0	1
w_3	0	-2	1	...	0	2	0	-1	...	0	1
$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
w_n	0	-2	0	...	1	2	0	-2	...	-1	1

Since the update right hand side constant in Row 1 is $-1 < 0$, the method now continues by making a single principal pivot step in position 1 in Tableau 6.6 (this replaces w_1 in the basic vector by z_1). The pivot element is inside a box. This leads to the canonical tableau in Tableau 6.7.

Tableau 6.7 Canonical Tableau after 2^{n-1} Pivot Steps are Carried Out, beginning with Tableau 6.5 in Principal Pivoting Method I.

Basic Variable	w_1	w_2	w_3	...	w_n	z_1	z_2	z_3	...	z_n	q
z_1	-1	0	0	...	0	1	0	0	...	0	1
z_2	2	-1	0	...	0	0	1	0	...	0	-1
w_3	2	-2	1	...	0	0	0	-1	...	0	-1
$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
w_n	2	-2	0	...	1	0	0	-2	...	-1	-1

Since some of the updated right hand side constants in Tableau 6.7 are still negative, the method continues. By the arguments mentioned above, when Principal Pivoting Method I is continued from Tableau 6.7, z_1 remains the basic variable in the first row until another complementary basic vector satisfying the property that the entries in Rows 2 to n in the updated right hand side constants vector corresponding to it are all nonnegative, is again reached. It can be verified that the principal subproblem obtained by eliminating Row 1 and the columns corresponding to the complementary pair of variables w_1, z_1 in Tableau 6.7 and interchanging the columns of the variables w_2 and z_2; is exactly the LCP $(-e_{n-1}, \widetilde{M}(n-1))$ with the exception that the variables in it are called $(z_2, w_3, \ldots, w_n)$; $(w_2, , z_3, \ldots, z_n)$. When Principal Pivoting Method I is continued from Tableau 6.7 the pivot steps obtained are exactly those that occur when

this principal subproblem is solved by Principal Pivoting Method I, until this principal subproblem is solved. Since this principal subproblem is the LCP $(-e_{n-1}, \widetilde{M}(n-1))$, by the induction hypothesis, this leads to an additional $\mathbf{2^{n-1}-1}$ pivot steps from Tableau 6.7. Since the variables in this principal subproblem are $(z_2, w_3, \ldots, w_n)$, $(w_2, z_3, \ldots, z_n)$, in that order, by the results mentioned earlier, $(w_2, w_3, \ldots, w_n)$ is the unique complementary feasible basic vector for this principal subproblem. So after continuing for an additional $\mathbf{2^{n-1}-1}$ pivot steps from Tableau 6.7, Principal Pivoting Method I reaches the complementary basic vector $(z_1, w_2, w_3, \ldots, w_n)$, which was verified to be a complementary feasible basic vector for the LCP in Tableau 6.5 and then the method terminates. So it took $\mathbf{2^{n-1}}$ pivot steps to reach Tableau 6.7 and an additional $\mathbf{2^{n-1}-1}$ pivot steps afterwards, before termination. Thus it requires a total of $\mathbf{2^{n-1}} + \mathbf{2^{n-1}-1} = \mathbf{2^n-1}$ pivot steps before termination, when applied on the LCP of order n in Tableau 6.5. Thus under the induction hypothesis, the statement of the theorem also holds for n. The statement of the theorem has already been verified for $n = 2, 3$. Hence, by induction, Theorem 6.4 is true for all $n \geqq 2$.

□

Exercise

6.1 Prove that the sequence of complementary basic vectors obtained when Principal Pivoting Method I is applied on the LCP in Tableau 6.5 is exactly the sequence $V_0, V_1, \ldots, V_{a(n)}$, obtained in the proof of Theorem 6.1. (Hint: Use an inductive argument as in the proof of Theorem 6.4).

So far, we have discussed the worst case computational complexity of complementary and principal pivot methods, which can handle a large class of LCPs. These results may not apply to other special algorithms for solving LCPs (q, M), in which the matrix M has special structure. An example of these is the algorithm of R. Chandrasekaran which can solve the LCP (q, M) when M is a Z-matrix (a square matrix $M = (m_{ij})$ is said to be a Z-matrix if $m_{ij} \leqq 0$ for all $i \neq j$) discussed in Section 8.1. This special algorithm for this special class of LCPs has been proved to terminate in at most n pivot steps.

The matrix $\widetilde{M}(n)$ used in the examples contructed above is lower triangular, it is a P-matrix, a nonnegative matrix, it is copositive plus and also PSD. So it has all the nice properties of matrices studied in LCP literature. In spite of it, complementary pivot methods take $\mathbf{2^n-1}$ or $\mathbf{2^n}$ pivot steps to solve the LCP of order n in the examples constructed above, all of which are associated with the matrix $\widetilde{M}(n)$.

We have shown that the computational requirements of the well known complementary and principal pivot methods exhibit an exponential growth rate in terms of the order of the LCP. Our analysis applies only to the worst case behavior of the methods on specially constructed simple problems. The performance of the algorithms on average practical problems using practical data may be quite different. The analysis

here is similar to the analysis of the worst case computational requirements of the simplex method for solving linear programs in Chapter 14 of [2.26].

The class of LCPs (q, M) where M is a PD and symmetric matrix is of particular interest because of the special structure of these problems, and also because they appear in many practical applications. It turns out that even when restricted to this special class of LCPs, the worst case computational requirements of complementary pivot methods exhibit an exponential growth rate in terms of the order of the LCP. See reference [6.3] of Y. Fathi and Exercises 6.2 to 6.5.

As mentioned in Section 2.8 the exponential growth of the worst case computational complexity as a function of the size of the problem does not imply that these algorithms are not useful for solving large scale practical problems. The exponential growth has been mathematically established on specially constructed problems with a certain pathological structure. This pathological structure does not seem to appear often in practical applications. As discussed in Section 2.8 and in Reference [2.36], the probabilistic average (or expected) computational complexity of some versions of the complementary pivot algorithm grows at most quadratically with n. Empirical computational tests seem to indicate that the number of pivot steps needed by these algorithms before termination grows linearly with n on an average.

6.5 Exercises

6.2 For $n \geqq 2$, let $M(n) = \left(\widetilde{M}(n)\right)\left(\widetilde{M}(n)\right)^T$.

$$M(n) = \begin{pmatrix} 1 & 2 & 2 & \dots & 2 & 2 \\ 2 & 5 & 6 & \dots & 6 & 6 \\ 2 & 6 & 9 & \dots & \vdots & \vdots \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 2 & 6 & \vdots & \dots & 1+4(n-2) & 2+4(n-2) \\ 2 & 6 & \vdots & \dots & 2+4(n-2) & 1+4(n-1) \end{pmatrix}$$

Prove that $M(n)$ is PD and symmetric. Solve the LCP $(-e_3, M(3))$ by Principal Pivoting Method I and verify that it takes $2^{\mathbf{3}} - 1 = 7$ pivot steps before termination. Solve The LCP $((-4, -7)^T, M(2))$ by the complementary pivot method and verify that it takes $2^{\mathbf{2}} = 4$ pivot steps before termination. Solve the parametric LCP $((4 - \gamma, 1 - \gamma)^T, M(2))$ by the parametric LCP algorithm and verify that it produces all the $2^{\mathbf{2}} = 4$ complementary basic vectors of this problem before solving the problem for all the values of the parameter γ.
(Y. Fathi [6.3])

6.3 Prove that Principal Pivoting Method I requires $2^{\mathbf{n}} - 1$ steps before termination when applied on the LCP $(-e_n, M(n))$, for any $n \geqq 2$.
(Y. Fathi [6.3])

6.4 Prove that there exists a column vector $q(n) \in \mathbf{R}^n$ (actually an uncountable number of such $q(n)$s exist) such that the straight line $\{x : x = q(n) - \gamma e_n, \gamma \text{ a real number}\}$ cuts across the interior of every one of the $2^{\mathbf{n}}$ complementary cones in the class $\mathcal{C}(M(n))$ for any $n \geqq 2$.
(Y. Fathi [6.3])

6.5 Prove that the parametric algorithm obtains all the $2^{\mathbf{n}}$ complementary basic vectors before termination, when applied to solve the LCP $(q(n) - \gamma e_n, M(n))$ for all γ for any $n \geqq 2$, where $q(n)$ is the column vector in $\mathbf{R}^n$ constructed in Exercise 6.4.
(Y. Fathi [6.3])

6.6 Prove that the complementary pivot method requires $2^{\mathbf{n}}$ pivot steps before termination when applied on the LCP $(q(n), M(n))$, for $n \geqq 2$, where $q(n)$ is the column vector in $\mathbf{R}^n$ constructed in Exercise 6.4.
(Y. Fathi [6.3])

6.7 Construct a class of LCPs with integer data, containing one problem of order n for each $n \geqq 2$, each associated with a PD matrix, such that the number of pivot steps required by Graves' principal pivoting method (Section 4.2) to solve the n^{th} problem in this class is an exponentially growing function of n.

6.8 Let $q(n) = (2^{\mathbf{n}} + 2, 2^{\mathbf{n}} + 4, \ldots, 2^{\mathbf{n}} + 2^{\mathbf{j}}, \ldots, 2^{\mathbf{n}} + 2^{\mathbf{n-1}}, -2^{\mathbf{n}})^T$ and

$$\overline{M}(n) = \begin{pmatrix} 1 & 2 & 2 & \ldots & 2 & -2 \\ 0 & 1 & 2 & \ldots & 2 & -2 \\ \vdots & 0 & 1 & \ldots & 2 & -2 \\ \vdots & \vdots & 0 & \ldots & \vdots & \vdots \\ \vdots & \vdots & 0 & \ldots & \vdots & \vdots \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 1 & -2 \\ 0 & 0 & 0 & \ldots & 0 & 1 \end{pmatrix}$$

Prove that the Dantzig-Cottle principal pivoting method of Section 4.3 requires $2^{\mathbf{n-1}}$ steps to solve the LCP $(q(n), \overline{M}(n))$.
(A. Gana [6.4])

6.9 Show that the variable dimension algorithm of Secton 2.6 requires $2^{\mathbf{n}} - 1$ steps to solve the LCP $\left(\tilde{q}(n), (\widetilde{M}(n))^T\right)$.
(A. Gana [6.4])

6.10 Define the matrix $M = (m_{ij})$ of order $n \times n$ by the following

$$\begin{aligned} m_{ii} &= \ \ 1 \quad \text{for all } i = 1 \text{ to } n \\ m_{ij} &= \ \ 2 \quad \text{if } j > i \text{ and } i + j \text{ is odd} \\ &= -1 \quad \text{if } j > i \text{ and } i + j \text{ is even} \\ &= -1 \quad \text{if } j < i \text{ and } i + j \text{ is odd} \\ &= \ \ 2 \quad \text{if } j < i \text{ and } i + j \text{ is even.} \end{aligned}$$

For example, the matrix M defined above, is the following for $n = 4$

$$M = \begin{pmatrix} 1 & 2 & -1 & 2 \\ -1 & 1 & 2 & -1 \\ 2 & -1 & 1 & 2 \\ -1 & 2 & -1 & 1 \end{pmatrix}.$$

Show that M is a P-matrix and a PSD matrix.

Let e be the column vector of all 1s in $\mathbf{R}^n$. Consider the LCP $(-e, M)$, where M is the matrix defined above. Show that the complementary feasible basic vector for this problem is

$$\begin{aligned} &(w_1, z_2, \ldots, z_n) \quad \text{if } n \text{ is even} \\ &(z_1, z_2, \ldots, z_n) \quad \text{if } n \text{ is odd.} \end{aligned}$$

Study the computational complexity of the various algorithms for solving LCPs discussed so far, on the LCP $(-e, M)$, where M is the matrix defined above.
(R. Chandrasekaran, J. S. Pang and R. Stone)

6.6 References

6.1 J. R. Birge and A. Gana, "Computational Complexity of Van der Heyden's Variable Dimension Algorithm and Dantzig-Cottle's Principal Pivoting Method for Solving LCPs", *Mathematical Programming,* 26 (1983) 316–325.

6.2 R. W. Cottle, "Observations on a Class of Nasty Linear Complementary Problems", *Discrete Applied Mathematics,* 2 (1980) 89–111.

6.3 Y. Fathi, "Computational Complexity of LCPs Associated with Positive Definite Symmetric Matrices", *Mathematical Programming*, 17 (1979) 335–344.

6.4 A. Gana, "Studies in the Complementary Problem", Ph.D. Dissertation, Department of Industrial and Operations Engineering, University of Michigan, Ann Arbor, Michigan (1982).

6.5 K. G. Murty, "Computational Complexity of Complementary Pivot Methods", *Mathematical Programming Study 7*, (1978) 61–73.

Chapter 7

NEAREST POINT PROBLEMS ON SIMPLICIAL CONES

Let $\mathbf{\Gamma} = \{B_{.1}, \ldots, B_{.n}\}$ be a given linearly independent set of column vectors in $\mathbf{R}^n$, and let $b \in \mathbf{R}^n$ be another given column vector. Let $B = (B_{.1} \vdots \ldots \vdots B_{\cdot n})$. For $x \in \text{Pos}(\mathbf{\Gamma})$, $\alpha = B^{-1}x \geqq 0$, is known as the **combination vector** corresponding to x. We consider the problem of finding the nearest point (in terms of the usual Euclidean distance) in the simplicial cone $\text{Pos}(\mathbf{\Gamma})$ to b. This problem will be denoted by the symbol $[\mathbf{\Gamma}; b]$ or $[B; b]$, and will be called a **nearest point problem** of order n. The optimum solution of this problem is unique, and if $b \notin \text{Pos}(\mathbf{\Gamma})$ the solution lies on the boundary of $\text{Pos}(\mathbf{\Gamma})$. If this point is x^*, then $\alpha^* = B^{-1}x^*$ is known as the **optimum combination vector** for $[\mathbf{\Gamma}; b]$. This problem is equivalent to the quadratic program: Minimize $(b - B\alpha)^T(b - B\alpha)$ over $\alpha = (\alpha_1, \ldots, \alpha_n)^T \geqq 0$. This is the quadratic program: Minimize $-b^TB\alpha + \frac{1}{2}\alpha^T(B^TB)\alpha$, subject to $\alpha = (\alpha_1, \ldots, \alpha_n)^T \geqq 0$. The solution of this can be obtained by solving the following LCP :

$$\begin{aligned} &u - (B^TB)\alpha = -B^Tb \\ &u \geqq 0, \quad \alpha \geqq 0 \\ &u^T\alpha = 0 \end{aligned}$$

where $u = (u_1, \ldots, u_n)^T$ is a column vector of variables in $\mathbf{R}^n$. Let $D = B^TB$. Since B is nonsingular, D is positive definite. This LCP has a unique complementary solution, and if this solution is (u^*, α^*), then α^* is the optimum solution for the quadratic program, and hence the optimum combination vector for the nearest point problem $[B; b]$. Also consider the following LCP

$$\begin{aligned} &w - Mz = q \\ &w \geqq 0,\ z \geqq 0 \\ &w^Tz = 0 \end{aligned} \tag{7.1}$$

where M is a positive definite symmetric matrix of order n. Let F be a nonsingular matrix such that $F^TF = M$ (for example, the transpose of the Cholesky factor of M). Now using earlier results, we conclude that if (w^*, z^*) is the unique solution of (7.1), then z^* is the optimum combination vector for the nearest point problem $[F; -(F^{-1})^Tq]$. Conversely if z^* is the optimum combination vector for the nearest point problem $[F; -(F^{-1})^Tq]$, then $(w^* = Mz^* + q, z^*)$ is the unique solution of (7.1). This clearly establishes that corresponding to each nearest point problem, there is an equivalent LCP associated with a positive definite symmetric matrix and vice versa. This equivalence relationship between the two problems will be used here to develop an algorithm for solving them. In the sequel (q, M) denotes the LCP (7.1) where M is a positive definite symmetric matrix of order n. B denotes a square matrix of order n satisfying $B^TB = M$ (as mentioned earlier, B could be chosen as the Cholesky factor of M). If we are given the LCP (7.1) to solve, we will choose B^T to be the Cholesky factor of M, unless some other matrix satisfying $B^TB = M$ is available, and $b = -(B^{-1})^Tq$, and $\mathbf{\Gamma} = \{B_{.1}, \ldots, B_{.n}\}$. For solving either the nearest point problem $[\mathbf{\Gamma}; b]$ or the LCP (q, M), the algorithm discussed here based on the results in [3.51,7.2] of K. G. Murty and Y. Fathi, operates on both of them (it carries out some geometric work on the nearest point problem, and some algebraic work on the LCP).

Example 7.1

Let

$$q = \begin{pmatrix} 14 \\ -11 \\ -7 \end{pmatrix}, \quad M = \begin{pmatrix} 3 & -2 & -1 \\ -2 & 2 & 1 \\ -1 & 1 & 1 \end{pmatrix}, B = \begin{pmatrix} 1 & 0 & 0 \\ 1 & -1 & 0 \\ -1 & 1 & 1 \end{pmatrix}, \quad b = \begin{pmatrix} -3 \\ -4 \\ 7 \end{pmatrix}.$$

The LCP (q, M) is

w_1	w_2	w_3	z_1	z_2	z_3	q
1	0	0	−3	2	1	14
0	1	0	2	−2	−1	−11
0	0	1	1	−1	−1	−7

$w_j, z_j \geqq 0$, and $w_j z_j = 0$ for all j

It can be verified that $B^TB = M$ and $b = -(B^{-1})^Tq$. So, the above LCP is equivalent to the problem of finding the nearest point in $\text{Pos}(B)$ to b.

It can be verified that the solution of the LCP (q, M) is $(w_1, w_2, w_3; z_1, z_2, z_3) = (3, 0, 0; 0, 4, 3)$. This implies that the vector $\alpha^* = (0, 4, 3)^T$ is the optimum combination vector for the nearest point problem $[B; b]$; that is, $4B_{.2} + 3B_{.3} = (0, -4, 7)^T$ is the nearest point in $\text{Pos}(B)$ to b. Conversely, given that $\bar{x} = (0, -4, 7)^T$ is the nearest point in $\text{Pos}(B)$ to b, we get $z^* = B^{-1}\bar{x} = (0, 4, 3)^T$, and $w^* = Mz^* + q = (3, 0, 0)^T$, and (w^*, z^*) is the solution of the LCP (q, M).

Some Results

Let $\mathbf{S} = \{B_{.j_1}, \ldots, B_{.j_r}\} \subset \boldsymbol{\Gamma}$. Define

$$\begin{aligned}
\mathbf{I}(\mathbf{S}) &= \text{ Index set of } \mathbf{S} = \{j_1, \ldots, j_r\} \\
\overline{\mathbf{I}(\mathbf{S})} &= \{1, \ldots, n\} \setminus \mathbf{I}(\mathbf{S}) \\
\mathbf{H}(\mathbf{S}) &= \left\{ y : y = \sum_{j \in \mathbf{I}(\mathbf{S})} \gamma_j B_{.j}; \gamma_j \text{ real number for all } j \in \mathbf{I}(\mathbf{S}) \right\} \\
B(\mathbf{S}) &= \text{The } n \text{ by } r \text{ matrix whose columns are } B_{.j_1}, \ldots, B_{.j_r} \\
w(\mathbf{S}) &= (w_{j_1}, \ldots, w_{j_r})^T \\
z(\mathbf{S}) &= (z_{j_1}, \ldots, z_{j_r})^T \\
q(\mathbf{S}) &= (q_{j_1}, \ldots, q_{j_r})^T \\
M(\mathbf{S}) &= B(\mathbf{S})^T B(\mathbf{S}), \text{ the principal submatrix of } M \text{ corresponding to } \mathbf{I}(\mathbf{S}) \ ,
\end{aligned}$$

$\mathbf{H}(\mathbf{S})$ as defined above is the linear hull of $\mathbf{S}$, it is the subspace of $\mathbf{R}^n$ spanned by the column vectors in $\mathbf{S}$. If $\mathbf{S} = \emptyset$, define $\mathbf{H}(\mathbf{S}) = \text{ Pos}(\mathbf{S}) = \{0\}$. For any $\mathbf{S} \subset \boldsymbol{\Gamma}$, Pos($\mathbf{S}$) is a face of Pos($\boldsymbol{\Gamma}$). The problem of finding the nearest point in Pos($\mathbf{S}$) to b (in terms of the usual Euclidean distance) will be denoted by $[\mathbf{S}; b]$. If $\mathbf{S} \neq \emptyset$, the nearest point in $\mathbf{H}(\mathbf{S})$ to b is denoted by $b(\mathbf{S})$, and this point is known as the **projection** or the **orthogonal projection** of b in $\mathbf{H}(\mathbf{S})$.

Theorem 7.1 *Let $\mathbf{S} \subset \boldsymbol{\Gamma}$ and $\mathbf{S} \neq \emptyset$. Then $b(\mathbf{S}) = B(\mathbf{S})\big(B(\mathbf{S})^T B(\mathbf{S})\big)^{-1} B(\mathbf{S})^T b$.*

Proof. Let $\mathbf{S} = \{B_{.j_1}, \ldots, B_{.j_r}\}$ and let $\gamma = (\gamma_1, \ldots, \gamma_r)^T$. The problem of finding the projection of b in $\mathbf{H}(\mathbf{S})$ is the unconstrained minimization problem: Minimize $(b - B(\mathbf{S})\gamma)^T (b - B(\mathbf{S})\gamma)$: $\gamma \in \mathbf{R}^r$, and the optimum solution of this unconstrained minimization problem is $\overline{\gamma} = \big(B(\mathbf{S})^T B(\mathbf{S})\big)^{-1} (B(\mathbf{S}))^T b$. Hence, $b(\mathbf{S}) = B(\mathbf{S})\overline{\gamma} = B(\mathbf{S})\big(B(\mathbf{S})^T B(\mathbf{S})\big)^{-1} (B(\mathbf{S}))^T b$.

□

Example 7.2

Let B be the matrix defined in Example 7.1, and b the vector from the same example. So

$$B = \begin{pmatrix} 1 & 0 & 0 \\ 1 & -1 & 0 \\ -1 & 1 & 1 \end{pmatrix} \ , \quad b = \begin{pmatrix} -3 \\ -4 \\ 7 \end{pmatrix}$$

Let $\mathbf{S} = \{B_{.1}, B_{.3}\}$. So in this case $\mathbf{I}(\mathbf{S}) =$ index set of $\mathbf{S} = \{1, 3\}$. So $\mathbf{H}(\mathbf{S})$ is the subspace $\{\gamma_1(1, 1, -1)^T + \gamma_2(0, 0, 1)^T : \gamma_1, \gamma_2 \text{ real numbers}\}$ of $\mathbf{R}^3$. The matrix $B(\mathbf{S})$ here is

$$B(\mathbf{S}) = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ -1 & 1 \end{pmatrix} \ , \quad M(\mathbf{S}) = \big(B(\mathbf{S})\big)^T B(\mathbf{S}) = \begin{pmatrix} 3 & -1 \\ -1 & 1 \end{pmatrix}$$

The projection $b(\mathbf{S})$ here can be verified to be $b(\mathbf{S}) = B(\mathbf{S}) \begin{pmatrix} -\frac{7}{2} \\ \\ \frac{7}{2} \end{pmatrix} = \left(-\frac{7}{2}, -\frac{7}{2}, 7\right)^T$. Since $b(\mathbf{S}) = B(-\frac{7}{2}, 0, \frac{7}{2})^T$, it is not in the cone $\text{Pos}(B)$.

Theorem 7.2 *For $\mathbf{S} \subset \boldsymbol{\Gamma}$, the nearest point in Pos($\mathbf{S}$) to b is the same as the nearest point in Pos($\mathbf{S}$) to $b(\mathbf{S})$.*

Proof. The case $\mathbf{S} = \emptyset$ is trivially verified to be true. So assume $\mathbf{S} \neq \emptyset$. For $x \in \mathbf{H}(\mathbf{S})$ by Pythagoras theorem $||b - x||^2 = ||b - b(\mathbf{S})||^2 + ||b(\mathbf{S}) - x||^2$. Since $\text{Pos}(\mathbf{S}) \subset \mathbf{H}(\mathbf{S})$, this equality obviously holds for all $x \in \text{Pos}(\mathbf{S})$. Hence the theorem follows.

□

Theorem 7.3 *Let $\mathbf{S} \subset \boldsymbol{\Gamma}$, $\mathbf{S} \neq \emptyset$. The optimum solution of $[\mathbf{S}; b]$ is in the relative interior of Pos($\mathbf{S}$) if and only if $b(\mathbf{S})$ is in the relative interior of Pos($\mathbf{S}$).*

Proof. $b(\mathbf{S})$ is in the relative interior of $\text{Pos}(\mathbf{S})$ if and only if $b(\mathbf{S}) = B(\mathbf{S})\overline{\gamma}$, where $\overline{\gamma} > 0$. As long as $b(\mathbf{S}) \in \text{Pos}(\mathbf{S})$, $b(\mathbf{S})$ is the optimum solution of $[\mathbf{S}; b]$, and hence in this case the statement of the theorem is true. If $b(\mathbf{S}) \notin \text{Pos}(\mathbf{S})$, by Theorem 7.2, the optimum solutions of $[\mathbf{S}; b]$ and $[\mathbf{S}; b(\mathbf{S})]$ are the same. $[\mathbf{S}; b(\mathbf{S})]$ is the nearest point problem in the subspace $\mathbf{H}(\mathbf{S})$, whose order is the same as the dimension of $\mathbf{H}(\mathbf{S})$, and hence in this case the optimum solution of $[\mathbf{S}; b(\mathbf{S})]$ lies on the relative boundary of $\text{Pos}(\mathbf{S})$.

□

Definition — Projection Face
Let $\mathbf{S} \subset \boldsymbol{\Gamma}$. $\text{Pos}(\mathbf{S})$ is a face of $\text{Pos}(\boldsymbol{\Gamma})$ of dimension $|\mathbf{S}|$. $\text{Pos}(\mathbf{S})$ is said to be a **Projection face** of $\text{Pos}(\boldsymbol{\Gamma})$, if $b(\mathbf{S}) \in \text{Pos}(\mathbf{S})$.

Example 7.3

Let B, b be as in in Example 7.2. As computed there, the projection of b in the linear hull of $\{B_{.1}, B_{.3}\}$ is not in the face $\text{Pos}\{B_{.1}, B_{.3}\}$, since it is $-\frac{7}{2}B_{.1} + \frac{7}{2}B_{.3}$, not a nonnegative combination of $B_{.1}$, $B_{.3}$. So, the face $\text{Pos}\{B_{.1}, B_{.3}\}$ is not a projection face.

On the other hand, consider the face $\text{Pos}\{B_{.2}, B_{.3}\}$. The projection of b in the linear hull of $\{B_{.2}, B_{.3}\}$ can be verified to be $4B_{.2} + 3B_{.3} = (0, -4, 7)^T$ which is in $\text{Pos}\{B_{.2}, B_{.3}\}$. So $\text{Pos}\{B_{.2}, B_{.3}\}$ is a projection face of $\text{Pos}(B)$.

Theorem 7.4 *Let $x^* = B\alpha^*$ be the optimum solution of $[\boldsymbol{\Gamma}; b]$. Let $\mathbf{I}(\mathbf{S}) = \{j_1, \ldots, j_r\} = \{j : j \text{ such that } \alpha_j^* > 0\}$, and $\mathbf{S} = \{B_{.j} : j \in \mathbf{I}(\mathbf{S})\}$. Then Pos($\mathbf{S}$) is a projection face of Pos($\boldsymbol{\Gamma}$).*

Proof. Obviously $x^* \in \text{Pos}(\mathbf{S})$. Since x^* is the nearest point in $\text{Pos}(\boldsymbol{\Gamma})$ to b, and since $\text{Pos}(\mathbf{S}) \subset \text{Pos}(\boldsymbol{\Gamma})$, clearly x^* is the nearest point in $\text{Pos}(\mathbf{S})$ to b. However, by the definition of $\mathbf{S}$, x^* is in the relative interior of $\text{Pos}(\mathbf{S})$. Hence, by Theorem 7.3, x^* must be the projection of b in $\mathbf{H}(\mathbf{S})$. Since $x^* \in \text{Pos}(\mathbf{S})$, this implies that $\text{Pos}(\mathbf{S})$ is a projection face of $\text{Pos}(\boldsymbol{\Gamma})$.

□

Exercises

7.1 Prove that the problem of finding the nearest point in the face Pos($\mathbf{S}$) of Pos($\mathbf{\Gamma}$) to b or $b(\mathbf{S})$, is equivalent to the principal subproblem of the LCP (7.1) in the variables $w(\mathbf{S})$, $z(\mathbf{S})$. Also show that if $(\hat{w}(\mathbf{S}) = q(\mathbf{S}) + M(\mathbf{S})\hat{z}(\mathbf{S}), \hat{z}(\mathbf{S}))$ is the solution of this principal subproblem, then $B(\mathbf{S})\hat{z}(\mathbf{S})$ is the nearest point in Pos($\mathbf{S}$) to b or $b(\mathbf{S})$; and conversely. Also prove that the face Pos($\mathbf{S}$) of Pos($\mathbf{\Gamma}$) is a projection face iff $z(\mathbf{S})$ is a complementary feasible basic vector for this principal subproblem.

7.2 If $\mathbf{S} \subset \mathbf{\Gamma}$ is such that Pos($\mathbf{S}$) is a projection face of Pos($\mathbf{\Gamma}$), prove that $b(\mathbf{S})$ is the nearest point in Pos($\mathbf{\Gamma}$) to b iff $(w(\mathbf{\Gamma}\backslash\mathbf{S}), z(\mathbf{S}))$ is a complementary feasible vector for (7.1).

Definitions and Notation

Let $\mathbf{K}_j$ denote the facet $\text{Pos}(B_{.1}, \ldots, B_{.j-1}, B_{.j+1}, \ldots, B_{.n})$ of Pos($\mathbf{\Gamma}$) for $j = 1$ to n. Let $x = \alpha_1 B_{.1} + \ldots + \alpha_n B_{.n} \in \text{Pos}(\mathbf{\Gamma})$. It follows that $\alpha_j = 0$ if and only if $x \in \mathbf{K}_j$, and $\alpha_j > 0$ if and only if $x \notin \mathbf{K}_j$, for all $j = 1$ to n. Given the two points $b \in \mathbf{R}^n$ and $\bar{x} \in \mathbf{R}^n$ such that $b \neq \bar{x}$, let the open ball $\mathbf{B}(b; \bar{x}) = \{x : ||b - x|| < ||b - \bar{x}||\}$. Consider the hyperplane $\mathbf{T}(b; \bar{x}) = \{x : (x - \bar{x})^T (b - \bar{x}) = 0\}$. The open half space $\{x : (x - \bar{x})^T (b - \bar{x}) > 0\}$ is called the **near side** of $\mathbf{T}(b; \bar{x})$, while the closed half space $\{x : (x - \bar{x})^T (b - \bar{x}) \leqq 0\}$ is called the **far side** of $\mathbf{T}(b; \bar{x})$. If the point $\bar{x}$ is chosen such that $0 \in \mathbf{T}(b; \bar{x})$, then $\bar{x}^T (B - \bar{x}) = 0$ and therefore for such $\bar{x}$ we have: $\mathbf{T}(b; \bar{x}) = \{x : x^T (b - \bar{x}) = 0\}$, near side of $\mathbf{T}(b; \bar{x}) = \{x : x^T (b - \bar{x}) > 0\}$, far side of $\mathbf{T}(b; \bar{x}) = \{x : x^T (b - \bar{x}) \leqq 0\}$. For points $\bar{x}$ satisfying $0 \in \mathbf{T}(b; \bar{x})$, we define the set $\mathbf{N}(\bar{x})$ by

$$\mathbf{N}(\bar{x}) = \{j : j \text{ such that } B_{.j}^T (b - \bar{x}) > 0\} \ .$$

So $\mathbf{N}(\bar{x})$ is the set of subscripts of the column vectors in $\mathbf{\Gamma}$ which are on the near side of $\mathbf{T}(b, \bar{x})$.

Let $V^j = 0$ if $b^T B_{.j} \leqq 0$, or $= \frac{B_{.j}(b^T B_{.j})}{||B_{.j}||^2}$ if $b^T B_{.j} > 0$. V_j is the nearest point on the ray of $B_{.j}$ to b, for all $j = 1$ to n. Also let l be such that $||V^l - b|| = \min\{||V^j - b|| : j = 1 \text{ to } n\}$. Break ties for the minimum in this equation arbitrarily. If $V^l \neq 0$, it is the orthogonal projection of b on the linear hull of $B_{.l}$.

Example 7.4

Let B, b be as given in Example 7.2. That is,

$$B = \begin{pmatrix} 1 & 0 & 0 \\ 1 & -1 & 0 \\ -1 & 1 & 1 \end{pmatrix}, \qquad b = \begin{pmatrix} -3 \\ -4 \\ 7 \end{pmatrix}$$

So $b^T B_{.1} = -14 < 0$, $b^T B_{.2} = 11 > 0$, $b^T B_{.3} = 7 > 0$. So if V^j is the nearest point to b on the ray of $B_{.j}$, we have $V^1 = 0$, $V^2 = (0, -\frac{11}{2}, \frac{11}{2})^T$, $V^3 = (0, 0, 7)^T$. Also, we verify that the nearest point among V^1, V^2, V^3 to b is V^2, so l as defined above, is 2 in this problem.

If we take $\bar{x} = V^2$, since $\bar{x}$ is the nearest point on the ray of $B_{.2}$ to b, the ray of $B_{.2}$ is a tangent line to the ball $\mathbf{B}(b; \bar{x})$ at its boundary point $\bar{x}$. See Figure 7.1. So the tangent plane $\mathbf{T}(b; \bar{x})$ to $\mathbf{B}(b; \bar{x})$ at its boundary point $\bar{x}$ contains the ray of $B_{.2}$. So in this example $\mathbf{N}(\bar{x}) = \{j : j \text{ such that } (b - \bar{x})^T B_{.j} > 0\} = \{3\}$. So the vector $B_{.3}$ is on the near side of $\mathbf{T}(b; \bar{x})$, and the vector $B_{.1}$ is on the far side of $\mathbf{T}(b; \bar{x})$, in this example.

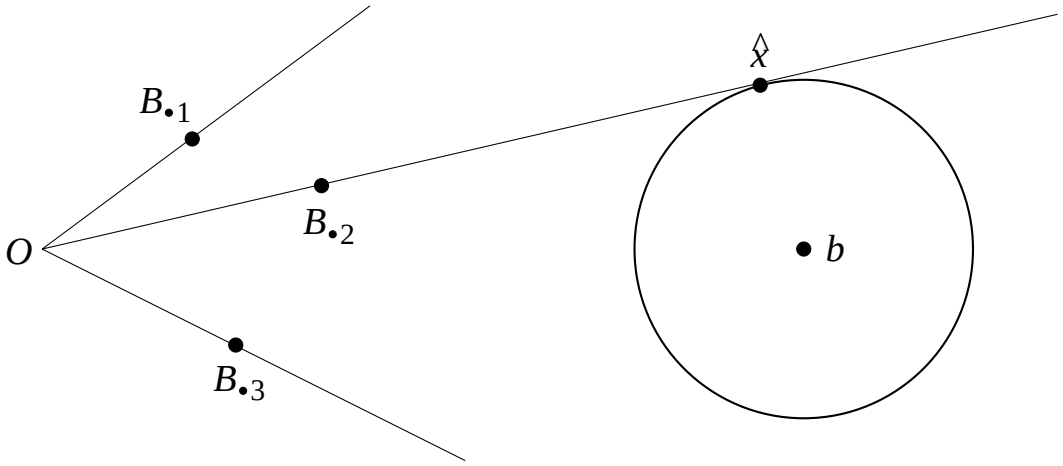


Figure 7.1

Theorem 7.5 *If $V^l = 0$, the nearest point in Pos($\mathbf{\Gamma}$) to b is 0.*

Proof. In this Case $b^T B_{.j} \leqq 0$ for all $j = 1$ to n. Hence the hyperplane $\{x : b^T x = 0\}$ for which the ray of b is the normal at 0, separates b and Pos($\mathbf{\Gamma}$). So 0 is the nearest point in Pos($\mathbf{\Gamma}$) to b.

□

Example 7.5

Let

$$B = \begin{pmatrix} 1 & 0 & 0 \\ 1 & -1 & 0 \\ -1 & 1 & 1 \end{pmatrix}, \qquad b = \begin{pmatrix} -1 \\ 0 \\ -1 \end{pmatrix}.$$

We have $b^T B_{.j} = 0, -1, -1$ respectively for $j = 1, 2, 3$. So, the nearest point on the ray of $B_{.j}$ is $V^j = 0$ for all $j = 1, 2, 3$. Hence in this case 0 is the nearest point in Pos(B) to b.

Thus 0 is the nearest point to b in Pos(B) iff $b^T B_{.j} \leqq 0$ for all $j = 1$ to n. So, in the sequel, we will assume that $b^T B_{.j} > 0$ for at least one j, and under this condition, V^l as defined above is not zero.

Theorem 7.6 *A point $\bar{x} \in$ Pos($\mathbf{\Gamma}$) is the nearest point in Pos($\mathbf{\Gamma}$) to b if and only if*

$$\begin{array}{l} 0 \in \mathbf{T}(b; \bar{x}) \text{ and} \\ (b - \bar{x})^T B_{.j} \leqq 0, \text{ for all } j = 1 \text{ to } n. \end{array} \tag{7.2}$$

Proof. Suppose $\bar{x}$ is the nearest point in Pos($\mathbf{\Gamma}$) to b. So, $\bar{x}$ is the orthogonal projection of b on the full line generated by $\bar{x}$, and hence $0 \in \mathbf{T}(b; \bar{x})$. Also, the hypothesis implies that the hyperplane $\mathbf{T}(b; \bar{x})$ strictly separates $\mathbf{B}(b; \bar{x})$ and Pos($\mathbf{\Gamma}$). So $(b - \bar{x})^T B_{.j} \leqq 0$ for all $j = 1$ to n.

Conversely suppose $\bar{x} \in$ Pos($\mathbf{\Gamma}$) satisfies 7.2. These conditions imply that $T(b; \bar{x})$ is the tangent hyperplane to the closure of $\mathbf{B}(b; \bar{x})$ at its boundary point $\bar{x}$, and that $\mathbf{T}(b; \bar{x})$ separates the closure of $\mathbf{B}(b; \bar{x})$ and Pos($\mathbf{\Gamma}$). So, under these conditions, $\bar{x}$ is the nearest point in Pos($\mathbf{\Gamma}$) to b.

□

Example 7.6

Let B, b be as given in Example 7.4, that is

$$B = \begin{pmatrix} 1 & 0 & 0 \\ 1 & -1 & 0 \\ -1 & 1 & 1 \end{pmatrix}, \qquad b = \begin{pmatrix} -3 \\ -4 \\ 7 \end{pmatrix} .$$

If $\bar{x} = V^2 = (0, -\frac{11}{2}, \frac{11}{2})^T$, we verified as in Example 7.4 that $(b - \bar{x})^T B_{.3} = (\frac{3}{2}) > 0$, and hence $\bar{x}$ is not the nearest point in Pos(B) to b.

Let $\hat{x} = (0, -4, 7)^T$, the orthogonal projection of b in the linear hull of $\{B_{.2}, B_{.3}\}$, which is the nearest point in the face Pos$\{B_{.2}, B_{.3}\}$ of Pos(B) to b, obtained in Example 7.3. Since $\hat{x}$ is the orthogonal projection of b in a subspace, the tangent plane $\mathbf{T}(b, \bar{x})$ contains this subspace (in this case $\mathbf{T}(b, \bar{x})$ is the linear hull of $\{B_{.2}, B_{.3}\}$ itself) and hence the origin 0. Also, it can be verified that $(b - \hat{x})^T B_{.j} = -3, 0, 0 \leq 0$, for $j = 1, 2, 3$. So $\mathbf{N}(\hat{x}) = \emptyset$ and $\hat{x}$ is the nearest point in Pos(B) to b in this example. See Figure 7.2.

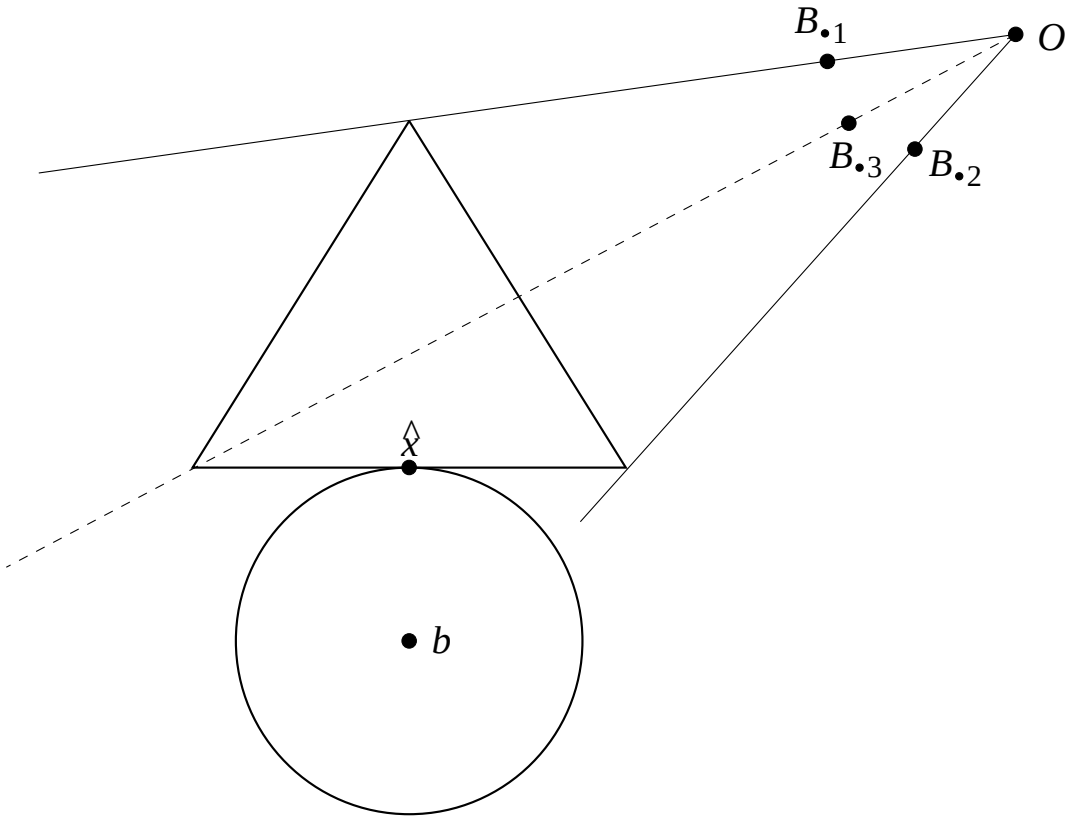


Figure 7.2 $\hat{x}$ is the nearest point in $\text{Pos}(B)$ to b.

Let α^* be the unknown optimum combination vector for $[\mathbf{\Gamma}; b]$. Let $\mathbf{J} = \{j : \alpha_j^* > 0\}$. $\mathbf{J}$ is called the set of **critical indices** for the LCP (q, M) and for the corresponding nearest point problem $[\mathbf{\Gamma}; b]$. It is clear that $\mathbf{J}$ is also the set of all j such that z_j is strictly positive in the unique solution of the LCP (q, M). Notice that if (w, z) is the unique solution of the LCP (q, M), then $w_j = 0$ for all $j \in \mathbf{J}$ and $z_j = 0$ for all $j \notin \mathbf{J}$, or equivalently if $y_j = z_j$ for all $j \in \mathbf{J}$, w_j for all $j \notin \mathbf{J}$, then $(y_1, \ldots, y_n)$ is a complementary feasible basic vector for this LCP. So if the set $\mathbf{J}$ can be found, the basic solution of (7.1) corresponding to the basic vector $(y_1, \ldots, y_n)$ defined above is the unique solution of this problem. Also by earlier results, the solution to the nearest point problem $[\mathbf{\Gamma}; b]$ is the orthogonal projection of b on the linear hull of $\{B_{.j} : j \in \mathbf{J}\}$. Hence if $\mathbf{J}$ is known, the solution of the LCP (q, M) and correspondingly the solution to the associated nearest point problem $[\mathbf{\Gamma}; b]$ can be easily found.

Even if a single critical index is known, this information can be used to reduce (q, M) to an LCP of order $n - 1$ as shown in the following theorem.

Theorem 7.7 *If a single critical index is known, (q, M) can be reduced to an LCP of order $n-1$.*

Proof. Without loss of generality suppose we know that 1 is a critical index. Then perform a single principal pivot step in (7.1) in position 1. Suppose this leads to

w_1	w_2	...	w_n	z_1	z_2	...	z_n	
$-\overline{m}_{11}$	0	...	0	1	$-\overline{m}_{12}$	...	$-\overline{m}_{1n}$	$\bar{q}_1$
$-\overline{m}_{21}$	1	...	0	0	$-\overline{m}_{22}$	...	$-\overline{m}_{2n}$	$\bar{q}_2$
$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
$-\overline{m}_{n1}$	0	...	1	0	$-\overline{m}_{n2}$	...	$-\overline{m}_{nn}$	$\bar{q}_n$

Let $\overline{\mathcal{M}} = (\overline{m}_{ij} : 2 \leqq i,\ j \leqq n)$ be the matrix of order $n-1$, and $\bar{q} = (\bar{q}_2, \ldots, \bar{q}_n)^T$, from the above Tableau. Eliminating the columns of w_1, z_1, and the first row from it leads to the principal subproblem in variables $\omega = (w_2, \ldots, w_n)$ and $\xi = (z_2, \ldots, z_n)$, which is an LCP of order $n-1$, denoted by $(\bar{q}, \overline{\mathcal{M}})$. Since M is positive definite and symmetric, so is $\overline{\mathcal{M}}$. If $(y_2, \ldots, y_n)$, where $y_j \in \{w_j, z_j\}$, is a complementary feasible basic vector for $(\bar{q}, \overline{\mathcal{M}})$, then, since $1 \in \mathbf{J}$, $(z_1, y_2, \ldots, y_n)$ is a complementary feasible basic vector for the original (q, M). Thus to solve (q, M), if we know that $1 \in \mathbf{J}$, it is enough if we solve the principal subproblem $(\bar{q}, \overline{\mathcal{M}})$ of order $n-1$. Therefore the fact that $1 \in \mathbf{J}$ has made it possible for us to reduce the LCP (q, M) of order n, into $(\bar{q}, \overline{\mathcal{M}})$ of order $n-1$.

□

We can also argue geometrically that the knowledge of a critical index reduces the dimensionality of the nearest point problem. If 1 is a critical index, then the nearest point to b in $\text{Pos}(\mathbf{\Gamma})$ is also the nearest point to b in $\text{Pos}(\mathbf{\Gamma} \cup \{-B_{.1}\})$. Define $\bar{b} = b - \frac{B_{.1}(b^T B_{.1})}{||B_{.1}||^2}$, $\bar{B}_{.j} = B_{.j} - \frac{B_{.1}(B_{.1})^T B_{.j}}{||B_{.1}||^2}$, for $j = 2, \ldots, n$. Let $\overline{\mathbf{\Gamma}} = \{\bar{B}_{.2}, \ldots, \bar{B}_{.n}\}$. For $2 \leqq j \leqq n$, $\bar{B}_{.j}$ is orthogonal to $B_{.1}$ and the cone $\text{Pos}(\mathbf{\Gamma} \cup \{-B_{.1}\})$ is the direct sum of the full line generated by $B_{.1}$ and the simplicial cone $\text{Pos}(\overline{\mathbf{\Gamma}})$. Solving $[\overline{\mathbf{\Gamma}}; \bar{b}]$ is an $(n-1)$ dimensional nearest point problem. If $\bar{x}^*$ is its solution, as embedded in $\mathbf{R}^n$, then $x^* = \bar{x}^* + \frac{B_{.1}(b^T B_{.1})}{||B_{.1}||^2}$ solves $[\mathbf{\Gamma}; b]$.

We will develop an algorithm for finding a critical index. When it is obtained, we can reduce (q, M) into a linear complementarity problem of lower order and apply the same approach on it.

Example 7.7

Consider the LCP (q, M) discussed in Example 7.1. In Example 7.9 we will establish the fact that 3 is a critical index for this LCP. Performing a principal pivot step in

position 3 in this LCP leads to the following :

w_1	w_2	w_3	z_1	z_2	z_3	
1	0	1	−2	1	0	7
0	1	−1	1	−1	0	−4
0	0	−1	−1	1	1	7

$w_j, z_j \geqq 0$ for all j. $w_j z_j = 0$ for all j

Since 3 is a critical index, we eliminate w_3, z_3 and the last row from the problem, leading to the principal subproblem

w_1	w_2	z_1	z_2	
1	0	−2	1	7
0	1	1	−1	−4

$w_j, z_j \geqq 0$ for all j. $w_j z_j = 0$ for all j

It can be verified that (w_1, z_2) is a complementary feasible basic vector for this principal subproblem. So, (w_1, z_2, z_3) is a complementary feasible basic vector for the original LCP (q, M).

Theorem 7.8 *Given $0 \neq \bar{x} \in Pos(\mathbf{\Gamma})$ satisfying $0 \in \mathbf{T}(b; \bar{x})$, if for some $i \in \{1, \ldots, n\}$, we have*

(i) $(b - \bar{x})^T B_{.i} > 0$, and either

(ii) $||\bar{x} - b|| \leqq ||V^i - b||$ and $\{\bar{x}, B_{.i}\}$ is linearly independent, or

(ii)' $b^T B_{.i} \leqq 0$;

then, the projection of b onto the linear hull of $\{\bar{x}, B_{.i}\}$ is in the relative interior of $Pos\{\bar{x}, B_{.i}\}$.

Proof. Since $\bar{x}$ is the closest point in $\mathbf{T}(b; \bar{x})$ to b and since $0 \in \mathbf{T}(b; \bar{x})$, $\bar{x}$ is the closest point on the ray of $\bar{x}$ to b.

If (ii)' holds, then $V^i = 0$ and hence in this case we have $||\bar{x} - b|| < ||V^i - b||$, and clearly $\{\bar{x}, B_{.i}\}$ is linearly independent. So under these conditions (ii)' implies (ii).

By linear independence, $\text{Pos}\{\bar{x}, B_{.i}\}$ is a two dimensional simplicial cone. Let p be the closest point in $\text{Pos}\{\bar{x}, B_{.i}\}$ to b. By (i), $B_{.i}$ is on the near side of $\mathbf{T}(b; \bar{x})$, and hence $\mathbf{B}(b; \bar{x}) \cap \text{Pos}\{\bar{x}, B_{.i}\} \neq \emptyset$. This implies that p is closer than $\bar{x}$ to b; and by (ii), p must be closer than V^i to b. So p is not contained on the rays of $\bar{x}$ or $B_{.i}$, and hence p must be in the relative interior of $\text{Pos}\{\bar{x}, B_{.i}\}$.

□

Theorem 7.9 *Let $\emptyset \neq \mathbf{S} \subset \overline{\mathbf{\Gamma}}$ be such that $\bar{x} = b(\mathbf{S}) \in Pos(\mathbf{S})$. Then $0 \in \mathbf{T}(b; \bar{x})$. Also, in this case if $\mathbf{N}(\bar{x}) \cap \overline{\mathbf{I}(\mathbf{S})} = \emptyset$, then $\mathbf{N}(\bar{x}) = \emptyset$, and $\bar{x}$ is the nearest point in $Pos(\mathbf{\Gamma})$ to b.*

Proof. Under the hypothesis $\mathbf{T}(b;\bar{x})$ contains $\mathbf{H}(\mathbf{S})$ and hence $0 \in \mathbf{T}(b;\bar{x})$. Also, by the properties of orthogonal projection, the line joining b and $\bar{x}$ is orthogonal to $\mathbf{H}(\mathbf{S})$, and hence $(b-\bar{x})^T B_{.j} = 0$ for all $j \in \mathbf{I}(\mathbf{S})$. So $\mathbf{N}(\bar{x}) \cap \overline{\mathbf{I}(\mathbf{S})} = \emptyset$ implies $\mathbf{N}(\bar{x}) = \emptyset$ in this case. By Theorem 7.6 these facts imply that $\bar{x}$ is the nearest point in $\text{Pos}(\mathbf{\Gamma})$ to b.

□

Example 7.8

Consider B, b given in Exercise 7.6. Let $\mathbf{S} = \{B_{.2}, B_{.3}\}$, $b(\mathbf{S}) = \hat{x} = (0,-4,7)^T$ given in Example 7.6 (computed in Example 7.3) and $\hat{x} \in \text{Pos}(\mathbf{S})$. In Example 7.6 we computed that $\mathbf{N}(\hat{x}) = \emptyset$ and so $\mathbf{N}(\hat{x}) \cap \overline{\mathbf{I}(\mathbf{S})} = \emptyset$. This implies that $\hat{x}$ is the nearest point in $\text{Pos}(B)$ to b.

Theorem 7.10 *Let $\bar{x} \in Pos(\mathbf{\Gamma})$ be such that $0 \in \mathbf{T}(b;\bar{x})$. If there exists an index j such that $(b-\bar{x})^T B_{.i} \leqq 0$ for all $i \neq j$, then $\mathbf{K}_j \cap \mathbf{B}(b;\bar{x}) = \emptyset$.*

Proof. Clearly under these conditions $x^T(b-\bar{x}) \leqq 0$ for all $x \in \mathbf{K}_j$; however $x^T(b-\bar{x}) > 0$ for all $x \in \mathbf{B}(b;\bar{x})$. Hence $\mathbf{K}_j \cap \mathbf{B}(b;\bar{x}) = \emptyset$.

□

Theorem 7.11 *Let $\bar{x} \in Pos(\mathbf{\Gamma})$ be such that $0 \in \mathbf{T}(b;\bar{x})$. If there exists an index j such that $(b-\bar{x})^T B_{.i} \leqq 0$ for all $i \neq j$ and $(b-\bar{x})^T B_{.j} > 0$, then j is a critical index of $[\mathbf{\Gamma}, b]$.*

Proof. By Theorem 7.6, $\bar{x}$ is not the nearest point in $\text{Pos}(\mathbf{\Gamma})$ to b. Let $\hat{x}$ be the nearest point in $\text{Pos}(\mathbf{\Gamma})$ to b. Then $\hat{x} \in \mathbf{B}(b;\bar{x})$. By Theorem 7.10 $\mathbf{K}_j \cap \mathbf{B}(b;\bar{x}) = \emptyset$. Hence $\hat{x} \notin \mathbf{K}_j$ and thus j is a critical index of $[\mathbf{\Gamma}; b]$.

□

Example 7.9

Consider B, b given in Example 7.4. If $\bar{x} = V^2$, we verified in Example 7.4 that $\mathbf{N}(\bar{x}) = \{3\}$. This implies that 3 is a critical index of $[B; b]$.

Here we describe a routine for selecting a critical index. This routine terminates once a critical index is identified. Later on we will discuss the algorithm for solving the LCP (q, M) where M is a PD symmetric matrix, or the associated nearest point problem, using this routine.

Routine for Selecting a Critical Index

This routine operates on the nearest point problem $[\mathbf{\Gamma}; b]$ which is equivalent to the given LCP (q, M). Clearly if $b \in \text{Pos}(\mathbf{\Gamma})$, the nearest point in $\text{Pos}(\mathbf{\Gamma})$ to b is the point

b itself; so we assume that $b \notin \text{Pos}(\mathbf{\Gamma})$ in the sequel. As mentioned earlier, we also assume that $V^l \neq 0$ (as otherwise, 0 is the nearest point in $\text{Pos}(\mathbf{\Gamma})$ to b).

The routine maintains a nonempty subset of $\mathbf{\Gamma}$ called the **current set** denoted by $\mathbf{S}$, and a point called the **current point** denoted by $\bar{x}$. $\bar{x} \in \text{Pos}(\mathbf{S})$ always. As these things change from step to step, the symbols $\mathbf{S}$, $\bar{x}$ may represent different things in different steps.

Initial Step: Set $\bar{x} = V^l$, and compute $\mathbf{N}(\bar{x})$. If $\mathbf{N}(\bar{x}) = \emptyset$, $\bar{x}$ is the nearest point in $\text{Pos}(\mathbf{\Gamma})$ to b, terminate. If $\mathbf{N}(\bar{x})$ is a singleton set, say i_1, i_1 is a critical index of $[\mathbf{\Gamma}; b]$, terminate this routine. If the cardinality of $\mathbf{N}(\bar{x})$ is greater than or equal to 2, choose $g \in \mathbf{N}(\bar{x})$; compute the orthogonal projection $\hat{b}$ of b onto the linear hull of $\{\bar{x}, B_{.g}\}$. Replace $\bar{x}$ by $\hat{b}$. Set $\mathbf{S} = \{B_{.l}, B_{.g}\}$. Go to Step 1.

Step 1: Let $\mathbf{S}$, $\bar{x}$ be the current entities. Compute $\mathbf{N}(\bar{x})$. If $\mathbf{N}(\bar{x}) = \emptyset$, $\bar{x}$ is the nearest point in $\text{Pos}(\mathbf{\Gamma})$ to b, terminate. If $\mathbf{N}(\bar{x})$ is a singleton set, say i_1, i_1 is a critical index of $[\mathbf{\Gamma}; b]$, terminate this routine. If the cardinality of $\mathbf{N}(\bar{x})$ is greater than or equals 2, go to Step 2 if $\mathbf{N}(\bar{x}) \cap \overline{\mathbf{I}(\mathbf{S})} \neq \emptyset$, or to Step 3 if $\mathbf{N}(\bar{x}) \cap \overline{\mathbf{I}(\mathbf{S})} = \emptyset$.

Step 2: Choose a $g \in \mathbf{N}(\bar{x}) \cap \overline{\mathbf{I}(\mathbf{S})}$. Compute $\hat{b}$, the orthogonal projection of b onto the linear hull of $\{\bar{x}, B_{.g}\}$. Replace $\mathbf{S}$ by $\mathbf{S} \cup \{B_{.g}\}$, and $\bar{x}$ by $\hat{b}$. Go back to Step 1.

Step 3: Compute $b(\mathbf{S})$. If $b(\mathbf{S}) \in \text{Pos}(\mathbf{S})$, replace $\bar{x}$ by $b(\mathbf{S})$ and go to Step 1. If $b(\mathbf{S}) \notin \text{Pos}(\mathbf{S})$, go to Step 4.

Step 4: Let the current point $\bar{x} = \sum(\alpha_j B_{.j} : j \in \mathbf{I}(\mathbf{S}))$, where $\alpha_j \geqq 0$ for all $j \in \mathbf{I}(\mathbf{S})$. Let $b(\mathbf{S}) = \sum[\gamma_j B_{.j} : j \in \mathbf{I}(\mathbf{S})]$. Since $b(\mathbf{S}) \notin \text{Pos}(\mathbf{S})$, $\gamma_j < 0$ for some $j \in \mathbf{I}(\mathbf{S})$. An arbitrary point on the line segment joining $\bar{x}$ to $b(\mathbf{S})$ can be written as $Q(\lambda) = (1-\lambda)\bar{x} + \lambda b(\mathbf{S})$, $0 \leqq \lambda \leqq 1$; or equivalently $Q(\lambda) = \sum[(((1-\lambda)\alpha_j + \lambda\gamma_j)B_{.j}) : j \in \mathbf{I}(\mathbf{S})]$. As λ increases from 0 to 1, $Q(\lambda)$ moves from $\bar{x}$ to $b(\mathbf{S})$. Let $\lambda = \bar{\lambda}$ be the largest value of λ for which $Q(\lambda)$ is in $\text{Pos}(\mathbf{S})$. So $Q(\bar{\lambda})$ is on the boundary of $\text{Pos}(\mathbf{S})$ and $Q(\lambda) \notin \text{Pos}(\mathbf{S})$ for $\lambda > \bar{\lambda}$. So $\bar{\lambda} = \max\{\lambda : (1-\lambda)\alpha_j + \lambda\gamma_j \geqq 0, \text{ for all } j \in \mathbf{I}(\mathbf{S})\}$. The point $(1-\bar{\lambda})\bar{x} + \bar{\lambda} b(\mathbf{S}) = Q(\bar{\lambda})$ is the last point in the cone $\text{Pos}(\mathbf{S})$ on the line segment joining $\bar{x}$ and $b(\mathbf{S})$, as you move away from $\bar{x}$ along this line segment. See Figure 7.3.

Let k be such that $(1-\bar{\lambda})\alpha_k + \bar{\lambda}\gamma_k = 0$. If there is more than one index in $\mathbf{I}(\mathbf{S})$ with this property, choose one of the them arbitrarily and call it k. $Q(\bar{\lambda})$ is the nearest point to $b(\mathbf{S})$ on the line segment joining $\bar{x}$ to $b(\mathbf{S})$ that lies in $\text{Pos}(\mathbf{S})$. So $Q(\bar{\lambda}) \in \text{Pos}(\mathbf{S} \setminus \{B_{.k}\})$. Delete $B_{.k}$ from $\mathbf{S}$. Also delete k from $\mathbf{I}(\mathbf{S})$ and include it in $\overline{\mathbf{I}(\mathbf{S})}$. Replace $\bar{x}$ by $Q(\bar{\lambda})$ and go to Step 3.

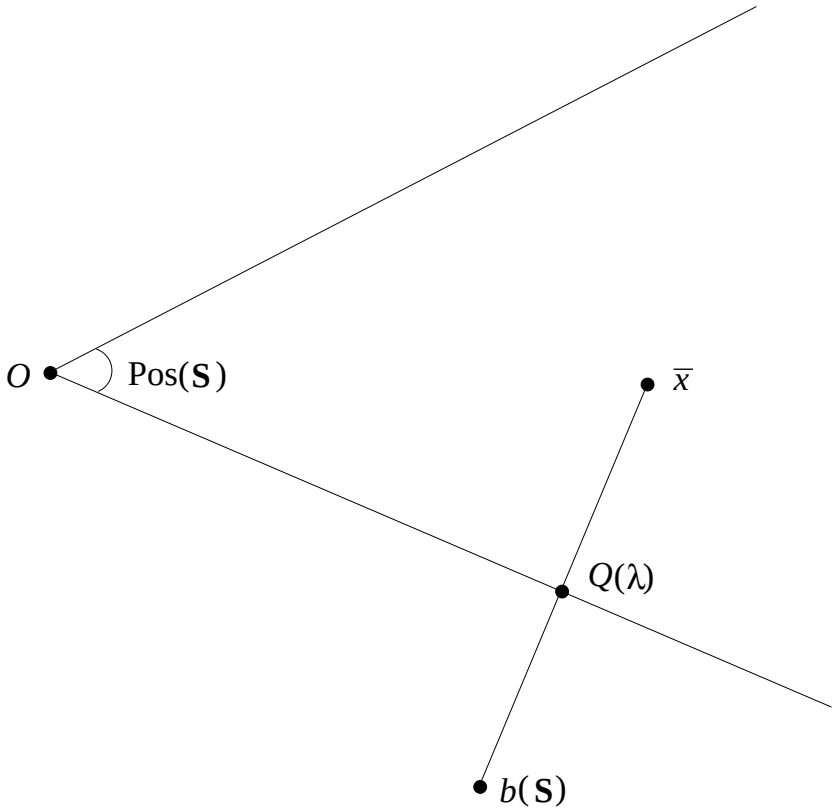


Figure 7.3

Discussion

If termination does not occur in the Initial Step, when we move to Step 1 we will have $||\bar{x} - b|| < ||V^l - b||$ by Theorem 7.8, and this property will continue to hold in all subsequent steps, since $||\bar{x} - b||$ never increases in the routine. Clearly $\bar{x} \in \text{Pos}(\mathbf{S})$ always. These facts imply that once the algorithm enters Step 1, the cardinality of $\mathbf{S}$ will always be greater than or equal 2.

While executing Step 4, if $\bar{\lambda}$ turns out to be zero, there is no change in the point $\bar{x}$, but the cardinality of the set $\mathbf{S}$ decreases by 1 at the end of this step. Thus a sequence of consecutive moves in the algorithm of the form Step 3 $\rightarrow$ Step 4 $\rightarrow$ Step 3 ..., must terminate after at most $(n-2)$ visits to Step 4, with $\bar{x}$ set equal to $b(\mathbf{S})$ for some projection face $\text{Pos}(\mathbf{S})$ in Step 3, and then the routine moves to Step 1. When this happens, while executing Step 1, by Theorem 7.9 either the routine itself terminates; or else Step 2 must be taken implying a strict decrease in $||\bar{x} - b||$ by Theorem 7.8 with the new $\bar{x}$ via Step 2, and thus the projection face $\text{Pos}(\mathbf{S})$ cannot repeat.

Whenever the routine visits Step 1, the current point $\bar{x}$ is the orthogonal projection of b onto a subspace of dimension 2 or more, and hence the property $0 \in \mathbf{T}(b; \bar{x})$ will hold then. Clearly, this property also holds in the Initial Step.

In the Initial Step, or in Step 1, if $\mathbf{N}(\bar{x}) = \emptyset$, $\bar{x}$ is the nearest point in $\text{Pos}(\mathbf{\Gamma})$ to b by Theorem 7.9. In these steps, if $\mathbf{N}(\bar{x})$ is a singleton set, the element in it is a critical

index for $[\mathbf{\Gamma}; b]$ by Theorem 7.11.

Since there are but a finite number of projection faces, these facts imply that if the routine does not terminate in the Initial Step, it terminates after a finite number of steps while executing Step 1.

When termination occurs in Step 1, it either finds the nearest point in $\text{Pos}(\mathbf{\Gamma})$ to b, in which case the problem is completely solved, or it finds a critical index of the problem. In the latter case an LCP of order $n-1$ can be constructed and the same routine can be applied to this smaller problem, as discussed in Theorem 7.7. The solution to the original problem then can be obtained using the solution of this smaller problem, as discussed in Theorem 7.7. Hence the unique solution of (q, M) can be obtained after at most n applications of the routine discussed above on LCPs of decreasing orders, each one associated with a positive definite symmetric matrix. We will now provide a summary of the whole algorithm.

Algorithm for Solving the LCP (q, M) When M is PD *Symmetric*

Step 0: Let (q, M) be the LCP and $[B; b]$ the corresponding nearest point problem. Check if $B^{-1}b \geqq 0$. If it is, $b \in \text{Pos}(B)$ and b itself is the nearest point in $\text{Pos}(B)$ to b. In this case z is a complementary feasible basic vector to the LCP (q, M) and the solution for it is $(w = 0, z = M^{-1}q)$. If this condition is not satisfied, continue.

Check if $b^T B \leqq 0$. If it is, the origin 0 is the nearest point in $\text{Pos}(B)$ to b. In this case w is a complementary feasible basic vector to the LCP (q, M), that is, $q \geqq 0$, and $(w = q, z = 0)$ is the solution of the LCP. If this condition is not satisfied, continue.

For $j = 1$ to n, define

$$V^j = \begin{cases} 0 & \text{if } b^T B_{.j} \leqq 0 \\ \left(\frac{b^T B_{.j}}{||B_{.j}||^2}\right) B_{.j} & \text{otherwise .} \end{cases}$$

Let V^l be the nearest among $V^1, \ldots, V^n$ to b. Break ties for l arbitrarily. Go to Step 1 with $\mathbf{S} = \{B_{.l}\}$, $\bar{x} = V^l$, $\mathbf{I}(\mathbf{S}) = \{l\}$.

Step 1: Let $\bar{x}$ be the current point and $\mathbf{S}$ the current subset of columns of B. Compute $\mathbf{N}(\bar{x}) = \{j : (b - \bar{x})^T B_{.j} > 0\}$.

If $\mathbf{N}(\bar{x}) = \emptyset$, $\bar{x}$ is the nearest point in $\text{Pos}(B)$ to b. Define for $j = 1$ to n

$$y_j = \begin{cases} z_j & \text{if } j \in \mathbf{I}(\mathbf{S}) \\ w_j & \text{otherwise .} \end{cases}$$

Then $y = (y_1, \ldots, y_n)$ is a complementary feasible basic vector for the LCP (q, M) and $(\bar{w} = M\bar{z} + q, \bar{z} = B^{-1}\bar{x})$ is the solution of the LCP. Terminate.

If $\mathbf{N}(\bar{x})$ is a singleton set, that is, if $\mathbf{N}(\bar{x}) = \{j_1\}$ for some j_1, j_1 is a critical index. Using it, reduce the LCP to one of order one less as in Theorem 7.7, and obtain the corresponding nearest point problem of dimension one less either by finding the Cholesky factor of the matrix associated with the reduced LCP or by using the

geometric procedure described following the proof of Theorem 7.7. With the reduced LCP and the reduced nearest point problem, go back to Step 0.

If the cardinality of $\mathbf{N}(\bar{x})$ is greater than or equal to 2, go to Step 2 if $\mathbf{N}(\bar{x}) \cap \overline{\mathbf{I}(\mathbf{S})} \neq \emptyset$, or to Step 3 otherwise.

Step 2: Select a $g \in \mathbf{N}(\bar{x}) \cap \overline{\mathbf{I}(\mathbf{S})}$. Compute $\hat{b}$, the orthogonal projection of b on the linear hull of $\{\bar{x}, B_{.g}\}$. Include $B_{.g}$ in $\mathbf{S}$, g in $\mathbf{I}(\mathbf{S})$, and replace $\bar{x}$ by $\hat{b}$ and go back to Step 1.

Step 3: Compute $b(\mathbf{S})$, the orthogonal projection of b on the linear hull of $\mathbf{S}$. If $b(\mathbf{S}) \in \text{Pos}(\mathbf{S})$, replace $\bar{x}$ by $b(\mathbf{S})$, and go back to Step 1 leaving $\mathbf{S}$, $\mathbf{I}(\mathbf{S})$ the same. If $b(\mathbf{S}) \notin \text{Pos}(\mathbf{S})$, go to Step 4.

Step 4: Let $\bar{x} = \sum_{j \in \mathbf{I}(\mathbf{S})} \alpha_j B_{.j}$ and $b(\mathbf{S}) = \sum_{j \in \mathbf{I}(\mathbf{S})} \gamma_j B_{.j}$. Now compute the value $\bar{\lambda} = \min\left\{\frac{\alpha_j}{(\alpha_j - \gamma_j)} : j \text{ such that } \gamma_j < 0\right\}$, and let k be an index which attains this minimum. Break ties for k arbitrarily. Replace $\bar{x}$ by $(1 - \bar{\lambda})\bar{x} + \bar{\lambda} b(\mathbf{S})$. Delete $B_{.k}$ from $\mathbf{S}$ and k from $\mathbf{I}(\mathbf{S})$, and go back to Step 3.

For solving LCPs (7.1) in which M is a given positive definite symmetric matrix, or equivalently the nearest point problem $[\mathbf{\Gamma}; b]$ where $\mathbf{\Gamma}$ is a given basis for $\mathbf{R}^n$; the approach discussed here seems to be the most efficient from a practical point of view. Empirical results on the computational efficiency of this approach are reported in Chapter 8.

7.1 Exercises

7.3 Let $\mathbf{\Gamma} = \{B_{.1}, \ldots, B_{.n}\}$ be a basis for $\mathbf{R}^n$ and b be another point in $\mathbf{R}^n$. Suppose it is required to find the nearest point in $\text{Pos}(\mathbf{\Gamma})$ to b in terms of the L_1-distance, also known as the rectilinear distance. The rectilinear distance between two points $x = (x_j)$, $y = (y_j)$ in $\mathbf{R}^n$ is defined to be $\sum_{j=1}^{n} (|x_j - y_j|)$. Show that this problem can be formulated as an LP. Given the nearest point in $\text{Pos}(\mathbf{\Gamma})$ to b in terms of the L_1 distance, can you draw from it any conclusions about the location of the nearest point in $\text{Pos}(\mathbf{\Gamma})$ to b in terms of the Euclidean distance? (explore questions like whether they lie in the same face etc.)

7.4 Let $\mathbf{\Gamma}$ be a subset consisting of a finite number of column vectors from $\mathbf{R}^n$, which is not linearly independent, and let $b \in \mathbf{R}^n$ be another column vector. It is required to find the nearest point in $\text{Pos}(\mathbf{\Gamma})$ to b. Modify the algorithm discussed above to solve this problem.

7.5 Let $\mathbf{K} \subset \mathbf{R}^n$ be a given convex polyhedron, and let $b \in \mathbf{R}^n$ be a given point. It is required to find the nearest point in $\mathbf{K}$ (in terms of the usual Euclidean distance) to b. $\mathbf{K}$ may be given in one of two forms:

(i) All the extreme points and extreme homogeneous solutions associated with $\mathbf{K}$ may be given, or

(ii) The constraints which define $\mathbf{K}$ may be given, for example $\mathbf{K} = \{x : Ax \geqq p, Dx = d\}$ where A, D, p, d are given.

Modify the algorithm discussed above, to find the nearest point in $\mathbf{K}$ to b, when $\mathbf{K}$ is given in either of the forms mentioned above.

7.6 Generalize the algorithm discussed above, to process the LCP (q, M) when M is PSD and symmetric.

7.7 Let $b \in \mathbf{R}^n$, $b > 0$ and let $\mathbf{K} = \{y : 0 \leqq y \leqq b\}$ be a rectangle. For $x \in \mathbf{R}^n$ let $P_{\mathbf{K}}(x)$ be the nearest point (in terms of the usual Euclidean distance) to x in $\mathbf{K}$. For any $x, y \in \mathbf{R}^n$, prove the following:

(1) The i^{th} coordinate of $P_{\mathbf{K}}(x)$ is $\min\{\max\{0, x_i\}, b_i\}$,

(2) $x \leqq y$ implies $P_{\mathbf{K}}(x) \leqq P_{\mathbf{K}}(y)$,

(3) $P_{\mathbf{K}}(x) - P_{\mathbf{K}}(y) \leqq P_{\mathbf{K}}(x - y)$,

(4) $P_{\mathbf{K}}(x + y) \leqq P_{\mathbf{K}}(x) + P_{\mathbf{K}}(y)$,

(5) $P_{\mathbf{K}}(x) + P_{\mathbf{K}}(-x) \leqq |x| = (|x_j|)$, with equality holding if $-b \leqq x \leqq b$.

(B. H. Ahn [7.1])

7.8 Let $f(x)$ be a real valued convex function defined on $\mathbf{R}^n$. Let $\bar{x} \in \mathbf{R}^n$, $\alpha \in \mathbf{R}^1$ be given. It is required to find a point that minimizes the distance $||x - \bar{x}||$ over $\{x : f(x) \leqq \alpha\}$. Develop an efficient algorithm for this problem. What changes are needed in this algorithm if $f(x) = (f_1(x), \ldots, f_m(x))^T$ where each $f_i(x)$ is a real valued convex function defined on $\mathbf{R}^n$, and $\alpha \in \mathbf{R}^m$?

7.9 Let B a square nonsingular matrix of order n. Let $M = B^T B$. Let $\mathbf{J} \subset \{1, \ldots, n\}$, with elements in $\mathbf{J}$ arranged in increasing order. Let $M_{\mathbf{JJ}}$ denote the principal submatrix of M corresponding to the subset $\mathbf{J}$. For any column vector $q \in \mathbf{R}^n$, let $q_{\mathbf{J}}$ denote the column vector of $(q_j : j \in \mathbf{J})$ with the entries in q_j arranged in the same order as the elements j are in $\mathbf{J}$.

It is required to find a point p in the interior of $\text{Pos}(B)$ satisfying :

Property 1: For every nonempty face $\mathbf{F}$ of $\text{Pos}(B)$, the orthogonal projection of p in the linear hull of $\mathbf{F}$, is in the relative interior of $\mathbf{F}$.

Prove that $p \in \mathbf{R}^n$ satisfies Property 1 iff $(M_{\mathbf{JJ}})^{-1} q_{\mathbf{J}} > 0$ for all subsets $\mathbf{J} \subset \{1, \ldots, n\}$, where $q = B^T p$.

If $n = 2$, prove that a point p satisfying Property 1 always exists. In this case, show that p can be taken to be any nonzero point on the bisector of the angle (that is less than 180°) created by the rays of $B_{.1}$ and $B_{.2}$ in $\mathbf{R}^2$.

For general n, let $A = B^{-1}$. Then $\{x : A_{i.} x = 0\}$ is the hyperplane $\mathbf{H}_i$ which is the linear hull of $\{B_{.1}, \ldots, B_{.i-1}, B_{.i+1}, \ldots, B_{.n}\}$. The generalization of finding a point

on the bisector of the angle between the rays of $B_{.1}$, $B_{.2}$ when $n = 2$, is to find a point p, satisfying the property that the shortest distances from p to each of the hyperplanes $\mathbf{H}_i$, $i = 1$ to n, are all equal. A point like this would be a positive scalar multiple of $d = Be$. Is the statement "if a point p satisfying Property 1 exists, $d = p$ is one such point" true?

Show that if

$$M = \begin{pmatrix} 6 & -4 & 1 & 0 & 0 & 0 \\ -4 & 6 & -4 & 1 & 0 & 0 \\ 1 & -4 & 6 & -4 & 1 & 0 \\ 0 & 1 & -4 & 6 & -4 & 1 \\ 0 & 0 & 1 & -4 & 6 & -4 \\ 0 & 0 & 0 & 1 & -4 & 6 \end{pmatrix}$$

and B is such that $B^T B = M$, there exists no point p satisfying Property 1.

Derive necessary and sufficient conditions on the matrix B to guarantee that a point p satisfying Property 1 exists.

(This problem came up in the algorithm discussed in Exercise 2.20. The numerical example is due to J. S. Pang)

7.10 Let M be a square matrix of order n, which is PSD, but not necessarily symmetric. Let $\widehat{M} = (M + M^T)/2$. Prove that $x^T \widehat{M}$ and $q^T x$ are constants over the solution set of the LCP (q, M).

7.11 $\{A_{.1}, \ldots, A_{.n+1}\}$ is a set of column vectors in $\mathbf{R}^n$ such that $\{A_{.2} - A_{.1}, \ldots, A_{.n+1} - A_{.1}\}$ is linearly independent. b is another column vector in $\mathbf{R}^n$. Let $\mathbf{K}$ be the n-dimensional simplex which is the convex hull of $\{A_{.1}, \ldots, A_{.n+1}\}$. Develop an efficient algorithm of the type discussed in this chapter, for finding the nearest point (in terms of the usual Euclidean distance) to b in $\mathbf{K}$.

7.12 Let $\mathbf{\Gamma} = \{A_{.1}, \ldots, A_{.m}\}$ be a given finite set of column vectors in $\mathbf{R}^n$. Let $\mathbf{K}$ be the convex hull of $\mathbf{\Gamma}$.

Suppose x^* is the point minimizing $\|x\|$ over $x \in \mathbf{K}$. For any $y \in \mathbf{R}^n$, $y \neq 0$, define

$h(y) =$ maximum value of $y^T x$, over $x \in \mathbf{K}$

$s(y) =$ a point in $\mathbf{\Gamma}$ which maximizes $y^T x$ over $x \in \mathbf{K}$. So, $h(y) = y^T s(y)$.

Incidentally, $h(y)$, $s(y)$ can be found by computing $y^T A_{.j}$ for each $j = 1$ to m and choosing $s(y)$ to be an $A_{.p}$ where p is such that $y^T A_{.p} =$ maximum $\{y^T A_{.j} : j = 1$ to $m\}$.

(i) Prove that x^* can be expressed as a convex combination of at most $n + 1$ vectors from $\mathbf{\Gamma}$.

(ii) If $0 \notin \mathbf{K}$, prove that x^* can be expressed as a convex combination of at most n vectors from $\mathbf{\Gamma}$.

(iii) For each $x \in \mathbf{K}$, prove that $\|x\|^{\mathbf{2}} + h(-x) \geqq 0$. Also prove that $\|x\|^{\mathbf{2}} + h(-x) = 0$ for $x \in \mathbf{K}$ iff $x = x^*$.

(iv) For any $x \in \mathbf{K}$, $x \neq x^*$, prove that $s(-x) - x$ is a descent direction for $\|x\|$.

(v) For any $x \in \mathbf{K}$ satisfying $\|x\|^{\mathbf{2}} + h(-x) > 0$, prove that there must exist a point $\overline{x}$ on the line segment joining x and $s(-x)$ such that $\|\overline{x}\| < \|x\|$.

(vi) Consider the following algorithm for minimizing the norm $\|x\|$ over $x \in \mathbf{K}$ by R. O. Barr and E. G. Gilbert. If $0 \in \mathbf{\Gamma}$, clearly x^*, the point minimizing $\|x\|$ over $x \in \mathbf{K}$, is 0 itself, so we assume that $0 \notin \mathbf{\Gamma}$. The algorithm operates with a subset $\mathbf{S} \subseteq \mathbf{\Gamma}$ satisfying $|\mathbf{S}| \leqq n + 1$ always, and $\mathbf{S}$ is the set of vectors of a simplex. The set $\mathbf{S}$ changes from step to step. Let the index set of $\mathbf{S}$ be $\mathbf{I}(\mathbf{S}) = \{j : A_{.j} \in \mathbf{S}\}$.

The algorithm needs a subroutine for minimizing $\|x\|$ over a simplex. If $\mathbf{\Gamma}$ is the set of vertices of a simplex (i. e., $\mathbf{K}$ is a simplex) the problem is solved by calling this subroutine once, terminate. So, we assume that $\mathbf{K}$ is not a simplex in the sequel.

Let rank $(\mathbf{\Gamma}) = r$. Initiate the algorithm with an arbitrary subset $\mathbf{S}$ of $r + 1$ or less vectors from $\mathbf{\Gamma}$ whose convex hull is a simplex (we can initiate the algorithm with $\mathbf{S} = \{A_{.l}\}$ where l is such that $\|A_{.l}\| =$ minimum $\{\|A_{.j}\| : j = 1 \text{ to } m\}$).

General Step: Let $\mathbf{S}$ be the current subset of vectors from $\mathbf{\Gamma}$, and $\mathbf{I}(\mathbf{S})$ its index set. Find $\overline{x}$, the point of minimum norm $\|x\|$, in the convex hull of $\mathbf{S}$ (for executing this, you need a subroutine to minimize the norm $\|x\|$ on a simplex).

If $\overline{x} = 0$, then $0 \in \mathbf{K}$, $x^* = 0$, terminate the algorithm.

If $\overline{x} \neq 0$, compute $\|\overline{x}\|^{\mathbf{2}} + h(-\overline{x})$. If $\|\overline{x}\|^{\mathbf{2}} + h(-\overline{x}) = 0$, then $x^* = \overline{x}$, terminate the algorithm.

If $\overline{x} \neq 0$ and $\|\overline{x}\|^{\mathbf{2}} + h(-\overline{x}) > 0$, let $\overline{x} = \sum(a_j A_{.j} : j \in \mathbf{I}(\mathbf{S}))$. Since $\overline{x}$ is the point of minimum norm in the convex hull of $\mathbf{S}$ and $\overline{x} \neq 0$, $\overline{x}$ must be a boundary point of the convex hull of $\mathbf{S}$, that is, $a_j = 0$ for at least one $j \in \mathbf{I}(\mathbf{S})$. Let $\mathbf{J} = \{j : j \in \mathbf{I}(\mathbf{S})$ and $a_j = 0\}$. Replace $\mathbf{S}$ by $\{s(-\overline{x})\} \cup (\mathbf{S} \setminus \{A_{.j} : j \in \mathbf{J}\})$, update $\mathbf{I}(\mathbf{S})$; and with the new $\mathbf{S}$, $\mathbf{I}(\mathbf{S})$, go to the next step. Prove that $\mathbf{S}$ always remains the set of vertices of a simplex in this algorithm, and that the algorithm finds x^* after at most a finite number of steps.

(See R. O. Barr, "An efficient computational procedure for generalized quadratic programming problems", SIAM Journal on Control 7 (1969) 415–429; and R. O. Barr and E. G. Gilbert, "Some efficient algorithms for a class of abstract optimization problems arising in optimal control", IEEE Transactions on Automatic Control, AC-14 (1969) 640–652. My thanks to S. Keerthi for bringing this and the next two problems to my attention).

7.13 Let $\mathbf{\Gamma} = \{A_{.1}, \ldots, A_{.m}\}$ be a finite set of column vectors from $\mathbf{R}^n$; and b, another given column vector in $\mathbf{R}^n$. Discuss how the Barr-Gilbert algorithm presented in Exercise 7.12, can be used to find the nearest point (in terms of the Euclidean distance) in the convex hull of $\mathbf{\Gamma}$ to b.

7.14 Let $\mathbf{\Gamma} = \{A_{.1}, \ldots, A_{.m}\}$, $\mathbf{\Delta} = \{B_{.1}, \ldots, B_{.t}\}$ be two finite sets of column vectors from $\mathbf{R}^n$. Let $\mathbf{K}$, $\mathbf{P}$ denote the convex hulls of $\mathbf{\Gamma}$, $\mathbf{\Delta}$ respectively. It is required to find $x^* \in \mathbf{K}$, $y^* \in \mathbf{P}$ such that

$$\|x^* - y^*\| = \text{ minimum } \{\|x - y\| : x \in \mathbf{K}, y \in \mathbf{P}\}.$$

Using the fact that $\mathbf{K} - \mathbf{P}$ (defined in Appendix 2) is a convex set, discuss how the Barr-Gilbert algorithm presented in Exercise 7.12, can be used to find x^*, y^*.

7.2 References

7.1 B. H. Ahn, "Iterative methods for linear complementarity problems with upper-bounds on primary variables", *Mathematical Programming*, 26 (1983) 295–315.

7.2 K. G. Murty and Y. Fathi, "A Critical Index Algorithm for Nearest Point Problems on Simplicial Cones", *Mathematical Programming*, 23 (1982) 206–215.

7.3 P. Wolfe, "Algorithm for a Least Distance Programming Problem", *Mathematical Programming Study 1*, (1974) 190–205.

7.4 P. Wolfe, "Finding the Nearest Point in a Polytope", *Mathematical Programming*, 11 (1976) 128–149.

7.5 D. R. Wilhelmsen, "A Nearest Point Algorithm for Convex Polyhedral Cones and Applications to Positive Linear Approximations", *Mathematics of Computation*, 30 (1976) 48–57.

Chapter 8

POLYNOMIALLY BOUNDED ALGORITHMS FOR SOME CLASSES OF LCPs

In this chapter we discuss algorithms for special classes of LCPs, whose computational complexity is bounded above by a polynomial in either the order or the size of the LCP. We consider the LCP (q, M) where M is either a Z-matrix, or a triangular P-matrix, or an integer PSD-matrix.

8.1 Chandrasekaran's Algorithm for LCPs Associated with Z-Matrices

Consider the LCP (q, M) of order n, where M is a Z-matrix. As discussed in Section 3.4, $M = (m_{ij})$ is a Z-matrix if all its off diagonal entries are nonpositive, that is $m_{ij} \leqq 0$ for all $i \neq j$. The algorithm discussed below by R. Chandrasekaran [8.2], terminates after at most n principal pivot steps, with either a solution of the LCP (q, M) or the conclusion that it has no solution.

The Algorithm

The initial tableau is (8.1)

w	z	
I	$-M$	q

(8.1)

Step 1: Start with the initial tableau and with $w = (w_1, \ldots, w_n)$ as the initial complementary basic vector. If this is a feasible basis (i. e., if $q \geqq 0$) it is a complementary feasible basis, terminate. Otherwise, go to the next step.

General Step: Let $\bar{q}$ be the present update right hand side constants vector. If $\bar{q} \geqq 0$, the present basic vector is a complementary feasible basic vector, terminate. Otherwise select a t such that $\bar{q}_t < 0$. Let $-\overline{m}_{tt}$ be the present update entry in the t^{th} row and the column vector of z_t. At this stage, the present basic variable in row t will be w_t (this follows from statement 5 listed below). If $-\overline{m}_{tt} \geqq 0$, there exists no nonnegative solution for (8.1) and consequently the LCP (q, M) has no solution, terminate. Otherwise if $-\overline{m}_{tt} < 0$, perform a principal pivot step in position t and go to the next step.

Using the fact that the initial matrix M is a Z-matrix, we verify that in the initial system (8.1), for any $t = 1$ to n, all the entries in row t are nonnegative with the exception of the entry in the column of z_t. From the manner in which the algorithm is carried out, the following facts can be verified to hold.

1. All pivot elements encountered during the algorithm are strictly negative.
2. For any t such that no pivot step has been performed in the algorithm so far in row t, all the entries in this row on the left hand portion of the present updated tableau are nonnegative, except, possibly the entry in the column of z_t. The infeasibility conclusion in the algorithm follows directly from this fact.
3. If s is such that a pivot step has been carried out in row s in the algorithm, in all subsequent steps, the updated entry in this row in the column of any nonbasic z_i is nonpositive.
4. Once a pivot step has been performed in a row, the updated right hand side constant in it remains nonnegative in all subsequent steps. This follows from statements 1 and 3.
5. Once a variable z_t is made a basic variable, it stays as a basic variable, and its value remains nonnegative in the solution, in all subsequent steps.
6. All basic vectors obtained in the algorithm are complementary, and the algorithm terminates either with the conclusion of infeasibility or with a complementary feasible basis.
7. At most one principal pivot step is carried out in each position, thus the algorithm terminates after at most n pivot steps. Thus the computational effort measured in terms of basic operations like multiplications, additions, comparisons of real numbers, is at most $\mathcal{O}(n^3)$.

From these facts we conclude that if the system "$w - Mz = q$, $w \geqq 0$, $z \geqq 0$" is feasible and M is a Z-matrix, then the LCP (q, M) has a complementary feasible solution and the above algorithm finds it. Hence, when M is a Z-matrix, the LCP (q, M) has a solution iff $q \in \text{Pos}(I \vdots -M)$, or equivalently, every Z-matrix is a Q_0-matrix.

R. W. Cottle and R. S. Sacher, and J. S. Pang [8.7, 8.8] discuss several large scale applications of the LCP basid on this algorithm.

Exercises

8.1 Solve the LCP with the following data by Chandrasekaran's algorithm.

$$M = \begin{pmatrix} 1 & -2 & 0 & -2 & -1 \\ -1 & 0 & -1 & -2 & 0 \\ -2 & -3 & 3 & 0 & 0 \\ 0 & -1 & -1 & -2 & -1 \\ -2 & 0 & -1 & -2 & 3 \end{pmatrix}, \qquad q = \begin{pmatrix} -4 \\ -4 \\ -2 \\ -1 \\ -2 \end{pmatrix}.$$

8.2 Is the complementary pivot method guaranteed to process the LCP (q, M) when M is a Z-matrix ?

8.3 Discuss an efficient method for computing all the complementary solutions of the LCP (q, M) when M is a Z-matrix.

8.2 A Back Substitution Method for the LCPs Associated with Triangular P-Matrices

A square matrix $M = (m_{ij})$ of order n is said to be a lower triangular matrix if $m_{ij} = 0$ for all $j \geqq i+1$. It is upper triangular if M^T is lower triangular. The square matrix M is said to be a triangular matrix if there exists a permutation of its rows and columns which makes it lower triangular. A triangular matrix satisfies the following properties.

(i) The matrix has a row that contains a single nonzero entry.

(ii) The submatrix obtained from the matrix by striking off the row containing a single nonzero entry and the column in which that nonzero entry lies, also satisfies property (i). The same process can be repeated until all the rows and columns of the matrix are struck off.

A lower triangular or an upper triangular matrix is a P-matrix iff all its diagonal entries are strictly positive. A triangular matrix is a P-matrix iff every one of its single nonzero entries identified in the process (i), (ii) above is the diagonal entry in its row and is strictly positive. Thus a triangular matrix is a P-matrix iff there exists a permutation matrix Q such that $Q^T MQ$ is a lower triangular matrix with positive diagonal entries.

Example 8.1

Let

$$M = \begin{pmatrix} 1 & 0 & 0 & 2 \\ 2 & 1 & 0 & 2 \\ 2 & 2 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix} , \qquad Q = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix} .$$

Verify that $Q^T M Q = \widetilde{M}(4)$ defined in equation (1.15) for $n = 4$, and hence M is a triangular P-matrix.

If M is a triangular P-matrix, the LCP (q, M) can be solved by the following back substitution method.

Identify the row in $M = (m_{ij})$ containing a single nonzero entry. Suppose it is row t. If $q_t \geqq 0$, make $w_t = q_t$, $z_t = 0 = \bar{z}_t$. On the other hand, if $q_t < 0$, make $w_t = 0$, $z_t = \frac{q_t}{-m_{tt}} = \bar{z}_t$. Add $\bar{z}_t M_{.t}$ to the right hand side constants vector q in (8.1), and then eliminate the columns of w_t, z_t and the t^{th} row from (8.1), thus converting (8.1) into a system of the same form in the remaining variables, on which the same process is repeated.

In this method, the value of one complementary pair of variables (w_i, z_i) are computed in each step, their values are substituted in the other constraints and the process repeated. The method finds the complete solution in n steps.

Example 8.2

Consider the LCP (q, M) with

$$M = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 2 & 2 & 1 \end{pmatrix} , \qquad q = \begin{pmatrix} -8 \\ -12 \\ -14 \end{pmatrix} .$$

It can be verified that this method leads to the values $(w_1, z_1) = (0, 8)$, $(w_2, z_2) = (4, 0)$, $(w_3, z_3) = (2, 0)$ in that order, yielding the solution $(w_1, w_2, w_3; z_1, z_2, z_3) = (0, 4, 2; 8, 0, 0)$. The same problem was solved by the complementary pivot algorithm in Example 2.10.

8.3 Polynomially Bounded Ellipsoid Algorithms for LCPs Corresponding to Convex Quadratic Programs

In the following sections we show that the ellipsoid algorithms for linear inequalities and LPs (see references [8.13], [2.26]) can be extended to solve LCPs associated with PSD

matrices with integer data, in polynomial time. As shown in Chapter 1 every convex quadratic programming problem can be transformed into an LCP associated with a PSD matrix, and hence the methods described here provide polynomially bounded algorithms for solving convex quadratic programs with integer data. These algorithms are taken from S. J. Chung and K. G. Murty [8.4]. Similar work also appeared in [8.14, 8.1] among other references. If the data in the problem is not integer but rational, it could be converted into an equivalent problem with integer data by multiplying all the data by a suitably selected positive integer, and solved by the algorithms discussed here in polynomial time.

In Sections 8.1, 8.2 we discussed algorithm for special classes of LCPs in which the computational effort required to solve an LCP of order n is at most $\mathcal{O}(n^3)$. These algorithms do not require the data in the problem to be integer or rational, it could even be irrational as long as the matrix M satisfies the property of being a Z-matrix or triangular P-matrix as specified and the required arithmetical operations can be carried out on the data with the desired degree of precision. Thus these algorithms discussed in Section 8.1, 8.2 are extremely efficient and practically useful to solve LCPs of the types discussed there. The ellipsoid algorithms discussed in the following sections have an entirely different character. They are polynomially bounded as long as M is an integer PSD-matrix, but their computational complexity is not bounded above by a polynomial in the order of the problem, but by a polynomial in the size of the problem (the size of the problem is the total number of digits in all the data when it is encoded using binary encoding). From Chapter 6 we know that in the worst case, the complementary and principal pivoting method discussed earlier are not polynomially bounded. However, in computational tests on practical, or randomly generated problems, the observed average computational effort required by ellipsoid method turned out to be far in excess of that required by complementary and principal pivoting methods. Also, in the ellipsoid methods, each computation has to be carried out to a large number of digits of precision, making it very hard to implement them on existing computers.

Thus the ellipsoid algorithms discussed in the following sections are not likely to be practically useful, at least not in their present forms. The major importance of these ellipsoid methods is theoretical, they made it possible for us to prove that convex quadratic programs, or equivalently LCPs associated with PSD-matrices with integer data, are polynomially solvable.

Size of an LCP

In this and in subsequent sections, we use the symbol L to denote the size of the problem istance, it is the total number of binary digits in all the data in the instance, assuming that all the data is integer. Given an integer α, the total number of binary digits in it (i. e., the number of bits needed to encode it in binary form) is approximately $\lceil 1 + \log_2(1 + |\alpha|) \rceil$, the celing of $\big(1 + \log_2(1 + |\alpha|)\big)$, that is, the positive integer just $\geqq \big(1 + \log_2(1 + |\alpha|)\big)$. Since the data in an LCP (q, M) of order n is n, q, M, we can

define the size of this LCP to be

$$L = \left\lceil \left(1 + \log_2 n\right) + \sum_{i,j=1}^{n} \left(1 + \log_2(1 + |m_{ij}|)\right) + \sum_{j=1}^{n} \left(1 + \log_2(1 + |q_j|)\right) \right\rceil .$$

An Ellipsoid in $\mathbf{R}^n$

An ellipsoid in $\mathbf{R}^n$ is uniquely specified by its center $p \in \mathbf{R}^n$ and a positive definite matrix D of order n. Given these, the ellipsoid corresponding to them is $\{x : (x - p)^T D^{-1}(x - p) \leqq 1\}$ and is denoted by $\mathbf{E}(p, D)$. Notice that if $D = I$, the ellipsoid $\mathbf{E}(p, D)$ is the solid spherical ball with p as center and the radius equal to 1. When D is positive definite, for $x, y \in \mathbf{R}^n$, the function $f(x, y) = (x - y)^T D^{-1}(x - y)$ is called the **distance between** x **and** y **with** D^{-1} **as the metric matrix** (if $D = I$, this becomes the usual **Euclidean distance**). The ellipsoid methods discussed in the following sections obtain a new ellipsoid in each step by changing the metric matrix. Hence these methods belong to the family of **variable metric methods**. Also, the formula for updating the metric matrix from step to step is of the form $D_{r+1} =$ a constant times $(D_r + C_r)$, where D_j is the metric matrix in step j for $j = r, r + 1$; and C_r is a square matrix of order n and rank 1 obtained by multiplying a column vector in $\mathbf{R}^n$ by its transpose. Methods which update the metric matrix by such a formula are called **rank one methods** in nonlinear programming literature. Rank one methods and variables metric methods are used extensively for solving convex unconstrained minimization problems in nonlinear programming. See references [10.2, 10.3, 10.9, 10.13]. The ellipsoid methods discussed in the following sections belong to these families of methods.

8.4 An Ellipsoid Algorithm for the Nearest Point Problem on Simplicial Cones

Let $B = (b_{ij})$ be a nonsingular square matrix of order n, and $b = (b_i)$ a column vector in $\mathbf{R}^n$. We assume that all the data in B, b is integer, and consider the nearest point problem $[B; b]$ discussed in Chapter 7. This is equivalent to the LCP $(\bar{q}, \overline{M})$ where $\overline{M} = B^T B$, $\bar{q} = -B^T b$, and so $\overline{M}$, $\bar{q}$ are integer matrices too, and $\overline{M}$ is PD and symmetric. If $b \in \text{Pos}(B)$, then the point b is itself the solution of $[B; b]$, and $(\bar{w} = 0, \bar{z} = B^{-1}b)$ is the unique solution of the LCP $(\bar{q}, \overline{M})$. So we assume that $b \notin \text{Pos}(B)$ (this implies that $b \neq 0$). Here we present an ellipsoid algorithm for solving this nearest point problem $[B; b]$ and the corresponding LCP $(\bar{q}, \overline{M})$. We begin with some results necessary to develop the algorithm.

Definitions

Let ε be a small positive number. Later on we specify how small ε should be. Let

$$
\begin{aligned}
\mathbf{K} &= \{x:\ B^{-1}x \geqq 0, B^T(x-b) \geqq 0\} \\
\mathbf{E} &= \{x:\ (x-\tfrac{b}{2})^T(x-\tfrac{b}{2}) \leqq \tfrac{b^Tb}{4}\} \\
\mathrm{Bd}(\mathbf{E}) &= \text{Boundary of } \mathbf{E} = \{x:\ (x-\tfrac{b}{2})^T(x-\tfrac{b}{2}) = \tfrac{b^Tb}{4}\} \\
\mathbf{E}_1 &= \left\{x : (x-\tfrac{b}{2})^T(x-\tfrac{b}{2}) \leqq \left(\varepsilon + \sqrt{\tfrac{b^Tb}{4}}\right)^{\mathbf{2}}\right\} \\
L_1 &= \left\lceil \big(1+\log_2 n\big) + \sum_{i,j=1}^{n}\Big(1+\log_2(|b_{ij}|+1)\Big) + \sum_{i=1}^{n}\Big(1+\log_2(|b_i|+1)\Big)\right\rceil \\
L_2 &= n(n+1)(L_1+1) \\
L_3 &= \big(n(2n+1)+1\big)L_1 \\
\bar{x} &= \text{Nearest point in } \mathrm{Pos}(B) \text{ to } b \\
\overline{M} &= (\overline{m}_{ij}) = B^TB \\
\bar{q} &= (\bar{q}_i) = -B^Tb \\
\bar{z} &= B^{-1}\bar{x} \\
\bar{w} &= \bar{q} + \overline{M}\bar{z} \\
\delta &= \tfrac{3}{4}2^{-\mathbf{L_2}}.
\end{aligned}
$$

Some Preliminary Results

Our nearest point problem $[B;b]$ is equivalent to the LCP $(\bar{q},\overline{M})$. Each $\overline{m}_{ij}$ or $\bar{q}_i$ is of the form $\gamma_1\gamma_2+\gamma_3\gamma_4+\ldots+\gamma_{2n-1}\gamma_{2n}$, where the γ's are entries from B, b, and hence are integer. So we have

$$
\begin{aligned}
\log_2|m_{ij}| &= \log_2(|\gamma_1\gamma_2+\ldots+\gamma_{2n-1}\gamma_{2n}|) \\
&< \log_2\big((|\gamma_1|+2)(|\gamma_2|+2)+\ldots+(|\gamma_{2n-1}|+2)(|\gamma_{2n}|+2)\big) \\
&\leqq \log_2\big((|\gamma_1|+2)(|\gamma_2|+2)\ldots(|\gamma_{2n}|+2)\big) \\
&= \sum_{t=1}^{2n}\log_2(|\gamma_t|+2) \\
&\leqq \sum_{t=1}^{2n}\big(1+\log_2(|\gamma_t|+1)\big) \\
&\leqq L_1\ .
\end{aligned}
$$

So the total number of digits needed to specify the data in the LCP $(\bar{q},\overline{M})$ in binary encoding is at most L_2.

From well known results the absolute value of the determinant of any square submatrix of B is at most $\frac{2^{\mathbf{L_1}}}{n}$. See Chapter 15 in [2.26]. So there exists a positive integer $\gamma < \frac{2^{\mathbf{L_1}}}{n}$ such that all the data in the system

$$
\begin{aligned}
\gamma B^{-1}x &\geqq 0 \\
B^T(x-b) &\geqq 0
\end{aligned}
\tag{8.2}
$$

are integers. The absolute value of each entry in γB^{-1} is $< \left(\frac{2^{L_1}}{n}\right)^2$ (since it is less than or equal to a subdeterminant of B times γ). Hence the size of (8.2) the total number of digits in the data in it, in binary encoding, is at most L_3.

Theorem 8.1 $\mathbf{K}$ *has nonempty interior.*

Proof. Proving this theorem is equivalent to showing that there exists an $x \in \mathbf{R}^n$ satisfying each of the constraints in the definition of $\mathbf{K}$ as a strict inequality. This holds iff the system

$$\begin{aligned} B^{-1}x &> 0 \\ B^T x - B^T b x_{n+1} &> 0 \\ x_{n+1} &> 0 \end{aligned}$$

has a feasible solution $(x, x_{n+1}) = X$. By Motzkin's theorem of the alternatives (Theorem 5 of Appendix 1) this system has a feasible solution X iff there exists no row vectors $\pi, \mu \in \mathbf{R}^n$, $\delta \in \mathbf{R}^1$ satisfying

$$\begin{aligned} \pi B^{-1} + \mu B^T &= 0 \\ -\mu B^T b + \delta &= 0 \\ (\pi, \mu, \delta) &\geq 0 \end{aligned} \tag{8.3}$$

From the first set of constraints in this system we have $\mu B^T B = -\pi \leqq 0$. Since $B^T B$ is PD, we know that $\mu B^T B \leqq 0$, $\mu \geqq 0$ implies that μ, must be 0 in any feasible solution of (8.3). This in turn implies that π, δ will have to be zero too, a contradiction. So (8.3) has no feasible solution, hence $\mathbf{K}$ has a nonempty interior.

□

Theorem 8.2 $\mathbf{K} \cap \mathbf{E} = \mathbf{K} \cap \text{Bd}(\mathbf{E}) = \{\bar{x}\}$.

Proof. By the results in Chapter 7, $(\overline{w}, \bar{z})$ is the solution of the LCP $(\bar{q}, \overline{M})$. So $\bar{z} = B^{-1}\bar{x} \geqq 0$, $0 \leqq \overline{w} = \bar{q} + \overline{M}\bar{z} = -B^T b + B^T B B^{-1}\bar{x} = B^T(\bar{x} - b)$. Also $(\bar{x} - \frac{b}{2})^T(\bar{x} - \frac{b}{2}) - (\frac{b^T b}{4}) = \bar{x}^T\bar{x} - \bar{x}^T b = \bar{x}^T(\bar{x} - b) = \bar{z}^T B^T(\bar{x} - b) = \bar{z}^T\overline{w} = 0$. So $\bar{x} \in \mathbf{K} \cap \mathbf{E}$.

Conversely, suppose $\hat{x} \in \mathbf{K} \cap \mathbf{E}$. Define $\hat{z} = B^{-1}\hat{x}$, $\widehat{w} = B^T(\hat{x} - b)$. Since $\hat{x} \in \mathbf{E}$ we have $0 \geqq (\hat{x} - \frac{b}{2})^T(\hat{x} - \frac{b}{2}) - (\frac{b^T b}{4}) = \hat{x}^T(\hat{x} - b) = \hat{z}^T\widehat{w}$. Since $\hat{x} \in \mathbf{K}$, we have $\hat{z} \geqq 0$, $\widehat{w} \geqq 0$, and hence $\hat{z}^T\widehat{w} \geqq 0$. These two together imply that $\hat{z}^T\widehat{w} = 0$ and we can verify that $\widehat{w} = B^T(\hat{x} - b) = \bar{q} + \overline{M}\hat{z}$. These facts together imply that $(\widehat{w}, \hat{z})$ is the solution of the LCP $(\bar{q}, \overline{M})$. Since M is PD, by Theorem 3, the LCP $(\bar{q}, \overline{M})$ has a unique solution and so $(\widehat{w}, \hat{z}) = (\overline{w}, \bar{z})$. So $\hat{x} = \bar{x}$. Thus $\mathbf{K} \cap \mathbf{E} = \{\bar{x}\}$. Also, for all $x \in \mathbf{K}$ we have $(x - \frac{b}{2})^T(x - \frac{b}{2}) = x^T(x - b) + (\frac{b^T b}{4}) = (B^{-1}x)^T B^T(x - b) + (\frac{b^T b}{4}) \geqq (\frac{b^T b}{4})$. This implies that $\mathbf{K} \cap \mathbf{E} = \mathbf{K} \cap \text{Bd}(\mathbf{E})$.

□

Theorem 8.3 $\bar{x}$ *is an extreme point of* $\mathbf{K}$.

Proof. Since M is PD, $(\overline{w}, \bar{z})$, the unique solution of the LCP $(\bar{q}, \overline{M})$ defined above, is a complementary BFS. So $\bar{z}$ is an extreme point of $\{z : -\overline{M}z \leqq \bar{q}, z \geqq 0\} = \mathbf{\Gamma}$. It can be verified that $z \in \mathbf{\Gamma}$ iff $x = Bz \in \mathbf{K}$. So there is a unique nonsingular linear

transformation between $\mathbf{\Gamma}$ and $\mathbf{K}$. This, and the fact that $\bar{z}$ is an extreme point of $\mathbf{\Gamma}$ implies that $\bar{x} = B^{-1}\bar{z}$ is an extreme point of $\mathbf{K}$.

□

Theorem 8.4 *If $(\widetilde{w} = (\widetilde{w}_i), \tilde{z} = (\tilde{z}_i))$ is any extreme point of*

$$\begin{aligned} & w - \overline{M}z = \bar{q} \\ & w \geqq 0, \quad z \geqq 0 \ , \end{aligned} \tag{8.4}$$

then $\widetilde{w}_i$, $\tilde{z}_i$, is either 0 or $> 2^{-\mathbf{L_2}}$, for each i.

Proof. As discussed above, L_2 is the size of the system (8.4). This result follows from the results discussed in Chapter 15 of [2.26].

□

Theorem 8.5 *The Euclidean lenght of any edge of $\mathbf{K}$ is $\geqq 2^{-\mathbf{L_3}}$.*

Proof. If the edge is unbounded, the theorem is trivially true. Each bounded edge of $\mathbf{K}$ is the line segment joining two distinct adjacent extreme points of $\mathbf{K}$. Let x^1, x^2 be two distinct adjacent extreme points of $\mathbf{K}$. Since $\mathbf{K}$ is the set of feasible solutions of (8.2), the results discussed in Chapter 15 of [2.26] imply that $x^1 = (\frac{u_{11}}{v_1}, \ldots, \frac{u_{n1}}{v_1})$, $x^2 = (\frac{u_{12}}{v_2}, \ldots, \frac{u_{n2}}{v_2})$ where all the u_{ij}'s are integers, v_1, v_2 are nonzero integers, all $|u_{ij}|$, $|v_1|$, $|v_2|$ are $\leqq \frac{2^{\mathbf{L_3}}}{n}$. Also, since $x^1 \neq x^2$, these facts imply that there exists a j satisfying $|x^1_j - x^2_j| \geqq 2^{-\mathbf{L_3}}$. This clearly implies that $||x^1 - x^2|| \geqq 2^{-\mathbf{L_3}}$.

□

Theorem 8.6 *If $\varepsilon < 2^{-\mathbf{2(n+1)^2 L_1}}$, the n-dimensional volume of $\mathbf{K} \cap \mathbf{E}_1 \geqq \varepsilon^{\mathbf{n}} 2^{-\mathbf{(n+1)L_3}}$.*

Proof. $\mathbf{K} \cap \mathrm{Bd}(\mathbf{E}) = \{\bar{x}\}$ and $\mathbf{K}$ has a nonempty interior. So $\mathbf{K} \cap \mathbf{E}_1$ contains all the points in $\mathbf{K}$ is an ε-neighbourhood of $\bar{x}$, and hence has a nonempty interior and a positive n-dimensional volume.

If one takes a sphere of radius α, a concentric sphere of radius $\alpha + \varepsilon$, and a hyperplane tangent to the smaller sphere at a boundary point x on it, then a tight upper bound on the distance between x and any point in the larger sphere on the side of the hyperplane opposite the smaller sphere is $\sqrt{2\alpha + \varepsilon^{\mathbf{2}}}$. Also the radius of $\mathbf{E}$ is $\sqrt{\frac{b^T b}{4}} < 2^{(\mathbf{L_1 - 1})}$. $\bar{x}$ is an extreme point of $\mathbf{K}$, and every edge of $\mathbf{K}$ through $\bar{x}$, has a length $\geqq 2^{-\mathbf{L_3}}$ by Theorem 8.5. These facts and the choice of ε here, together imply that every edge of $\mathbf{K}$ through $\bar{x}$ intersects the boundary of $\mathbf{E}_1$. Let $V_1, \ldots, V_n$ be points along the edges of $\mathbf{K}$ through $\bar{x}$ that intersect the boundary of $\mathbf{E}_1$, at a distance of at most 1 but greater than ε from $\bar{x}$, such that $\{\bar{x}, V_1, \ldots, V_n\}$ is affinely independent. The portion of the edge between $\bar{x}$ and V_i lies inside $\mathbf{E}_1$ for at least a length of ε. See Figure 8.1. If $V_i(\varepsilon)$ is the point on the edge joining $\bar{x}$ and V_i at a distance of ε from $\bar{x}$, the volume of $\mathbf{E}_1 \cap \mathbf{K}$ is greater than or equal to the volume of the simplex whose vertices are $\bar{x}$, $V_i(\varepsilon)$ for $i = 1$ to n. From the choice of V_i, $V_i(\varepsilon) - \bar{x} = \gamma(V_i - \bar{x})$ where

$\gamma \geqq \varepsilon$. So in this case the volume of $\mathbf{E}_1 \cap \mathbf{K}$ is greater than or equal to

$$\begin{aligned} & \frac{1}{n!} \left| \text{determinant of } \begin{pmatrix} \varepsilon(V_1 - \bar{x}) & \vdots & \dots & \vdots & \varepsilon(V_n - \bar{x}) \end{pmatrix} \right| \\ = & \frac{\varepsilon^{\mathbf{n}}}{n!} \left| \text{ determinant of } \begin{pmatrix} (V_1 - \bar{x}) & \vdots & \dots & \vdots & (V_n - \bar{x}) \end{pmatrix} \right| \\ = & \frac{\varepsilon^{\mathbf{n}}}{n!} \left| \text{ determinant of } \begin{pmatrix} 1 & 1 & \vdots & \dots & \vdots & 1 \\ \bar{x} & V_1 & \vdots & \dots & \vdots & V_n \end{pmatrix} \right| \\ > & \varepsilon^{\mathbf{n}} 2^{-(\mathbf{n+1})\mathbf{L_3}} \ . \end{aligned}$$

using the results from Chapter 15 in [2.26].

□

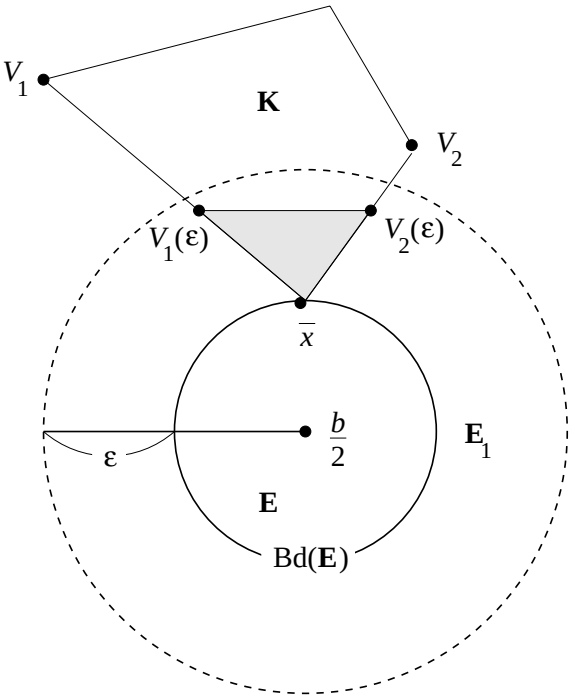


Figure 8.1 The volume of $\mathbf{E}_1 \cap \mathbf{K}$ is greater than or equal to the volume of the shaded simplex.

Theorem 8.7 *Let $\hat{x} \in \mathbf{E}_1 \cap \mathbf{K}$, $\hat{z} = B^{-1}\hat{x}$, $\widehat{w} = B^T(\hat{x} - b)$. Then, for all $j = 1$ to n*

$$\begin{aligned} |\hat{x}_j - \bar{x}_j| & \leqq 2^{\mathbf{L_1}} \sqrt{\varepsilon} \\ |\hat{z}_j - \bar{z}_j| & \leqq n 2^{\mathbf{2L_1}} \sqrt{\varepsilon} \\ |\widehat{w}_j - \overline{w}_j| & \leqq n 2^{\mathbf{2L_1}} \sqrt{\varepsilon} \end{aligned}$$

Proof. As mentioned earlier, the absolute value of any entry in B^{-1} is $\leqq 2^{\mathbf{L_1}}$, and the same fact obviously holds for B^T. The radius of $\mathbf{E}$ is $\frac{b^T b}{4} < 2^{\mathbf{L_1 - 1}}$. The results in this theorem follow from these facts and the definitions of $\mathbf{E}$, $\mathbf{E}_1$, $\widehat{w}$, $\hat{z}$.

□

Theorem 8.8 *Let $\hat{x} \in \mathbf{E}_1 \cap \mathbf{K}$ and $\hat{z} = B^{-1}\hat{x}$. If $\varepsilon \leqq 2^{-\mathbf{2(n+1)^2(L_1+1)}}$, then*

$$\begin{aligned} \hat{z}_j &\leqq (\tfrac{1}{4})2^{-\mathbf{L_2}}, \qquad \text{for } j \text{ such that } \bar{z}_j = 0 \\ \hat{z}_j &\geqq (\tfrac{3}{4})2^{-\mathbf{L_2}} = \delta, \text{ for } j \text{ such that } \bar{z}_j > 0 \ . \end{aligned}$$

Proof. This follows from Theorems 8.7 and 8.4.

□

The Algorithm

Fix $\varepsilon = 2^{-\mathbf{2(n+1)^2(L_1+1)}}$. Consider the following system of constraints.

$$-B^{-1}x \leqq 0, \qquad B^T(x - b) \leqq 0 \tag{8.5}$$

$$\left(x - \frac{b}{2}\right)^T \left(x - \frac{b}{2}\right) \leqq \left(\varepsilon + \sqrt{\frac{b^T b}{4}}\right)^{\mathbf{2}} \tag{8.6}$$

Any point $\hat{x} \in \mathbf{R}^n$ satisfying both (8.5) and (8.6) is in $\mathbf{K} \cap \mathbf{E}_1$. We use an ellipsoid method to first find such a point $\hat{x}$. Then using $\hat{x}$ we compute $\bar{x}$ in a final step.

Define $x^1 = \frac{b}{2}$, $A_1 = I\left(\varepsilon + \sqrt{\frac{b^T b}{4}}\right)^{\mathbf{2}}$, where I is the unit matrix of order n, $N = 8(n+1)^{\mathbf{4}}(L_1 + 1)$. Go to Step 2.

General Step $r+1$

Let x^r, A^r, $\mathbf{E}_r = \mathbf{E}(x^r, A_r)$ be respectively the center, positive definite symmetric matrix, and the ellipsoid at the beginning of this step. If x^r satisfies both (8.5), (8.6), terminate the ellipsoid method, call x^r as $\hat{x}$ and with it go to the final step described below. If x^r violates (8.5) select a constraint in it that it violates most, breaking ties arbitrarily, and suppose it is $ax \leqq d$. If x^r satisfies (8.5) but violates (8.6), find the point of intersection ξ^r, of the line segment joining x^1 and x^r with the boundary of $\mathbf{E}_1$. So $\xi^r = \lambda x^1 + (1-\lambda)x^r$ where $\lambda = 1 - \frac{\varepsilon + \sqrt{\frac{b^T b}{4}}}{||x^r - x^1||}$. Find the tangent plane of $\mathbf{E}_1$ at its boundary point ξ^r, and find out the half-space determined by this hyperplane which does not contain the point x^r. Suppose this half-space is determined by the constraint "$ax \leqq d$". See Figure 8.2.

Now define

$$\begin{aligned} \gamma_r &= \frac{d - ax^r}{\sqrt{aA_r a^T}} \\ x^{r+1} &= x^r - \left(\frac{1 - \gamma_r n}{1+n}\right) \frac{A_r a^T}{\sqrt{aA_r a^T}} \\ A_{r+1} &= \frac{(1-\gamma_r^2)n^2}{n^2-1}\left(A_r - \left(\frac{2}{n+1}\right)\left(\frac{1-\eta\gamma_r}{1-\gamma_r}\right)\frac{(A_r a^T)(A_r a^T)^T}{aA_r a^T}\right) \end{aligned} \tag{8.7}$$

where the square root of a quantity always represents the positive square root of that quantity. With x^{r+1}, A_{r+1}, $\mathbf{E}_{r+1} = \mathbf{E}(x^{r+1}, A_{r+1})$ move to the next step in the ellipsoid method.

After at most N steps, this ellipsoid method will terminate with the point x^r in the terminal step lying in $\mathbf{E}_1 \cap \mathbf{K}$. Then go to the final step discussed below.

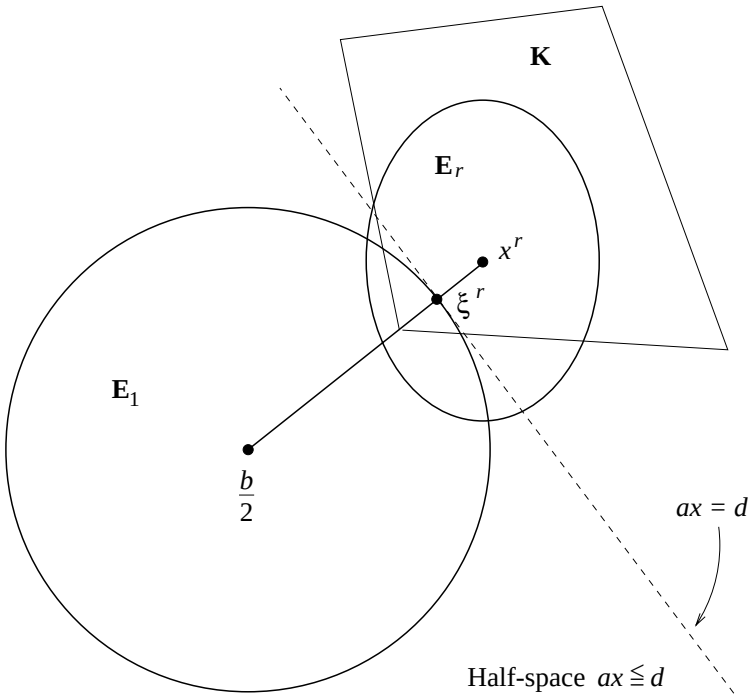


Figure 8.2 Construction of "$ax \leqq d$" when x^r satisfies (8.5) but violates (8.6).

Final Step : Let the center of the ellipsoid in the terminal step be $\hat{x}$ (this is the point x^r in the last step r of the ellipsoid method). Let $\hat{z} = B^{-1}\hat{x}$. Let $\mathbf{J} = \{j : j$ such that $\hat{z}_j \geqq \delta\}$. Let $y_j = z_j$ if $j \in \mathbf{J}$, w_j if $j \notin \mathbf{J}$ and let $y = (y_1, \ldots, y_n)$. Then y is a complementary feasible basic vector for the LCP $(\bar{q}, \overline{M})$, and the BFS of (8.4) corresponding to y is the solution of this LCP. If this solution is $(\overline{w}, \bar{z})$, $\bar{x} = B\bar{z}$ is the nearest point in Pos(B) to b.

Definition We denote by e, the base of natural logarithms. $e = 1 + \sum_{n=1}^{\infty} \frac{1}{n!}$, it is approximately equal to 2.7.

Proof of the Algorithm

Let x^r, A_r, $\mathbf{E}_r = \mathbf{E}(x^r, A_r)$, be the center, positive definite symmetric matrix, and the ellipsoid at the beginning of step $r+1$. The inequality "$ax \leqq d$" is choosen in this step $r+1$ in such a way that x^r violates it. In the hyperplane "$ax = d$" decrease d until a value d_1 is reached such that the translate "$ax = d_1$" is a tangent plane to the

ellipsoid $\mathbf{E}_r$, and suppose the boundary point of $\mathbf{E}_r$ where this is a tangent plane is η_r. Then $\mathbf{E}_{r+1} = \mathbf{E}(x^{r+1}, A_{r+1})$ is the minimum volume ellipsoid that contains $\mathbf{E}_r \cap \{x : ax \leqq d\}$, the shaded region in Figure 8.3, it has η_r as a boundary point and has the same tangent plane at η_r as $\mathbf{E}_r$. From the manner in which the inequality "$ax \leqq d$" is selected, it is clear that if $\mathbf{E}_r \supset \mathbf{E}_1 \cap \mathbf{K}$, then $\mathbf{E}_{r+1} \supset \mathbf{E}_1 \cap \mathbf{K}$. Arguing inductively on r, we conclude that every ellipsoid $\mathbf{E}_r$ constructed during the algorithm satisfies $\mathbf{E}_r \supset \mathbf{E}_1 \cap \mathbf{K}$. From Theorem 8.6, the volume of $\mathbf{E}_1 \cap \mathbf{K}$ is $\geqq 2^{-\mathbf{4n(n+1)^2(L_1+1)}}$. From the results in Chapter 15 of [2.26] we know that the volume of $\mathbf{E}_r$ gets multiplied by a factor of $e^{-\frac{1}{2(n+1)}}$ or less, after each step in the ellipsoid method. $\mathbf{E}_1$ is a ball whose radius is $(\varepsilon + \sqrt{\frac{b^T b}{4}})$, and $b^T b < 2^{\mathbf{2L_1}}$. So the volume of $\mathbf{E}_1$ is at most $2^{\mathbf{2nL_1}}$. The algorithm terminates in step r, if the center x^r satisfies (8.5), (8.6) and that is, it is a point in $\mathbf{E}_1 \cap \mathbf{K}$. If termination does not occur up to step $N = 8(n+1)^4(L_1+1)$, the volume of $\mathbf{E}_N$ is at most $2^{\mathbf{2L_1 n}} e^{-\frac{N}{2(n+1)}} < 2^{-\mathbf{4n(n+1)^2(L_1+1)}}$. From the fact that the volume of $\mathbf{E}_1 \cap \mathbf{K} > 2^{-\mathbf{4n(n+1)^2(L_1+1)}}$ this is a contradiction to $\mathbf{E}_N \supset \mathbf{E}_1 \cap \mathbf{K}$. So for some $r \leqq N$, we will have $x^r \in \mathbf{E}_1 \cap \mathbf{K}$, and in that step the ellipsoid method terminates. The validity of the remaining portion of the algorithm follows from Theorem 8.7, 8.8, 2.9. Since the ellipsoid method terminates after at most $N = 8(n+1)^4(L_1+1)$ steps, the algorithm is obviously polynomially bounded.

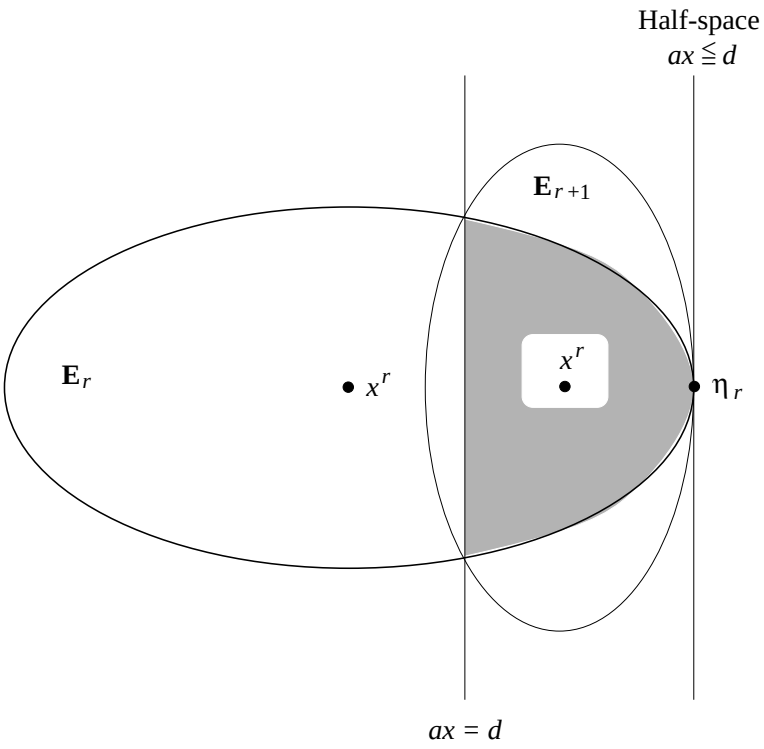


Figure 8.3 Construction of the new ellipsoid $\mathbf{E}_{r+1}$

In practice, it is impossible to run the algorithm using exact arithmetic. To run the algorithm using finite precision arithmetic, all computations have to be carried out

to a certain number of significant digits as discussed in [8.13], and the ellipsoid have to be expanded by a small amout in each iteration (this is achieved by multiplying the matrix A_r in each step by a number slightly larger than one in each step). As pointed out in [2.26] if each quantity is computed correct to $61nL_1$ bits of precision, and D_{r+1} multiplied by $(1 + \frac{1}{16n^2})$ before being rounded, all the results continue to hold.

Computational Comparison

Y. Fathi [8.10] did a comparative study in which this ellipsoid algorithm has been compared with the algorithm discussed in Chapter 7 for the nearest point problem. We provide a summary of his results here. In the study the matrix B was generated randomly, with its entries to be integers between -5 and $+5$. The b-vector was also generated randomly with its entries to be integers between -20 and $+20$. Instead of using computer times for the comparison, he counted the number of iterations of various types and from it estimated the total number of multiplication and division operations required before termination on each problem. Problems with $n = 10$, 20, 30, 40, 50 were tried and each entry in the table is an average for 50 problems. Double precision was used. It was not possible to take the values of ε and δ as small as those recomended in the algorithm. Mostly he tried $\varepsilon, \delta = 0.1$ (the computational effort before termination in the ellipsoid algorithms reported in the table below refers to $\varepsilon, \delta = 0.1$), and with this, sometimes the complementary basic vector obtained at termination of the algorithm turned out to be infeasible (this result is called an **unsuccessful run**). He noticed that if the values of these tolerances were decreased, the probability of an unsuccessful run decreases; but the computational effort required before termination increases very rapidly.

	Average Number of Multiplication and Division Operations Required Before Termination in	
n	The Algorithm of Chapter 7	The Ellipsoid Algorithm
10	Too small	33,303
20	16,266	381,060
30	42,592	1,764,092
40	170,643	5,207,180
50	324,126	11,286,717

These empirical results suggest that the ellipsoid algorithm cannot compete with the algorithm discussed in Chapter 7 for the nearest problem, in practical efficiency. The same comment seems to hold for the other ellipsoid algorithms discussed in the following sections.

8.5 An Ellipsoid Algorithm for LCPs Associated with PD Matrices

In this section $M = (m_{ij})$ denotes a given PD matrix of order n (symmetric or not) with integer entries, and $q = (q_i)$ denotes a given nonzero integer column vector in $\mathbf{R}^n$. We consider the LCP (q, M).

Definitions

Let ε be a small positive number. Later on we specify how small ε should be. Let

$$\begin{aligned}
\mathbf{K} &= \{z:\ Mz + q \geqq 0,\ z \geqq 0\}.\\
(\overline{w} = M\bar{z} + q, \bar{z}) &= \text{unique solution of the LCP } (q, M).\\
f(z) &= z^T(Mz + q).\\
\mathbf{E} &= \{z:\ f(z) \leqq 0\}.\\
\text{Bd}(\mathbf{E}) &= \text{Boundary of } \mathbf{E} = \{z:\ f(z) = 0\}.\\
L &= \left\lceil (1 + \log_2 n) + \sum_{i,j}\big(1 + \log_2(|m_{ij}| + 1)\big) + \sum_i \big(1 + \log_2(|q_i| + 1)\big)\right\rceil\\
\mathbf{E}_\varepsilon &= \{z:\ z^T(Mz + q) \leqq \varepsilon\} \text{ for } \varepsilon > 0.\\
\mathbf{E}_0 &= \{z:\ z^T z \leqq 2^{\mathbf{2L}}\}.
\end{aligned}$$

Since M is a PD matrix, $\mathbf{E}$ defined above is an ellipsoid.

Some Preliminary Results

Theorem 8.9 *The set* $\mathbf{K} = \{z : Mz + q \geqq 0, z \geqq 0\}$ *has nonempty interior.*

Proof. Remembering that M is a PD matrix, the proof of this theorem is similar to the proof of Theorem 8.1 of Section 8.4.

□

Theorem 8.10 $\mathbf{E} \cap \mathbf{K} = \text{Bd}(\mathbf{E}) \cap \mathbf{K} = \{\bar{z}\}$.

Proof. This follows directly from the definitions.

□

Theorem 8.11 *$\bar{z}$ is an extreme point of* $\mathbf{K}$. *Also, every extreme point z of* $\mathbf{K}$ *other than $\bar{z}$ satisfies* $f(z) > 2^{-\mathbf{2L}}$.

Proof. Since $(\overline{w}, \bar{z})$ is a BFS of: $w - Mz = q$, $w \geqq 0$, $z \geqq 0$; $\bar{z}$ is an extreme point of $\mathbf{K}$. Also, L is the size of this system. Since $(\overline{w}, \bar{z})$ is the unique solution of the LCP (q, M), at every extreme point z of $\mathbf{K}$ other than $\bar{z}$, we must have $f(z) > 0$. Using arguments similar to these in Theorem 8.4 of Section 8.4, we conclude that for each i, either z_i is 0 or $> 2^{-\mathbf{L}}$, and $M_{i.}z + q_i$ is 0 or $> 2^{-\mathbf{L}}$, at every extreme point z of $\mathbf{K}$. Combining these results we conclude that every extreme point z of $\mathbf{K}$ other than $\bar{z}$ satisfies $f(z) > 2^{-\mathbf{2L}}$.

□

Theorem 8.12 *For $0 < \varepsilon \leqq 2^{-2L}$, the n-dimensional volume of $\mathbf{E}_0 \cap \mathbf{E}_\varepsilon \cap \mathbf{K}$ is $\geqq \varepsilon^n 2^{-3(n+1)L}$.*

Proof. Obviously $\bar{z} \in \mathbf{E}_\varepsilon \cap \mathbf{K}$, and by Theorem 8.11, no other extreme point z of $\mathbf{K}$ lies in $\mathbf{E}_\varepsilon \cap \mathbf{K}$ for $0 < \varepsilon \leqq 2^{-2L}$. So for every value of ε in the specified range, every edge of $\mathbf{K}$ through $\bar{z}$ intersects $\mathbf{E}_\varepsilon$. Also, since $\mathbf{K}$ has a nonempty interior by Theorem 8.9, $\mathbf{E}_\varepsilon \cap \mathbf{K}$ has a positive n-dimensional volume, $\mathbf{K}$ might be unbounded, but by the results in Chapter 15 of [2.26], at every extreme point of $\mathbf{K}$, both z_i and $M_i.z + q_i$ are $\leqq \frac{2^L}{n}$ for each i. Let $\widehat{\mathbf{K}} = \{z : 0 \leqq z_j \leqq \frac{2^L}{n}, 0 \leqq M_j.z + q_j \leqq \frac{2^L}{n}$, for $j = 1$ to $n\}$. By the above facts, every edge of $\widehat{\mathbf{K}}$ through z is either an edge $\mathbf{K}$ (if it is a bounded edge of $\mathbf{K}$), or a portion of an edge of $\mathbf{K}$ (if it is an unbounded edge of $\mathbf{K}$). Let $z^1, \ldots, z^n$ be adjacent extreme points of $\bar{z}$ in $\widehat{\mathbf{K}}$, such that $\{\bar{z} : z^1, \ldots, z^n\}$ is affinely independent. The above facts imply that all these points $\bar{z}$, z^t, $t = 1$ to n are in $\mathbf{E}_0$. Since M is PD, $f(z)$ is convex. Let $\lambda = \varepsilon 2^{-2L}$. So for each $t = 1$ to n, $f(\bar{z} + \lambda(z^t - \bar{z})) \leqq (1-\lambda)f(\bar{z}) + \lambda f(z^t) = \lambda f(z^t) = \lambda \sum_{i=1}^n z_i^t (M_i.z^t + q_i) \leqq \lambda \sum_{i=1}^n \left(\frac{2^L}{n}\right)\left(\frac{2^L}{n}\right) \leqq \varepsilon$. This implies that the line segment $[\bar{z}, \bar{z} + \lambda(z^t - \bar{z})]$ completely lies inside $\mathbf{E}_0 \cap \mathbf{E}_\varepsilon \cap \mathbf{K}$. So the volume of $\mathbf{E}_0 \cap \mathbf{E}_\varepsilon \cap \mathbf{K} \geqq$ the volume of the simplex whose vertices are $\bar{z}$, $\bar{z} + \lambda(z^t - \bar{z})$, $t = 1$ to n, which is

$$\begin{aligned} &= \frac{1}{n!} \left| \text{ determinant of } \begin{pmatrix} \lambda(z^1 - \bar{z}) & \vdots & \ldots & \vdots & \lambda(z^t - \bar{z}) \end{pmatrix} \right| \\ &\geqq \lambda^n 2^{-(n+1)L}, \text{ by results similar to those in the proof of Theorem 8.6} \\ &\geqq \varepsilon^n 2^{-(3n+1)L} . \end{aligned}$$

□

Theorem 8.13 *Let $\varepsilon_0 = 2^{-(6L+1)}$. For any point $\hat{z} \in \mathbf{E}_0 \cap \mathbf{E}_{\varepsilon_0} \cap \mathbf{K}$, we have:*

$$\begin{array}{rrcl} \textit{either} & \hat{z}_i \leqq \sqrt{\varepsilon_0} & < & 2^{-3L} \\ \textit{or} & M_i.\hat{z} + q_i \leqq \sqrt{\varepsilon_0} & < & 2^{-3L} . \end{array}$$

Proof. For any i, if both $\hat{z}_i$ and $M_i.\hat{z} + q_i$ are $> \sqrt{\varepsilon_0}$, then $\hat{z}(M\hat{z} + q) > \varepsilon_0$, contradiction to the fact that $\hat{z} \in \mathbf{E}_0 \cap \mathbf{E}_{\varepsilon_0} \cap \mathbf{K}$.

□

Theorem 8.14 *Let $\hat{z}$ by any point in $\mathbf{E}_0 \cap \mathbf{E}_{\varepsilon_0} \cap \mathbf{K}$. Define*

$$y_i = \begin{cases} w_i & \text{if } \hat{z}_i < 2^{-3L} \\ z_i & \text{if } \hat{z}_i \geqq 2^{-3L} . \end{cases}$$

Then $(y_1, \ldots, y_n)$ is a complementary feasible basic vector for the LCP (q, M).

Proof. Let $\mathbf{J}_1 = \{i : \hat{z}_i \geqq 2^{-3L}\}$, $\mathbf{J}_2 = \{i : \hat{z}_i < 2^{-3L}\}$. So $\mathbf{J}_1 \cap \mathbf{J}_2 = \emptyset$ and $\mathbf{J}_1 \cup \mathbf{J}_2 = \{1, \ldots, n\}$, and by Theorem 8.13, $M_i.\hat{z} + q_i < 2^{-3L}$ for $i \in \mathbf{J}_1$.

In [8.11] P. Gács and L. Lovász proved the following lemma :

Consider the system of constraints

$$A_i.x \leqq b_i, \quad i = 1 \text{ to } m \tag{8.8}$$

with integer data, and let l be the size of this system. Suppose $\hat{x}$ is a solution of

$$A_i.x \leqq b_i + 2^{-\mathbf{l}}, \quad i = 1 \text{ to } m$$

such that $A_i.x \geqq b_i$, $i = 1$ to k, and let $\{A_{p_1}., \ldots, A_{p_r}.\} \subset \{A_1., \ldots, A_k.\}$ be such that it is linearly independent and it spans $\{A_1., \ldots, A_m.\}$ linearly. Let $\bar{x}$ be any solution of the system of equations

$$A_{p_t}.x = b_{p_t}, \quad t = 1 \text{ to } r \ .$$

Then $\bar{x}$ is a solution (8.8). See also Chapter 15 in [2.26]. We will use this lemma in proving this theorem. Consider the system :

$$\begin{aligned} -M_i.z &\leqq q_i + 2^{-\mathbf{3L}}, \quad i = 1 \text{ to } n \\ -z_i &\leqq 0 + 2^{-\mathbf{3L}}, \quad i = 1 \text{ to } n \\ M_i.z &\leqq -q_i + 2^{-\mathbf{3L}}, \quad i \in \mathbf{J}_1 \\ z_i &\leqq 0 + 2^{-\mathbf{3L}}, \quad i \in \mathbf{J}_2 \ . \end{aligned} \tag{8.9}$$

We know that $\hat{z}$ solves this system and in addition $\hat{z}$ also satisfies $M_i.\hat{z} \geqq -q_i$, $i \in \mathbf{J}_1$ and $\hat{z} \geqq 0$, $i \in \mathbf{J}_2$. Also, since M is PD, the set $\{M_i. : i \in \mathbf{J}_1\} \cup \{I_{.i} : i \in \mathbf{J}_2\}$ is linearly independent and linearly spans all the row vectors of the constraint coefficient matrix of the system (8.9). From the lemma of P. Gács and L. Lovász mentioned above, these facts imply that if $\tilde{z}$ is a solution of the system of equations :

$$\begin{aligned} M_i.z &= -q_i, \ i \in \mathbf{J}_1 \\ z_i &= 0, \quad i \in \mathbf{J}_2 \end{aligned} \tag{8.10}$$

then $\tilde{z}$ also satisfies :

$$\begin{aligned} -M_i.z &\leqq q_i, \ i = 1 \text{ to } n \\ -z_i &\leqq 0, \ i = 1 \text{ to } n \end{aligned}$$

So $\tilde{z} \geqq 0$, $\widetilde{w} = M\tilde{z} + q \geqq 0$ and since $\tilde{z}_i = 0$ for $i \in \mathbf{J}_2$ and $M_i.z + q_i = 0$ for $i \in \mathbf{J}_1$ we have $f(\tilde{z}) = 0$ (since $\mathbf{J}_1 \cap \mathbf{J}_2 = \emptyset$ and $\mathbf{J}_1 \cup \mathbf{J}_2 = \{1, \ldots, n\}$). So $(\widetilde{w}, \tilde{z})$ is the solution of the LCP (q, M). Since $\tilde{z}$ is the solution of (8.10), $(\widetilde{w}, \tilde{z})$ is the BFS of the system: $w - Mz = q$, $w \geqq 0$; $z \geqq 0$; corresponding to the basic vector y. So y is a complementary feasible basic vector for the LCP (q, M).

□

The Algorithm

Fix $\varepsilon = \varepsilon_0 = 2^{-(\mathbf{6L+1})}$. So $\mathbf{E}_0 = \mathbf{E}(0, 2^{\mathbf{2L}}I)$. Define $N = 2(n+1)^{\mathbf{2}}(11L+1)$ in this section. With $z^0 = 0$, $A_0 = 2^{\mathbf{2L}}I$, $\mathbf{E}(z^0, A_0)$ go to Step 1.

General Step $r+1$ **:** Let z^r, A_r, $\mathbf{E}_r = \mathbf{E}(z^r, A_r)$; be respectively the center, PD symmetric matrix, and the ellipsoid at the beginning of this step. If z^r satisfies :

$$\begin{aligned} -Mz - q &\leqq 0 \\ -q &\leqq 0 \end{aligned} \tag{8.11}$$

$$z^T(Mz+q) \leqq \varepsilon \tag{8.12}$$

terminate the ellipsoid algorithm, call z^r as $\hat{z}$ and go to the **final step** described below. If z^r violates (8.11), select a constraint in it that it violates most, breaking ties arbitrarily, and suppose it is "$az \leqq d$". If z^r satisfies (8.11) but violates (8.12), let ξ^r be the point of intersection of the line segment joining the center of the ellipsoid $\mathbf{E}_{\varepsilon_0}$ (this is, $z' = -\left(\frac{M+M^T}{2}\right)^{-1}\left(\frac{q}{2}\right)$) and z^r with the boundary $\mathbf{E}_{\varepsilon_0}$. Therefore $\xi^r = \lambda z' + (1-\lambda)z^r$, where λ is the positive root of the equation $(\lambda z' + (1-\lambda)z^r)^T M(\lambda z' + (1-\lambda)z^r) + q = \varepsilon_0$. Let $az = d$ by the equation of the tangent hyperplane to $\mathbf{E}_{\varepsilon_0}$ at ξ^r, where the equation is written such that the half-space $az \leqq d$ does not contain z^r. Define γ_{r+1}, A_{r+1}, as in (8.7) and

$$z^{r+1} = z^r - \left(\frac{1-\gamma_r n}{1+n}\right)\left(\frac{A_r a^T}{\sqrt{aA_r a^T}}\right)$$

With z^{r+1}, A_{r+1}, $\mathbf{E}_{r+1} = \mathbf{E}(z^{r+1}, A_{r+1})$, move to the next step in the ellipsoid algorithm.

After at most N steps, this ellipsoid algorithm will terminate with the point z^r in the terminal step lying in $\mathbf{E}_0 \cap \mathbf{E}_{\varepsilon_0} \cap \mathbf{K}$. Then go to the final step described below.

Final Step: Let the center of the ellipsoid in the terminal step by $\hat{z}$. Using $\hat{z}$, find the complementary BFS as outlined in Theorem 8.14.

Proof of the Algorithm

The updating formulas used in this ellipsoid algorithm are the same as those used in the algorithm of Section 8.4. Hence using the same arguments as in Section 8.4, we can verify that $\mathbf{E}_r \supset \mathbf{E}_0 \cap \mathbf{E}_{\varepsilon_0} \cap \mathbf{K}$ for all r. The volume of $\mathbf{E}_0$ is $< 2^{\mathbf{2Ln}}$. After each step in the ellipsoid algorithm, the volume of the current ellipsoid $\mathbf{E}_r$ gets multiplied by a factor of $e^{-\frac{1}{2(n+1)}}$ or less. So if the ellipsoid algorithm does not terminate even after N steps, the volume of $\mathbf{E}_N \leqq e^{-(\mathbf{n+1})(\mathbf{11L+1})}2^{\mathbf{2Ln}} < 2^{-\mathbf{L(9n+1)-n}}$, contradiction to the fact that $\mathbf{E}_N \supset \mathbf{E}_0 \cap \mathbf{E}_{\varepsilon_0} \cap \mathbf{K}$ and Theorem 8.12. So for some $r \leqq N$, we will have $z^r \in \mathbf{E}_0 \cap \mathbf{E}_{\varepsilon_0} \cap \mathbf{K}$, and in that step the ellipsoid algorithm terminates. Hence the algorithm is obviously polynomially bounded.

Comments made in Section 8.4 about the precision of computation required, remain valid here also.

8.6 An Ellipsoid Algorithm for LCPs Associated with PSD Matrices

In this section we consider the LCP (q, M) where M denotes a given PSD matrix of order n (symmetric or not) with integer entries, and q denotes a given integer column vector in $\mathbf{R}^n$.

Definitions

Let $\mathbf{K}$, $\mathbf{E}$, Bd($\mathbf{E}$), L, $\mathbf{E}_\varepsilon$ be as defined in Section 8.5. Let $\mathbf{E}_0 = \{z : z^T z \leqq 2^{\mathbf{2(L+1)}}\}$. Since M is only PSD here, $\mathbf{K}$ may have no interior, in fact $\mathbf{K}$ may even be empty. Also $\mathbf{E}$, $\mathbf{E}_\varepsilon$ may not be ellispoids. Let $e_n = (1, \ldots, 1)^T \in \mathbf{R}^n$.

Some Preliminary Results

Theorem 8.15 *In this case the LCP (q, M) has a solution iff $\mathbf{K} \neq \emptyset$. If $\mathbf{K} \neq \emptyset$, there exists a solution, $(\overline{w}, \bar{z})$, to the LCP (q, M) where $\bar{z}$ is an extreme point of $\mathbf{K}$. When $\mathbf{K} \neq \emptyset$, the LCP (q, M) may have many solutions, but the set of all solutions is a convex set which is $\mathbf{E} \cap \mathbf{K} = \text{Bd}(\mathbf{E}) \cap \mathbf{K}$.*

Proof. Since M is PSD, the fact that (q, M) has a solution iff $\mathbf{K} \neq \emptyset$ follows from Theorem 2.1. When $\mathbf{K} \neq \emptyset$, the complementary pivot algorithm produces a solution $(\overline{w}, \bar{z})$, to the LCP (q, M) which is a BFS and this implies that $\bar{z}$ is an extreme point of $\mathbf{K}$. The set of all solutions of the LCP (q, M) is obviously Bd($\mathbf{E}$) $\cap$ $\mathbf{K}$, and from the definition of $\mathbf{K}$, and $\mathbf{E}$ here it is clear that in this case Bd($\mathbf{E}$) $\cap$ $\mathbf{K}$ = $\mathbf{E} \cap \mathbf{K}$, and since both $\mathbf{E}$ and $\mathbf{K}$ are convex sets ($\mathbf{E}$ is convex because M is PSD), this set is convex. □

Theorem 8.16 *When $\mathbf{K} \neq \emptyset$, $\mathbf{E}_0 \cap \mathbf{E}_\varepsilon \cap \mathbf{K}$ contains all the extreme points z of $\mathbf{K}$ such that $(w = Mz + q, z)$ is a solution of the LCP (q, M).*

Proof. By the results discussed in Chapter 15 of [2.26] if $(\overline{w}, \bar{z})$ is solution of (q, M) which is BFS, then $z \in \mathbf{E}_0$. The rest follows from Theorem 8.15. □

In this case $\mathbf{E}_0 \cap \mathbf{E}_\varepsilon \cap \mathbf{K}$ may not contain all the z which lead to solutions of the LCP (q, M), Theorem 8.16 only guarantees that $\mathbf{E}_0 \cap \mathbf{E}_\varepsilon \cap \mathbf{K}$ contains all the z which are extreme points of $\mathbf{K}$ that lead to solutions of (q, M). Since M is PSD, the set of solutions of the LCP (q, M) may in fact be unbounded and hence all of it may not lie in $\mathbf{E}_0$.

Theorem 8.17 *If z_i is positive in some solution of (q, M), then its complement w_i is zero in all solutions of (q, M). Similarly if w_i is positive in some solutions of (q, M), then z_i is zero in all solutions of (q, M).*

Proof. By Theorem 8.15, the set of all solutions of (q, M) is convex set. So if (w^1, z^1), (w^2, z^2) are two solutions of (q, M) satisfying the properties that $z_i^1 > 0$ and $w_i^2 > 0$, then the other points on the line segment joining (w^1, z^1), (w^2, z^2) cannot be solutions of (q, M) (because they violate the complementarity constraint $w_i z_i = 0$) contradicting the fact that the set of solutions of (q, M) is a convex set.

□

Theorem 8.18 *If $\tilde{z}$ is an extreme point of $\mathbf{K}$, for each i either $\tilde{z}_i = 0$ or $2^{-\mathbf{L}} \leqq \tilde{z}_i \leqq \frac{2^{\mathbf{L}}}{n}$. Also either $M_i.\tilde{z} + q_i$ is zero or $2^{-\mathbf{L}} \leqq M_i.\tilde{z} + q_i \leqq \frac{2^{\mathbf{L}}}{n}$. Also at every extreme point $\tilde{z}$ of $\mathbf{K}$ that does not lead to a solution of (q, M), we will have $f(\tilde{z}) = \tilde{z}^T(Mz + q) > 2^{-\mathbf{2L}}$.*

Proof. Similar to the proof of Theorem 8.11 in Section 8.5.

□

Theorem 8.19 $\mathbf{K} \neq \emptyset$ *iff the set of solutions of*

$$\begin{aligned} Mz + q &\geqq -2^{-\mathbf{10L}} e \\ z &\geqq -2^{-\mathbf{10L}} e \end{aligned} \tag{8.13}$$

has a nonempty interior.

Proof. By the results of P. Gács and L. Lovász in [8.11] (also see Chapter 15 in [2.26]), (8.13) is feasible iff $\mathbf{K} \neq \emptyset$. Also any point in $\mathbf{K}$ is an interior point of the set of feasible solutions of (8.13).

□

Let $\mathbf{K}_1$ denote the set of feasible solutions of (8.13).

Theorem 8.20 *Let $\varepsilon_0 = 2^{-(\mathbf{6L+1})}$. For any point $\hat{z} \in \mathbf{E}_0 \cap \mathbf{E}_{\varepsilon_0} \cap \mathbf{K}_1$, we have for each $i = 1$ to n, either $\hat{z}_i < 2^{-\mathbf{3L}}$, or $M_i.\hat{z} + q_i < 2^{-\mathbf{3L}}$.*

Proof. Suppose that $\hat{z}_i \geqq 2^{-\mathbf{3L}}$ and $M_i.\hat{z} + q_i \geqq 2^{-\mathbf{3L}}$. Since $\hat{z} \in \mathbf{E}_{\varepsilon_0}$, $\hat{z}^T(M\hat{z} + q) \leqq 2^{-(\mathbf{6L+1})}$. Then we have $\sum_{t=1, t\neq i}^{n} \hat{z}_t(M_t.\hat{z} + q_t) \leqq 2^{-(\mathbf{6L+1})} - 2^{-\mathbf{6L}} \leqq -2^{-(\mathbf{6L+1})}$. But from (8.13) and the definition of $\mathbf{E}_0$ we arrive at the contradiction $\sum_{t=1, t\neq i}^{n} \hat{z}_t(M_t.\hat{z} + q_t)$ $\geqq -(n-1)2^{-\mathbf{10L}}(2^{\mathbf{2L+1}} + 2^{\mathbf{L}}) > -2^{-(\mathbf{6L+1})}$.

□

Theorem 8.21 *Let $\varepsilon_0 = 2^{-(\mathbf{6L+1})}$. If $\mathbf{K} \neq \emptyset$, the n-dimensional volume of $\mathbf{E}_0 \cap \mathbf{E}_{\varepsilon_0} \cap \mathbf{K}_1$ is $\geqq 2^{-\mathbf{11nL}}$.*

Proof. Assume $\mathbf{K} \neq \emptyset$. So (q, M) has a solution. Let $(\overline{w}, \bar{z})$ be a complementary BFS of (q, M). So, by Theorem 8.16, $\bar{z} \in \mathrm{Bd}(\mathbf{E}) \cap \mathbf{K}$. For $\lambda > 0$ define the hypercube; $\mathbf{C}_\lambda = \{z : z \in \mathbf{R}^n, |z_j - \bar{z}_j| \leqq \frac{\lambda}{2} \text{ for all } j = 1 \text{ to } n\}$. Then, clearly, the n-dimensional volume of $\mathbf{C}_\lambda$ is λ^n. We will now prove that $\mathbf{C}_\lambda \subset \mathbf{K}_1 \cap \mathbf{E}_0 \cap \mathbf{E}_{\varepsilon_0}$ for $\lambda \leqq 2^{-\mathbf{11L}}$. Since the radius of $\mathbf{E}_0$ is $2^{\mathbf{L+1}}$, $\mathbf{C}_\lambda \subset \mathbf{E}_0$ by the definition of $\mathbf{C}_\lambda$ and the fact that $||\bar{z}|| < 2^{\mathbf{L}}$ from Theorem 8.18. Let $\hat{z}$ be any point in $\mathbf{C}_\lambda$. Since $\bar{z}_i \geqq 0$, $M_i.\bar{z} + q_i \geqq 0$ for all $i = 1$ to n, we have; $\hat{z}_i \geqq \bar{z}_i - \frac{\lambda}{2} \geqq -\frac{\lambda}{2} \geqq -2^{-\mathbf{10L}}$; $M_i.\hat{z} + q_i \geqq M_i.\bar{z} + q_i - \frac{\lambda}{2} \sum_{j=1}^{n} |m_{ij}| \geqq -2^{-(\mathbf{11L+1})} \times$

$2^{\mathbf{L}} \geqq -2^{-\mathbf{10L}}$. So $\mathbf{C}_\lambda \subset \mathbf{K}_1$. Also, since $\bar{z}^T(M\bar{z}+q) = 0$ (since $(\overline{w} = M\bar{z}+q, \bar{z})$ solves (q, M)), we have: $\hat{z}^T(M\hat{z}+q) = (\hat{z}-\bar{z})^T(M\bar{z}+q+M^T\bar{z}) + (\hat{z}-\bar{z})^T M(\hat{z}-\overline{z}) \leqq \frac{\lambda}{2}n(2^{\mathbf{L}} + 2^{\mathbf{L}}2^{\mathbf{L}}) + (\frac{\lambda}{2})^{\mathbf{2}}\sum_{i,j}|m_{ij}| \leqq 2^{-(\mathbf{11L+1})}n2^{\mathbf{2L+2}} + n^{\mathbf{2}}2^{\mathbf{L-2(11L+1)}} \leqq \varepsilon_0$. This implies that $\mathbf{C}_\lambda \subset \mathbf{E}_{\varepsilon_0}$. Hence $\mathbf{C}_\lambda \subset \mathbf{K}_1 \cap \mathbf{E}_0 \cap \mathbf{E}_{\varepsilon_0}$. Now letting $\lambda = 2^{-\mathbf{11L}}$, the volume of $\mathbf{C}_\lambda$ is $2^{-\mathbf{11L}}$, and these facts imply the theorem.

□

Let $\hat{z}$ be any point in $\mathbf{E}_0 \cap \mathbf{E}_{\varepsilon_0} \cap \mathbf{K}_1$. Define

$$\begin{aligned} \mathbf{J}_1^- &= \{i : M_i.\hat{z} + q_i \leqq 0\}\ , & \mathbf{J}_1^+ &= \{i : 0 < M_i.\hat{z} + q_i \leqq 2^{-\mathbf{3L}}\}\ , \\ \mathbf{J}_2^- &= \{i : \hat{z}_i \leqq 0\}\ , & \mathbf{J}_2^+ &= \{i : 0 < \hat{z}_i \leqq 2^{-\mathbf{3L}}\}\ . \end{aligned}$$

Then by Theorem 8.20, $\mathbf{J}_1^- \cup \mathbf{J}_1^+ \cup \mathbf{J}_2^- \cup \mathbf{J}_2^+ = \{1, \ldots, n\}$. Furthermore, $\hat{z}$ is a solution of :

$$\begin{aligned} -M_i.z &\leqq q_i + 2^{-\mathbf{3L}}, && i = 1 \text{ to } n \\ -z_i &\leqq 2^{-\mathbf{3L}}, && i = 1 \text{ to } n \\ M_i.z &\leqq -q_i + 2^{-\mathbf{3L}}, && \text{for } i \in \mathbf{J}_1^+ \\ z_i &\leqq 2^{-\mathbf{3L}}, && \text{for } i \in \mathbf{J}_2^+ \end{aligned} \tag{8.14}$$

Theorem 8.22 *Let $\hat{z}$ be any point in $\mathbf{E}_0 \cap \mathbf{E}_{\varepsilon_0} \cap \mathbf{K}_1$. Let I be the unit matrix of order n. Using the constructive procedure described by P. Gács and L. Lovász in [8.11] (see also Theorem 15.7, Chapter 15 of [2.26]) obtain a new solution, which we will denote by the same symbol $\hat{z}$, such that if $\mathbf{J}_1^-$, $\mathbf{J}_1^+$, $\mathbf{J}_2^-$, $\mathbf{J}_2^+$ are the index sets corresponding to this new $\hat{z}$, then the new $\hat{z}$ also satisfies (8.14), and there exists a linearly independent subset, $\mathbf{D} \subset \{M_i. : i \in \mathbf{J}_1^- \cup \mathbf{J}_1^+\} \cup \{I_i. : i \in \mathbf{J}_2^- \cup \mathbf{J}_2^+\}$ such that $\mathbf{D}$ spans linearly $\{M_i. : i = 1 \text{ to } n\} \cup \{I_i. : i = 1 \text{ to } n\}$. Furthermore, if $\bar{z}$ is a solution of :*

$$\begin{aligned} -M_i.z &= q_i, \text{ for } i \text{ such that } M_i. \in \mathbf{D} \\ z_i &= 0, \text{ for } i \text{ such that } I_i. \in \mathbf{D} \end{aligned}$$

then $(\overline{w} = M\bar{z}+q, \bar{z})$ is a solution of the LCP (q, M).

Proof. This theorem follows from the results of P. Gács and L. Lovász in [8.11] (or Theorem 15.7, Chapter 15 in [2.26]) applied on (8.14). We know that $\hat{z}$ satisfies :

$$\begin{aligned} -M_i.\hat{z} &\geqq q_i, && \text{for } i \in \mathbf{J}_1^- \\ M_i.\hat{z} &\geqq -q_i, && \text{for } i \in \mathbf{J}_1^+ \\ -\hat{z}_i &\geqq 0, && \text{for } i \in \mathbf{J}_2^- \\ \hat{z}_i &\geqq 0, && \text{for } i \in \mathbf{J}_2^+ \end{aligned}$$

By these results, $\bar{z}$ is a solution of

$$\begin{aligned} -Mz &\leqq q \\ -z &\leqq 0\ . \end{aligned}$$

Furthermore, $\bar{z}$ satisfies :

$$\begin{aligned} M_i.\bar{z} &= -q_i, && \text{for } i \in \mathbf{J}_1^- \cup \mathbf{J}_1^+ \\ \bar{z}_i &= 0, && \text{for } i \in \mathbf{J}_2^- \cup \mathbf{J}_2^+ \end{aligned}$$

by the spanning property of $\mathbf{D}$ and these results. Also, since $\{1, \ldots, n\}$ is the union of $\mathbf{J}_1^-$, $\mathbf{J}_1^+$, $\mathbf{J}_2^-$, $\mathbf{J}_2^+$, at least one of w_i or z_i is zero for each $i = 1$ to n. All these facts together clearly imply that $(\overline{w}, \bar{z})$ is a solution of the LCP (q, M).

□

The Algorithm

Apply the ellipsoid algorithm discussed in Section 8.5 to get a point $\hat{z}$ in $\mathbf{E}_0 \cap \mathbf{E}_{\varepsilon_0} \cap \mathbf{K}_1$, initiating the algorithm with $z^0 = 0$, $A_0 = 2^{\mathbf{2(L+1)}} I$, $\mathbf{E}_0 = \mathbf{E}(z^0, A_0)$. In this case $\mathbf{K}$ could be $\emptyset$. This could be recognized in the ellipsoid algorithm in two different ways. For any r, if the quantity γ_r in step r of the ellipsoid algorithm turns out to be $\leqq -1$, it is an indication that the set $\mathbf{E}_0 \cap \mathbf{E}_{\varepsilon_0} \cap \mathbf{K}_1 = \emptyset$, terminate, in this case $\mathbf{K} = \emptyset$ and the LCP (q, M) has no solution (for a proof of this see Chapter 15 of [2.26]). If $\gamma_r > -1$, compute x^{r+1}, A_{r+1} and continue. The volume of $\mathbf{E}_0$ here is $< 2^{\mathbf{2n(L+1)}}$, and if $\mathbf{K} \neq \emptyset$, the volume of $\mathbf{E}_0 \cap \mathbf{E}_{\varepsilon_0} \cap \mathbf{K}_1$ is $> 2^{-\mathbf{11nL}}$ by Theorem 8.21. Hence if $\mathbf{K} \neq \emptyset$, this ellipsoid algorithm will terminate in at most $2(n+1)^{\mathbf{2}}(13L+1)$ steps with a point $\hat{z} \in \mathbf{E}_0 \cap \mathbf{E}_{\varepsilon_0} \cap \mathbf{K}_1$. So, if the ellipsoid algorithm did not find a point in $\mathbf{E}_0 \cap \mathbf{E}_{\varepsilon_0} \cap \mathbf{K}_1$ even after $2(n+1)^{\mathbf{2}}(13L+1)$ steps, we can conclude that $\mathbf{K} = \emptyset$, that is, that the LCP (q, M) has no solution. On the other hand, if a point $\hat{z}$ in $\mathbf{E}_0 \cap \mathbf{E}_{\varepsilon_0} \cap \mathbf{K}_1$ is obtained in the ellipsoid algorithm, then using it, obtain a solution $(\overline{w}, \bar{z})$ of the LCP (q, M) as discussed in Theorem 8.22.

8.7 Some NP-Complete Classes of LCPs

The ellipsoid algorithm discussed in Section 8.4, 8.5, 8.6 can only process LCPs associated with PSD matrices (the class of these LCP is equivalent to the class of convex quadratic programs). In [8.6, 8.15] it was shown that certain LCPs satisfying special properties can be solved as linear programs, and these LCPs are therefore polynomially solvable using the ellipsoid algorithm (see Chapter 15 in [2.26]) on the resulting linear programs.

For the general LCP, the prospects of finding a polynomially bounded algorithm are not very promising, in view of the result in [8.3] where it is shown that this problem is $\mathcal{NP}$-complete. See reference [8.12] for the definition of $\mathcal{NP}$-completeness. Let $a_1, \ldots, a_n, a_0$ be positive integers and let M_{n+2} and $q(n+2)$ be the following matrices :

$$M_{n+2} = \begin{pmatrix} -I_n & 0 & 0 \\ e_n^T & -n & 0 \\ -e_n^T & 0 & -n \end{pmatrix} \quad , \quad q(n+2) = \begin{pmatrix} a_1 \\ \vdots \\ a_n \\ -a_0 \\ a_0 \end{pmatrix}$$

where I_n denotes the unit matrix of order n, and e_n is the column vector in $\mathbf{R}^n$ all of whose entries are 1. Also consider the 0-1 equality constrained Knapsack feasibility problem :

$$\begin{aligned} &\sum_{i=1}^{n} a_i x_i = a_0 \\ &x_i = 0 \text{ or } 1 \quad \text{for all } i = 1 \text{ to } n \ . \end{aligned} \tag{8.15}$$

If $(\tilde{w}, \tilde{z})$ is a solution of the LCP $(q(n+2), M_{n+2})$, define $\tilde{x}_i = \frac{\tilde{z}_i}{a_i}$, $i = 1$ to n, and verify that $\tilde{x} = (\tilde{x}_1, \ldots, \tilde{x}_n)^T$ is a feasible solution of the Knapsack problem (8.15). Conversely of $\hat{x} = (\hat{x}_1, \ldots, \hat{x}_r)^T$ is a feasible solution of (8.15), define $\widehat{w}_{n+1} = \hat{z}_{n+1} = \widehat{w}_{n+2} = \hat{z}_{n+2} = 0$ and $\hat{z}_i = a_i \hat{x}_i$, $\widehat{w}_i = a_i(1 - \hat{x}_i)$, $i = 1$ to n; and verify that $(\widehat{w} = (\widehat{w}_1, \ldots, \widehat{w}_{n+2}), \hat{z} = (\hat{z}_1, \ldots, \hat{z}_{n+2}))$ is a solution of the LCP $(q(n+2), M_{n+2})$. Since the problem of finding whether a feasible solution for (8.15) exists is a well known $\mathcal{NP}$-complete problem (see [8.12]), the problem of checking whether the LCP $(q(n+2), M_{n+2})$ has a solution is $\mathcal{NP}$-complete. Also, since the matrix M_{n+2} is negative definite, the class of LCPs associated with negative definite or negative semidefinite matrices are $\mathcal{NP}$-hard. Also M_{n+2} is lower triangular. This shows that the class of LCPs associated with lower or upper triangular matrices is $\mathcal{NP}$-hard, if negative entries appear in the main diagonal.

Let M be a given negative definite matrix with integer entries, and let $q \in \mathbf{R}^n$ be a given integer column vector. In this case the LCP (q, M) may not have a solution; and even if it does, the solution may not be unique. From the results in Chapter 3 we know that the number of distinct solutions of the LCP (q, M) in this case is finite. Define :

$$\mathbf{K} = \{z : \ z \geqq 0, \ Mz + q \geqq 0\}$$
$$\mathbf{E} = \{z : \ z^T(Mz + q) \geqq 0\}$$

Since M is negative definite, $\mathbf{E}$ is an ellipsoid. Let Bd($\mathbf{E}$) = boundary of $\mathbf{E} = \{z : z^T(Mz + q) = 0\}$.

Clearly any point $z \in \text{Bd}(\mathbf{E}) \cap \mathbf{K}$ satisfies the property that $(w = Mz + q, z)$ is a solution of the LCP (q, M) and vice versa. So solving the LCP (q, M) is equivalent to the probem of finding a point in Bd($\mathbf{E}$) $\cap$ $\mathbf{K}$. However, in this case $\mathbf{K} \subset \mathbf{E}$, and in general, Bd($\mathbf{E}$) $\cap$ $\mathbf{K} \subset \mathbf{E} \cap \mathbf{K}$. See Figure 8.4. So the nice property that $\mathbf{E} \cap \mathbf{K}$ = Bd($\mathbf{E}$) $\cap$ $\mathbf{K}$ which held for LCPs associated with PSD matrices does not hold here anymore, which makes the LCP associated with a negative definite matrix much harder. In this case (i. e., with M being negative definite), it is possible to find a point in $\mathbf{E} \cap \mathbf{K}$ using an ellipsoid algorithm (actually since $\mathbf{K} \subset \mathbf{E}$ here, a point in $\mathbf{K}$ can be found by the ellipsoid algorithm of Chapter 15 of [2.26] and that point will also lie in $\mathbf{E}$), but the point in $\mathbf{E} \cap \mathbf{K}$ obtained by the algorithm may not be on the boundary of $\mathbf{E}$, and hence may not lead to a solution of the LCP (q, M). In fact, finding a point in Bd($\mathbf{E}$) $\cap$ $\mathbf{K}$ is a concave minimization problem, and that's why it is $\mathcal{NP}$-hard.

The status of the LCPs (q, M) where M is a P-but not PSD matrix, is unresolved. In this case the LCP (q, M) is known to have a unique solution by the results in Chapter 3, but the sets $\{z : z^T(Mz + q) \leqq 0\}$ are not ellipsoids. The interesting question is whether a polynomially bounded algorithm exists for solving this special

class of LCPs. This still remains an open question. It is also not known whether these LCPs are $\mathcal{NP}$-hard.

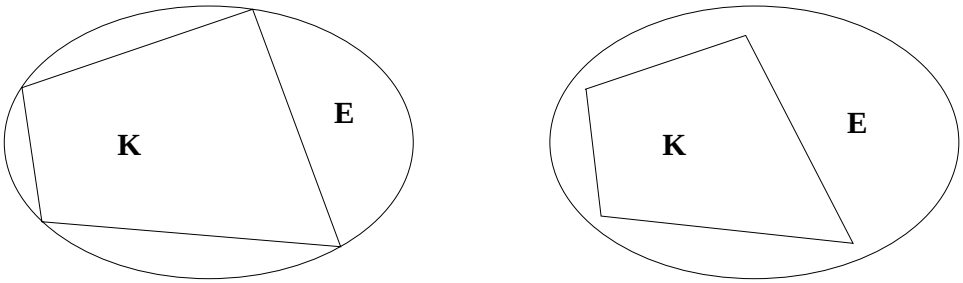


Figure 8.4 When M is negative definite, **E** and **K** may be as in one of the figures given here. Points of **K** on the boundary of **E**, if any, lead to solutions of the LCP (q, M).

8.8 An Ellipsoid Algorithm for Nonlinear Programming

In [8.9] J. Ecker and M. Kupferschmid discussed an application of the ellipsoid algorithm to solve NLPs of the following form :

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leqq 0, \quad i = 1 \text{ to } m \end{array}$$

where all the $f_i(x)$ are differentiable functions defined on $\mathbf{R}^n$, and we assume that $n > 1$.

For the convergence of the ellipsoid algorithm, we need to specify an initial ellipsoid whose intersection with a neighborhood of an optimum solution has positive n-dimensional volume. This requirement prevents the algorithm from being used in a simple way for problems having equality constraints, but the penalty transformation discussed in Section 2.7.6 can be used for them.

It is assumed that lower and upper bounds are avaible on each variable. l, u are these lower and upper bound vectors. The initial ellispoid is chosen to be the one of smallest volume among those ellipsoids with center $x^0 = \frac{l+u}{2}$ and containing $\{x : l \leqq x \leqq u\}$. Let this be $\mathbf{E}_0 = \{x : (x - x^0)^T D_0^{-1} (x - x^0) \leqq 1\} = \mathbf{E}_0(x^0, D_0)$, where

$$D_0 = \frac{n}{4} \begin{pmatrix} (u_1 - l_1)^2 & 0 & 0 & \dots & 0 \\ 0 & (u_2 - l_2)^2 & 0 & \dots & 0 \\ 0 & 0 & \ddots & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & (u_n - l_n)^2 \end{pmatrix} .$$

Suppose we have $\mathbf{E}_r(x^r, D_r)$. If x^r is infeasible, choose a violated constraint, say the i^{th}, where $f_i(x^r) > 0$. In case x^r is infeasible, the index i of the selected constraint is that of the first violated constraint encountered under a search of the constraints in cyclical order beginning with the constraint selected in the previous step. If x^r is feasible and $\nabla f_0(x^r) = 0$, terminate, x^r is optimal to NLP (under convexity assuptions, it is a stationary point otherwise). If x^r is feasible and $\nabla f_0(x^r) \neq 0$, choose the index i to be zero.

Having selected the index i (corresponding to a violated constraint if x^r is infeasible, or the objective function if x^r is feasible and $\nabla f_0(x^r) \neq 0$), let $\mathbf{H}_r$ be the hyperplane

$$\mathbf{H}_r = \{x : \ -(\nabla f_i(x^r))(x - x^r) = 0\} \ .$$

The hyperplane $\mathbf{H}_r$ supports the contour $f_i(x) = f_i(x^r)$ and divides the ellipsoid in half. The center x^{r+1} of the next ellipsoid $\mathbf{E}_{r+1}$ and the PD matrix D_{r+1} used in defining $\mathbf{E}_{r+1}$ are determined by the updating formulae

$$\begin{aligned} h &= \frac{\nabla f_i(x^r)}{||\nabla f_i(x^r)||} \\ d &= \frac{-D_r h^T}{+\sqrt{h D_r h^T}} \\ x^{r+1} &= x^r + \frac{d}{n+1} \\ D_{r+1} &= \frac{n^2}{n^2 - 1}\left(D_r - \frac{2}{n+1} d d^T\right) \ . \end{aligned}$$

The best point obtained during the algorithm and its objective value are maintained. Various stopping rules can be employed, such as requiring the difference between successive best values to be sufficiently small, etc.

The method is best suited for solving the NLP above, when all the functions $f_i(x)$ are convex. If a nonconvex function is used to generate the hyperplane H_r that cuts $\mathbf{E}_r$ in half, the next ellipsoid may not contain the optimal point, and the algorithm may converge to a point that is not even stationary.

In computational tests carried out by J. G. Ecker and M. Kupferschmid [8.9], this method performed very well.

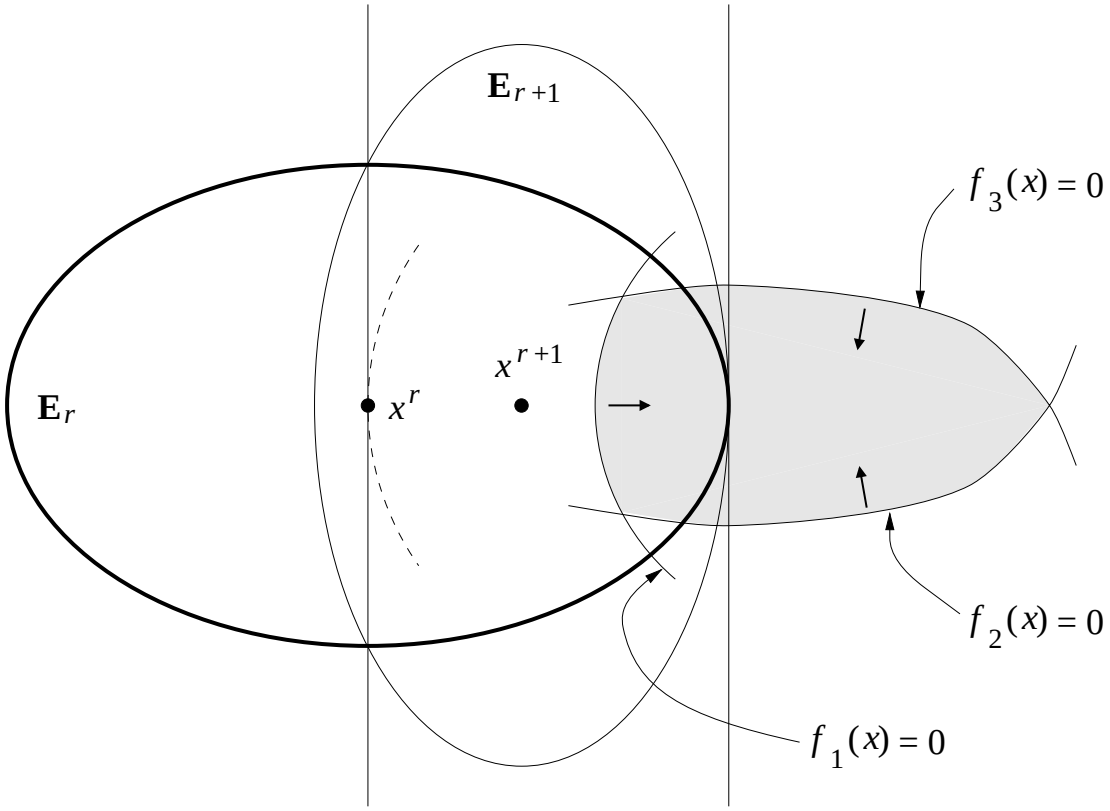


Figure 8.5 Construction of the new ellipsoid when x^r is infeasible. The arrow on constraint surface $f_i(x) = 0$ indicates the feasible side, that is satisfying $f_i(x) \leqq 0$. $f_1(x) \leqq 0$ is violated at x^r and is selected.

8.9 Exercises

8.4 Let A, D, b, d be given matrices of orders $m_1 \times n$, $m_2 \times n$, $m_1 \times 1$, $m_2 \times 1$ respectively with integer entries. Let F be a given PD symmetric matrix of order n with integer entries. Define.

$$\mathbf{K}_1 = \{x : \ Ax \geqq b\}$$
$$\mathbf{K}_2 = \{x : \ Dx \geqq d\}$$
$$\mathbf{E} = \{x : \ x^T F x \leqq 1\} \ .$$

Construct polynomially bounded algorithms for checking whether

(i) $\mathbf{K}_1 \subset \mathbf{K}_2$

(ii) $\mathbf{E} \subset \mathbf{K}_1$.

Does a polynomially bounded algorithm exists for checking whether $\mathbf{K}_1 \subset \mathbf{E}$? Why ?

8.5 Consider the quadratic program

$$\begin{array}{ll} \text{minimize} & cx + \frac{1}{2} x^T D x \\ \text{subject to} & x \leqq b \end{array}$$

where $b > 0$ and D is a Z-matrix of order n. Express the KKT optimality conditions for this problem in the form of a special type of linear complementarity problem, and develop a special direct method for solving it, based on Chandrasekaran's algorithm discussed in Section 8.1.
(J. S. Pang [8.17])

8.6 Study the computational complexity of the problem of checking whether the ellipsoid $\mathbf{E} = \{x : (x - \bar{x})^T D(x - \bar{x}) \leqq 1\}$ where D is given integer PD symmetric matrix and $\bar{x}$ is a given noninteger rational point, contains an integer point.

8.7 Show that the LCP (q, M) is equivalent to the following piecewise linear concave function minimization problem.

$$\begin{array}{ll} \text{minimize} & \sum_{j=1}^{n} (\text{minimum}\{0, M_j.z - z_j + q_j\} + z_j) \\ \text{subject to} & Mz + q \geqq 0 \\ & z \geqq 0. \end{array}$$

(O. L. Mangasarian)

8.10 References

8.1 I. Adler, R. P. McClean and J. S. Proven, "An Application of the Khachiyan-Shor Algorithm to a Class of Linear Complementarity Problems", Cowles Foundation Discussion Paper no. 549, Yale University, New Haven, Connecticut 06520, USA (1981).

8.2 R. Chandrasekaran, "A Special Case of the Complementary Pivot Problem", *Opsearch*, 7 (1970) 263–268.

8.3 S. J. Chung, "A Note on the Complexity of LCP: The LCP is Strongly $\mathcal{NP}$-complete", Technical Report 79–2, Department of Industrial and Operations Engineering, University of Michigan, Ann Arbor, Michigan 48109 (1979).

8.4 S. J. Chung and K. G. Murty, "Polynomially Bounded Ellipsoid Algorithms for Convex Quadratic Programming", in O. L. Mangasarian, R. R. Meyer, S. M. Robinson (editors), *Nonlinear Programming 4*, Academic Press, (1981) 439–485.

8.5 R. W. Cottle and M. S. Goheen, "A Special Class of Large Quadratic Programs", in O. L. Mangasarian, R. R. Meyer and S. M. Robinson eds., *Nonlinear Programming 3*, Academic Press (New York 1978) 361–390.

8.6 R. W. Cottle and J. S. Pang, "On Solving Linear Complementarity Problems as Linear Programs", *Mathematical Programming Study 7*, (1978) 88–107.

8.7 R. W. Cottle and J. S. Pang, "A Least-Element Theory of Solving Linear Complementarity Problems as Linear Programs", *Mathematics of Operations Research*, 3, 2, pp. 155–170, May 1978.

8.8 R. W. Cottle and R. S. Sacher, "On the Solution of Large Structured Linear Complementarity Problems: The Tridiagonal Case", *Journal of Applied Mathematics and Optimization*, 3 (1977) 321–340.

8.9 J. G. Ecker and M. Kupferschmid, "An Ellipsoid Algorithm for Nonlinear Programming", *Mathematical Programming*, 27 (1983) 83–106.

8.10 Y. Fathi, "Comparative Study of the Ellipsoid Algorithm and Other Algorithms for the Nearest Point Problem", Department of Industrial and Operations Engineering, The University of Michigan, Ann Arbor, Michigan 48109, USA (1980).

8.11 P. Gács and L. Lovász, "Khacian's Algorithm for Linear Programming", *Mathematical Progamming Study 14*, (1981) 61–68.

8.12 M. R. Garey and D. S. Johnson, "*Computers and Intractability, A Guide to the Theory of $\mathcal{NP}$-completeness*", W. H. Freeman and Company, New York (1979).

8.13 L. G. Hacijan, "A Polynomial Algorithm in Linear Programming", *Soviet Mathematics Doklady*, 20 (1979) 191–194.

8.14 M. K. Kozlov, S. P. Taransov and L. G. Hachijan, "Polynomial Solvability of Convex Quadratic Programming", *Soviet Mathematics Doklady*, 20 (1979).

8.15 O. L.Mangasarian, "Linear Complementarity Problem Solvable by a Single Linear Program", *Mathematical Programming* 10 (1976) 263–270.

8.16 O. L. Mangasarian, "Simplified Characterizations of Linear Complementarity Problems Solvable as Linear Programs", *Mathematics of Operations Research* 4 (1979) 268–273.

8.17 J. S. Pang, "On a Class of Least-Element Complementarity Problems", *Mathematical Programming* 16 (1979) 111-126.

8.18 J. S. Pang and R. Chandrasekaran, "Linear Complementarity Problems Solvable by a Polynomially Bounded Pivoting Algorithm", *Mathematical Programming Study* 25 (1985) 13–27.

8.19 R. Saigal, "A Note on a Special Linear Complementarity Problem", *Opsearch 7* (1970) 175–183.

8.20 R. Saigal, "The Lexicographic Variant Solves Certain Random Linear Programs in Average Steps Bounded Above by a Linear Function of m", Department of Industrial Engineering, Northwestern University, Evanston, IL–60201, USA, December 1984.

Chapter 9

ITERATIVE METHODS FOR LCP's

9.1 Introduction

The name **iterative method** usually refers to a method that provides a simple formula for computing the $(r+1)^{th}$ point as an explicit function of the r^{th} point: $x^{r+1} = f(x^r)$. The method begins with an initial point x^0 (quite often x^0 can be chosen arbitrarity, subject to some simple constraints that may be specified, such as $x^0 \geqq 0$, etc.) and generates the sequence of points $\{x^0, x^1, x^2, \ldots\}$ one after the other using the above formula. The method can be terminated whenever one of the points in the sequence can be recognized as being a solution to the problem under consideration. If finite termination does not occur, mathematically the method has to be continued indefinitely. In some of these methods, it is possible to prove that the sequence $\{x^r\}$ converges in the limit to a solution of the problem under consideration, or it may be possible to prove that every accumulation point of the sequence $\{x^r\}$ is a solution of the problem. In practice, it is impossible to continue the method indefinitely. In such cases, the sequence is computed to some finite length, and the final solution accepted as an approximate solution of the problem.

In this chapter we consider the LCP (q, M) which is to find $w, z \in \mathbf{R}^n$ satisfying

$$\begin{aligned} w - Mz &= q \\ w, z &\geqq 0 \\ w^T z &= 0 \end{aligned} \tag{9.1}$$

where M, q are given matrices of orders $n \times n$ and $n \times 1$, respectively. We discuss several iterative methods for solving this LCP (q, M). All the methods that we have discussed so far for solving this problem (the pivotal methods and the ellipsoid methods) have the finite termination property. In contrast, the iterative methods discussed here do not in general terminate in a finite number of steps (even though the special structure of the problem discussed in Section 9.2, makes it possible to construct a modification of the iterative method discussed there that terminates after a finite amount of work). However, these iterative methods have the advantage of being extremely simple and easy to program (much more so than all the methods discussed so far in this book) and hold promise for tackling very large problems that have no special structure (other than possibly symmetry and/or positive definiteness as required by the algorithm).

Most of the algorithms for solving nonlinear programming problems are iterative in nature (see references [10.9, 10.13, 10.33]) and the iterative methods discussed here can be interpreted as specializations of some nonlinear programming algorithms applied to solve a quadratic program equivalent to the LCP.

The word **sequence** here usually refers to an infinite sequence. An infinite sequence of points $\{x^r : r = 1, 2, \ldots\}$ in $\mathbf{R}^n$ is said to converge in the **limit** to the given point x^* if, for each $\varepsilon > 0$, there exists a positive integer N such that $\|x^r - x^*\| < \varepsilon$ for all $r \geqq N$. As an example the sequence in $\mathbf{R}^1$, $\{x^r :$ where $x^r = \frac{1}{r}, r \geqq 1$ and integer$\}$ converges to zero. However, the sequence $\{x^r :$ where $x^r = \frac{1}{r}$ if $r = 2s$ for some positive integer s, and $x^r = 1$ if $r = 2s + 1$ for some positive integer $s\}$ does not converge. A point $x^* \in \mathbf{R}^n$, is said to be a **limit point** or an **accumulation point** for the infinite sequence $\{x^r : r = 1, 2, \ldots\}$ of points in $\mathbf{R}^n$, if for every $\varepsilon > 0$ and positive integer N, there exists a positive integer $r > N$ such that $\|x^r - x^*\| < \varepsilon$. If x^* is a limit point of the sequence $\{x^r : r = 1, 2, \ldots\}$, then there exists a subsequence of this sequence, say $\{x^{r_k} : k = 1, 2, \ldots\}$, which converges in the limit to x^*, where $\{r_k : k = 1, 2, \ldots\}$ is a monotonic increasing sequence of positive integers. If the sequence $\{x^r : r = 1, 2, \ldots\}$ converges in the limit to x^*, then x^* is the only limit point for this sequence. A sequence that does not converge may have no limit point (for example, the sequence of positive integers in $\mathbf{R}^1$ has no limit point) or may have any number of limit points. As an example, consider the sequence of numbers in $\mathbf{R}^1$, $\{x^r :$ where $x^r = \frac{1}{r}$, if $r = 2s$ for some positive integer s, otherwise $x^r = 1 + \frac{1}{r}$, if $r = 2s + 1$ for some non-negative integer $s\}$. This sequence has two limit points, namely 0 and 1. The subsequence $\{x^{2s} : s = 1, 2, \ldots\}$ of this sequence converges to the limit point 0, while the subsequence $\{x^{2s+1} : s = 1, 2, \ldots\}$ converges to the limit point 1.

The discussion in this section also needs knowledge of some of the basic properties of compact subsets of $\mathbf{R}^n$. See [9.21].

9.2 An Iterative Method for LCPs Associated with PD Symmetric Matrices

The method discussed in this section is due to W. M. G. Van Bokhoven [9.22]. We consider the LCP (q, M) where M is assumed to be a PD symmetric matrix. For $q \geqq 0$, $(w = q, z = 0)$ is the unique solution of the LCP (q, M). So we only consider the case $q \ngeqq 0$. For any vector $x = (x_j) \in \mathbf{R}^n$ we denote by $|x|$ the vector $(|x_j|)$ in this section. The symbol I denotes the identity matrix of order n. We will now discuss the main result on which the method is based.

Theorem 9.1 *Let M be PD and symmetric. The LCP (q, M) is equivalent to the fixed point problem of determining $x \in \mathbf{R}^n$ satisfying*

$$f(x) = x \tag{9.2}$$

where $f(x) = b + B|x|$, $b = -(I + M)^{-1}q$, $B = (I + M)^{-1}(I - M)$.

Proof. In (9.1) transform the variables by substituing

$$w_j = |x_j| - x_j, \ z_j = |x_j| + x_j, \quad \text{for each } j = 1 \text{ to } n \tag{9.3}$$

We verify that the constraints $w_j \geqq 0$, $z_j \geqq 0$ for $j = 1$ to n automatically hold, from (9.3). Also substituing (9.3) in "$w - Mz - q = 0$", leads to $f(x) - x = 0$. Further, $w_j z_j = 0$ for each $j = 1$ to n, by (9.3). So any solution x of (9.2) automatically leads to a solution of the LCP (q, M) through (9.3). Conversely suppose (w, z) is the solution of the LCP (q, M). Then $x = \frac{1}{2}(z - w)$ can be verified to be the solution of (9.2). □

Some Matrix Theoretic Results

If A is square matrix of order n, its norm, dented by $\|A\|$, is defined to be the Supremum of $\{\frac{\|Ax\|}{\|x\|} : x \in \mathbf{R}^n, x \neq 0\}$. From this definition, we have $\|Ax\| \leqq \|A\| \cdot \|x\|$ for all $x \in \mathbf{R}^n$. See references [9.9, 9.10, 10.33].

Since M is symmetric and PD, all its eigenvalues are real and positive (see references [9.8, 9.9, 9.10, 10.33] for definition and results on eigenvalues of square matrices). If $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of M, then the eigenvalues of $B = (I + M)^{-1}(I - M)$ are given by $\mu_i = \frac{(1-\lambda_i)}{(1+\lambda_i)}$, $i = 1$ to n; and hence all μ_i are real and satisfy $|\mu_i| < 1$ for all i (since $\lambda_i > 0$). Since B is also symmetric we have $\|B\| = \text{Maximum}\{|\mu_i| : i = 1 \text{ to } n\} < 1$.

The Iterative Scheme for Solving (9.2)

The scheme begins with an initial point $x^1 \in \mathbf{R}^n$ chosen arbitrarily (say $x^1 = 0$). For $r \geqq 2$ define

$$x^{r+1} = f(x^r) = b + B|x^r| \ . \tag{9.4}$$

The equation (9.4) defines the iterative scheme. Beginning with the initial point $x^1 \in \mathbf{R}^n$ chosen arbitrarily, generate the sequence $\{x^1, x^2, \ldots\}$ using (9.4) repeatedly. This iteration is just the successive substitution method discussed in Section 2.7.2 for computing the Brouwer's fixed point of $f(x)$. We will now prove that the sequence generated $\{x^1, x^2, \ldots\}$ converges in the limit to the unique fixed point x^* of (9.2).

Convergence Theorems

Theorem 9.2 *When M is* PD *and symmetric, the sequence of points $\{x^r\}$ defined by (9.4) converges in the limit to x^*, the unique solution of (9.2), and the solution $(w^*, z*)$ of the LCP (q, M) can be obtained from x^* from the transformation (9.3).*

Proof. For any $x, y \in \mathbf{R}^n$ we have $\|f(x)-f(y)\| = \|B(|x|-|y|)\| \leqq \|B\| \cdot \|(|x|-|y|)\| < \|x-y\|$, since $\|(|x|-|y|)\| \leqq \|x-y\|$ and $\|B\| < 1$ as discussed above. So $f(x)$ is a contraction mapping (see reference [9.20]) and by Banach contraction mapping theorem the sequence $\{x^r\}$ generated by (9.4) converges in the limit to the unique solution x^* of (9.2). The rest follows from Theorem 9.1.

□

We will denote $\|B\|$ by the symbol ρ. We known that $\rho < 1$, and it can actually be computed by well known matrix theoretic algorithms.

Theorem 9.3 *If x^* is the unknown solution of (9.2), $\|x^*\| \geqq \frac{\|b\|}{(1+\rho)}$.*

Proof. From (9.2) $\|x^*\| = \|(b + B|x^*|)\| \geqq \|b\| - \|(B|x^*|)\| \geqq \|b\| - \rho\|x^*\|$. So $\|x^*\| \geqq \frac{\|b\|}{(1+\rho)}$.

□

Theorem 9.4 *Let x^r be the r^{th} point obtained in the iterative scheme (9.4) and let x^* be the unique solution of (9.2). Then for $r \geqq 1$, $\|x^* - x^{r+1}\| \leqq \left(\frac{\rho^{\mathbf{r}}}{1-\rho}\right)\|x^2 - x^1\|$.*

Proof. We have $x^* - x^{r+1} = f(x^*) - f(x^r)$. So $\|x^* - x^{r+1}\| = \|f(x^*) - f(x^r)\| \leqq \rho\|x^* - x^r\|$ (by the argument used in the proof of Theorem 9.2, since $\|B\| = \rho$). Applying the same argument repeatedly we get

$$\|x^* - x^{r+1}\| \leqq \rho^{\mathbf{r}}|x^* - x^1\| \ . \tag{9.5}$$

Now, for $r > 2$ we have $x^{r+1} - x^r = f(x^r) - f(x^{r-1})$. So we have $\|x^{r+1} - x^r\| = \|f(x^r) - f(x^{r-1})\| \leqq \rho\|x^r - x^{r-1}\|$. Using this argument repeatedly, we get

$$\|x^{r+1} - x^r\| \leqq \rho^{\mathbf{r-1}}\|x^2 - x^1\|, \text{ for } r > 2 \ . \tag{9.6}$$

We also have $x^* - x^1 = x^* - x^2 + (x^2 - x^1)$. So we have $\|x^* - x^1\| \leqq \|x^* - x^2\| + \|x^2 - x^1\|$. Using this same argument repeatedly, and the fact that the x^* = limit x^t as t tends to ∞, (and therefore limit $\|x^* - x^t\|$ as t tends to ∞ is 0), we get $\|x^* - x^1\| \leqq \sum_{t=1}^{\infty} \|x^{t+1} - x^t\| \leqq \|x^2 - x^1\|\left(\sum_{t=0}^{\infty} \rho^t\right)$ (from (9.6)) $= \frac{\|x^2 - x^1\|}{(1-\rho)}$. Using this in (9.5) leads to $\|x^* - x^{r+1}\| \leqq \left(\frac{\rho^{\mathbf{r}}}{1-\rho}\right)\|x^2 - x^1\|$ for $r \geqq 1$.

□

Theorem 9.5 *If $x^1 = 0$, we have $\|x^* - x^{r+1}\| \leqq \rho^{\mathbf{r}} \left(\frac{\|b\|}{1-\rho}\right)$.*

Proof. Follows from Theorem (9.4). □

Theorem 9.6 *If $x^1 = 0$, we have for $r \geqq 1$, $\|x^{r+1}\| \geqq \|b\| \left(\frac{1}{1+\rho} - \frac{\rho^{\mathbf{r}}}{1-\rho}\right)$.*

Proof. We know that $\|x^*\| - \|x^{r+1}\| \leqq \|x^* - x^{r+1}\|$. So $\|x^{r+1}\| \geqq \|x^*\| - \|x^* - x^{r+1}\|$. The result follows from this and Theorems 9.3, 9.5. □

How to Solve the LCP (q,M) in a Finite Number of Steps Using the Iterative Scheme (9.4)

Initiate the iterative scheme (9.4) with $x^1 = 0$. Then for $r > 1$ from Theorem 9.6, we know that there must exist an i satisfying

$$|x_i^{r+1}| \geqq \frac{\|b\|}{\sqrt{n}} \left(\frac{1}{1+\rho} - \frac{\rho^{\mathbf{r}}}{1-\rho}\right) . \tag{9.7}$$

But from Theorem 9.5, for the same i, we must have $|x_i^* - x_i^{r+1}| \leqq \|b\| \left(\frac{\rho^{\mathbf{r}}}{1-\rho}\right)$. So if r is such that $\frac{1}{\sqrt{n}} \left(\frac{1}{1+\rho} - \frac{\rho^{\mathbf{r}}}{1-\rho}\right) > \frac{\rho^{\mathbf{r}}}{(1-\rho)}$, that is $r > N = \left\lceil \log\left(\frac{(1-\rho)}{(1+\sqrt{n})(1+\rho)}\right) / \log\rho \right\rceil$ for the same i satisfying (9.7) we must have both x_i^{r+1} and x_i^* nonzero, and both have the same sign. Hence, after $N+1$ iterations of (9.4) we know at least one i for which x_i^* is nonzero, and its sign. If x_i^* is known to be negative, from (9.3), the variable w_i is positive in the solution of the LCP (q, M) (and consequently $z_i = 0$). On the other hand, if x_i^* is known to be positive, from (9.3), the variable z_i is positive and consequently $w_i = 0$ in the solution of the LCP (q, M). Using this information, the LCP (q, M) can be reduced to another LCP of order $(n-1)$ as discussed in Chapter 7. Since N defined above is finite and can be computed once the matrix B is known, after a finite number of steps of the iterative scheme (9.4), we can identify a basic variable in the complementary feasible basic vector for the LCP (q, M), and reduce the remaining problem into an LCP of order $(n-1)$, and repeat the method on it. The same thing is repeated until a complementary feasible basic vector for the LCP (q, M) is fully identified. In [9.22] W. M. G. Van Bokhoven has shown that the total number of steps that the iterative method has to be carried out before a basic variable in the complementary feasible basic vector for any of the principal subproblems in this process is identified, is at most N, where N is the number depending on the original matrix M, given above. So after at most nN steps of the iterative scheme (9.4) applied either on the original problem or one of its principal subproblems, a complementary feasible basic vector for the LCP (q, M) will be identified.

Exercise

9.1 Consider the LCP (q, M) where

$$M = \begin{pmatrix} 0 & A^T \\ -A & 0 \end{pmatrix}$$

which comes from transforming an LP into an LCP. Here M is neither PD nor even symmetric, but is PSD. Show that $(I + M)^{-1}$ exists in this case. Define, as before $b = -(I + M)^{-1}q$, $B = (I + M)^{-1}(I - M)$. Apply the transformation of variables as in (9.3) in this LCP, and show that it leads to the fixed point problem (9.2). Consider in this following iterative scheme for solving this fixed point problem in this case.

$$\begin{aligned} x^1 &= 0 \\ x^{r+1} &= \frac{b + x^r + B|x^r|}{2} \ . \end{aligned} \tag{9.8}$$

Show that if the LCP (q, M) has a solution, then the sequence $\{x^r\}$ generated by (9.8) converges to a solution of the fixed point problem and that the limit of this sequence leads to a solution of the LCP (q, M) in this case through the transformation (9.3). (W. M. G. Van Bokhoven [9.22]).

9.3 Iterative Methods for LCPs Associated with General Symmetric Matrices

In this section we consider the LCP (q, M), in which the only assumption made is that M is a symmetric matrix. The method and the results discussed here are due to O. L. Mangasarian [9.12], even through in some cases these turn out to be generalizations of the methods developed in references [10.33]. We begin with some basic definitions. We assume that $q \ngeqq 0$, as otherwise $(w = q, z = 0)$ is a solution of the LCP (q, M).

A square matrix $P = (p_{ij})$ is said to be **strictly lower triangular** if $p_{ij} = 0$ for $i \leqq j$. It is said to be **strictly upper triangular** if $p_{ij} = 0$ for all $i \geqq j$. Given a square matrix $M = (m_{ij})$ it can be written as the sum of three matrices $M = L+G+U$, where

$$L = \begin{pmatrix} 0 & 0 & \dots & 0 & 0 \\ m_{21} & 0 & \dots & 0 & 0 \\ m_{31} & m_{32} & \ddots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ m_{n1} & m_{n2} & \dots & m_{n,n-1} & 0 \end{pmatrix} , \quad G = \begin{pmatrix} m_{11} & 0 & \dots & 0 \\ 0 & m_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & m_{nn} \end{pmatrix} ,$$

$$U = \begin{pmatrix} 0 & m_{12} & m_{13} & \dots & m_{1,n-1} & m_{1,n} \\ 0 & 0 & m_{23} & \dots & m_{2,n-1} & m_{2,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & m_{n-1,n} \\ 0 & 0 & 0 & \dots & 0 & 0 \end{pmatrix}$$

The matrices L, G, U defined above, are respectively known as the **strictly lower triangular part**, the **diagonal part** and the **strict upper triangular part** of the given square matrix M. If M is symmetric we will have $L^T = U$.

Let $z = (x_j) \in \mathbf{R}^n$ be any column vector. We denote by $x^+ = (x_j^+)$ where $x_j^+ =$ Maximum $\{0, x_j\}$, for each $j = 1$ to n. The vector x^+ can be verified to be the nearest point in the nonnegative orthant to x.

The Iterative Method

Let $x^0 \geqq 0$ be an arbitrarily chosen initial point in the nonnegative orthant of $\mathbf{R}^n$. The iterative method is defined by the formula

$$z^{r+1} = \lambda\big(z^r - \omega E^r(Mz^r + q + K^r(z^{r+1} - z^r))\big)^+ + (1-\lambda)z^r \tag{9.9}$$

for $r = 0, 1, \ldots$, where λ, ω are parameters satisfying $0 < \lambda \leqq 1$, $\omega > 0$, whose values have to be chosen; for each r, K^r is a strictly lower triangular or strictly upper triangular matrix, and E^r is a positive diagonal matrix, which together satisfy

$$\begin{aligned} &E^r > \alpha I \\ &y^T\big((\lambda\omega E^r)^{-1} + K^r - \tfrac{M}{2}\big)y > \gamma\|y\|^2, \text{ for all } y \in \mathbf{R}^n \end{aligned} \tag{9.10}$$

for some positive numbers α, γ. Also $\{E^r : r = 0, 1, \ldots\}$, $\{K^r : r = 0, 1, \ldots\}$ are bounded sequences of matrices. When K^r is strictly lower triangular, (9.9) yields,

$$\begin{aligned} z_1^{r+1} &= \lambda\big(z_1^r - \omega E_{11}^r(M_1.z^r + q_1)\big)^+ + (1-\lambda)z_1^r, \text{ and} \\ z_j^{r+1} &= \lambda\big(z_j^r - \omega E_{jj}^r(M_j.z^r + q_j + \sum_{l=1}^{j-1} K_{jl}^r(z_l^{r+1} - z_l^r))\big)^+ + (1-\lambda)z_j^r, \text{ for } j = 2 \text{ to } n, \end{aligned}$$

where E_{jj}^r is the j^{th} diagonal entry in the diagonal matrix E^r and K_{jl}^r is the $(j,l)^{th}$ entry in K^r. So in this case z_j^{r+1} can be computed, very conveniently, in the specific order $j = 1, 2, \ldots, n$. If K^r is strictly upper triangular, (9.9) yields

$$\begin{aligned} z_n^{r+1} =& \lambda\big(z_n^r - \omega E_{nn}^r(M_n.z^r + q_n)\big)^+ + (1-\lambda)z_n^r, \text{ and} \\ z_j^{r+1} =& \lambda\big(z_j^r - \omega E_{jj}^r(M_j.z^r + q_j + \sum_{l=j+1}^{n} K_{jl}^r(z_l^{r+1} - z_l^r))\big)^+ + \\ & (1-\lambda)z_j^r, \text{ for } j = n-1 \text{ to } 1, \end{aligned}$$

and so in this case z_{j+1}^r can be computed very conveniently in the specific order $j = n, n-1, \ldots, 2, 1$.

How is the Iterative Method Obtained ?

The formula (9.9) for the iterative method is obtained by considering the quadratic programming problem

$$\begin{array}{ll} \text{Minimize} & f(z) = \frac{1}{2} z^T M z + q^T z \\ \text{Subject to} & z \geqq 0 \end{array} \tag{9.11}$$

In this section $f(z)$ denotes the function defined in (9.11). Remembering that M is a symmetric matrix, it can be verified that every KKT point for (9.11) leads to a solution of the LCP (q, M) and vice versa. The iteration (9.9) comes from an SOR (Successive Overrelaxation) type of gradient-projection algorithm for solving (9.11). We will discuss the choice for the parameters λ, ω and the matrices E^r, K^r in (9.9), later on. We will now characterize the convergence properties of the iterative method defined by (9.9).

Convergence Theorems

Theorem 9.7 *Let E be a diagonal matrix with positive diagonal entries. Then $(\overline{w} = M\bar{z} + q, \bar{z})$ is a solution of the LCP (q, M) iff $\bar{z}$ satisfies*

$$\left(z - \omega E(Mz + q)\right)^+ - z = 0, \text{ for some or all } \omega > 0 \ . \tag{9.12}$$

Proof. Suppose $(\overline{w} = M\bar{z} + q, \bar{z})$ is a solution of the LCP (q, M). Let $\omega > 0$ be arbitrary. If j is such that $\bar{z}_j = 0$, $M_j.\bar{z} + q_j \geqq 0$, we have $(\bar{z}_j - \omega E_{jj}(M_j.\bar{z} + q_j))^+ - \bar{z}_j = (-\omega E_{jj}(M_j.\bar{z} + q_j))^+ = 0$. If j is such that $M_j.\bar{z} + q_j = 0$ and $\bar{z}_j \geqq 0$, we have $(\bar{z}_j - \omega E_{jj}(M_j.\bar{z} + q_j))^+ - \bar{z}_j = \bar{z}_j - \bar{z}_j = 0$. So in this case $\bar{z}$ satisfies (9.12).

Conversely suppose $\bar{z} \in \mathbf{R}^n$ satisfies (9.12). Then $\bar{z} = (\bar{z} - \omega E(M\bar{z} + q))^+ \geqq 0$. Also, if for some j, we have $M_j.\bar{z} + q_j < 0$, then from (9.12), $0 = \left(\bar{z}_j - \omega E_{jj}(M_j.\bar{z} + q_j)\right)^+ - \bar{z}_j = -\omega E_{jj}(M_j.\bar{z} + q_j)$, a contradiction. So $M\bar{z} + q \geqq 0$ too. Now, for any j between 1 to n, if $\bar{z}_j - \omega E_{jj}(M_j.\bar{z} + q_j) \geqq 0$, we have $0 = (\bar{z}_j - \omega E_{jj}(M_j.\bar{z} + q_j))^+ - \bar{z}_j = -\omega E_{jj}(M_j.\bar{z} + q_j)$, and hence we must have $M_j.\bar{z} + q_j = 0$. On the other hand if $\bar{z}_j - \omega E_{jj}(M_j.\bar{z} + q_j) < 0$, we have $0 = (\bar{z}_j - \omega E_{jj}(M_j.\bar{z} + q_j))^+ - \bar{z}_j = -\bar{z}_j$, and hence we must have $\bar{z}_j = 0$. Thus depending on whether $\bar{z}_j - \omega E_{jj}(M_j.\bar{z} + q_j)$ in nonnegative or negative, we must have $M_j.\bar{z} + q_j$ or $\bar{z}_j$ equal to zero. So $\bar{z}^T(M\bar{z} + q) = 0$. Together with the nonnegativity proved above, we conclude that $(\overline{w} = M\bar{z} + q, \bar{z})$ is a solution of the LCP (q, M).

□

Theorem 9.8 *Let E be a diagonal matrix with positive diagonal entries and let $z \in \mathbf{R}^n$. Then $(z^+ - z)^T E^{-1}(z^+ - y) \leqq 0$ for all $y \geqq 0$.*

Proof. We have $(z^+ - z)^T E^{-1}(z^+ - y) = \sum_{j=1}^{n}((z_j^+ - z_j)(z_j^+ - y_j)/E_{jj})$. Here E_{jj} is the j^{th} diagonal entry of the matrix E. If j is such that $z_j \geqq 0$, then $z_j^+ - z_j = 0$. If j is such that $z_j < 0$, then $(z_j^+ - z_j)(z_j^+ - y_j)/E_{jj} = z_j y_j / E_{jj} \leqq 0$ since $y_j \geqq 0$. So $(z^+ - z)^T E^{-1}(z^+ - y)$ is the sum of non-postive quantities, and hence is non-positive.

□

Theorem 9.9 *Let $\{z^r : r = 1, 2, \ldots\}$ be the sequence of points obtained under the iterative scheme (9.9). If $\bar{z}$ is an accumulation point of this sequence, then $(\overline{w} = M\bar{z} + q, \bar{z})$ is a solution of the LCP (q, M).*

Proof. Since the initial point $z^0 \geqq 0$, and from (9.9) we conclude that $z^r \geqq 0$ for all $r = 1, 2, \ldots$. From strightforward manipulation it can be verified that

$$\begin{aligned}
&f(z^{r+1}) - f(z^r) = \\
&= \big(\omega E^r(Mz^r + q)\big)^T(\omega E^r)^{-1}(z^{r+1} - z^r) \\
&\quad + (z^{r+1} - z^r)^T M \tfrac{(z^{r+1} - z^r)}{2} \\
&= \Big(\tfrac{z^{r+1} - (1-\lambda)z^r}{\lambda} - z^r + \omega E^r(Mz^r + q \\
&\quad + K^r(z^{r+1} - z^r))\Big)^T(\omega E^r)^{-1}(z^{r+1} - z^r) + \\
&\quad + (z^{r+1} - z^r)\big(\tfrac{M}{2} - (\lambda\omega E^r)^{-1} - K^r\Big)(z^{r+1} - z^r) \\
&= \lambda\big(\tfrac{z^{r+1} - (1-\lambda)z^r}{\lambda} - (z^r - \omega E^r(Mz^r + q \\
&\quad + K^r(z^{r+1} - z^r)))\big)^T(\omega E^r)^{-1}\big(\tfrac{z^{r+1} - (1-\lambda)z^r}{\lambda} - z^r\big) + \\
&\quad + (z^{r+1} - z^r)^T\big(\tfrac{M}{2} - (\lambda\omega E^r)^{-1} - K^r\big)(z^{r+1} - z^r)
\end{aligned} \tag{9.13}$$

From (9.9) we know that $\frac{z^{r+1} - (1-\lambda)z^r}{\lambda} = \big(z^r - \omega E^r(Mz^r + q + K^r(z^{r+1} - z^r))\big)^+$. Also $\lambda > 0$. Using these, and Theorem 9.8, we conclude that the first term in the right hand side of (9.13) is $\leqq 0$. So $f(z^{r+1}) - f(z^r) \leqq (z^{r+1} - z^r)^T\big(\frac{M}{2} - (\lambda\omega E^r)^{-1} - K^r\big)(z^{r+1} - z^r)$. So,

$$\begin{aligned}
f(z^r) - f(z^{r+1}) &\geqq (z^{r+1} - z^r)^T\big((\lambda\omega E^r)^{-1} + K^r - \frac{M}{2}\big)(z^{r+1} - z^r) \\
&\geqq \gamma\|z^{r+1} - z^r\|^{\mathbf{2}}
\end{aligned} \tag{9.14}$$

The last inequality (9.14) follows from the conditions (9.10). Since $\gamma > 0$, (9.14) implies that $f(z^r) - f(z^{r+1}) \geqq 0$. Hence $\{f(z^r) : r = 1, 2, \ldots\}$ is a monotone non-increasing sequence of real numbers.

Let $\bar{z}$ be an accumulation point of $\{z^r : r = 0, 1, \ldots\}$. So there exists a sequence of positive integers such that the subsequence of z^r with r belonging to this sequence of integers converges to $\bar{z}$. Since $\{E^r : r = 0, 1, \ldots\}$, $\{K^r : r = 0, 1, \ldots\}$ are bounded sequences of matrices, we can again find a subsequence of the above sequence of positive integers satisfying the property that both the subsequences of E^r and K^r with r belonging to this subsequence converge to limits. Let $\{r_t : t = 1, 2, \ldots\}$ be this final subsequence of positive integers. So limit z^{r_t} as t tends to ∞ is $\bar{z}$. Also limits of E^{r_t}, K^{r_t} as t tends to ∞ exist, and denote these limits respectively by E and K. Since each E^r is a diagonal matrix satisfying $E^r \geqq \alpha I$, for some positive α for all r, we know that $E =$ limits E^{r_t} as t tends to ∞, is itself a diagonal matrix with positive diagonal entries. Since $f(z)$ is continuous, we have $f(\bar{z}) =$ limit $f(z^{r_t})$ as t tends to $+\infty$. Since $\{f(z^r) : r = 0, 1, \ldots\}$ is non-increasing sequence of real numbers, and

its subsequence $\{f(z^{r_t}) : t = 1, 2, \ldots\}$ converges to the limit $f(\bar{z})$, we conclude that $\{f(z^r) : r = 0, 1, \ldots\}$ is a non-increasing sequence of real numbers bounded below by $f(\bar{z})$. Hence the sequence $\{f(z^r) : r = 0, 1, \ldots\}$ itself converges. This and (9.14) together imply that $0 = \lim_{t \to +\infty} \left(f(z^{r_t}) - f(z^{1+r_t})\right) \geqq \lim_{t \to +\infty} \gamma \|z^{1+r_t} - z^{r_t}\|^{\mathbf{2}} \geqq 0$. From this and the fact that the sequence $\{z^{r_t} : t = 1, 2, \ldots\}$ converges to $\bar{z}$, we conclude that the sequence $\{z^{1+r_t} : t = 1, 2, \ldots\}$ also converges to $\bar{z}$. These facts imply that

$$\begin{aligned} 0 &= \lim_{t \to +\infty} \|z^{1+r_t} - z^{r_t}\| \\ &= \lambda \lim_{t \to +\infty} \|(z^{r_t} - \omega E^{r_t}(Mz^{r_t} + q + K^{r_t}(z^{1+r_t} - z^{r_t})))^+ - z^{r_t}\| \\ &= \lambda \|(\bar{z} - \omega E(M\bar{z} + q))^+ - \bar{z}\| \ . \end{aligned}$$

So we have $(\bar{z} - \omega E(M\bar{z} + q))^+ - \bar{z} = 0$. So by Theorem 9.7, $(\overline{w} = M\bar{z} + q, \bar{z})$ is a solution of the LCP (q, M).

□

Theorem 9.9 does not guarantee that the sequence $\{z^r : r = 0, 1, \ldots\}$ generated by the iterative method (9.9) has any limit points. When additional conditions are imposed, it is possible to guarantee that this sequence has some limit points.

Theorem 9.10 *Let M be a symmetric and copositive matrix of order n. Suppose $\{z^s : s = 1, 2, \ldots\}$ is an unbounded sequence (i. e., limit $\|z^s\|$ as s tends to ∞ is ∞) satisfying $z^s \geqq 0$ and $f(z^s) \leqq \alpha$ for all $s = 1, 2, \ldots$, where α is a constant. Then, there exists a subsequence $\{z^{s_t} : t = 1, 2, \ldots\}$ such that the sequence $\{y^{s_t} : y^{s_t} = \frac{z^{s_t}}{\|z^{s_t}\|}, t = 1, 2, \ldots\}$ converges to a point $\bar{y}$ satisfying $\bar{y} > 0$, $\bar{y}^T M \bar{y} = 0$, $q^T \bar{y} \geqq 0$. If, in addition, M is copositive plus, then $\bar{y}$ also satisfies $M\bar{y} = 0$, and in this case either (9.15) or (9.16) have no solution $z \in \mathbf{R}^n$.*

$$Mz + q > 0 \tag{9.15}$$

$$Mz \quad > 0 \tag{9.16}$$

Proof. Since $\|z^s\|$ diverges to $+\infty$, and $z^s \geqq 0$, we have $z^s \geq 0$ when s is sufficiently large. Eliminating some of the terms in the sequence $\{z^s : s = 1, 2, \ldots\}$ at the beginning of it, if necessary, we can therefore assume that $z^s \geq 0$ for all s in the sequence. So $\|z^s\| > 0$ and hence $y^s = \frac{z^s}{\|z^s\|}$ is defined for all s. The sequence $\{y^s : s = 1, 2, \ldots\}$ is an infinite sequence of points lying on the boundary of the unit sphere in $\mathbf{R}^n$ (i. e., satisfying $\|y^s\| = 1$ for all s), and hence if has a limit point $\bar{y}$, and there exists a subsequence $\{y^{s_t} : t = 1, 2, \ldots\}$ coverging to $\bar{y}$. Clearly $\|\bar{y}\| = 1$, and since $y^s \geq 0$ for all s, we have $\bar{y} \geq 0$. From the conditions satisfied by the sequence $\{z^s : s = 1, 2, \ldots\}$ we have

$$\frac{\alpha}{\|z^{s_t}\|^{\mathbf{2}}} \geqq \frac{f(z^{s_t})}{\|z^{s_t}\|^{\mathbf{2}}} = \frac{1}{2}(y^{s_t})^T M y^{s_t} + \frac{q^T y^{s_t}}{\|z^{s_t}\|} \ .$$

Taking the limit in this as t tends to $+\infty$, we have $0 \geqq (\frac{1}{2})\bar{y}^T M \bar{y}$, and since M is copostive and $\bar{y} \geq 0$, this implies that $\bar{y}^T M \bar{y} = 0$. Also, we have $\frac{\alpha}{\|z^{s_t}\|} \geqq \frac{f(z^{s_t})}{\|z^{s_t}\|} =$

$(\frac{1}{2})\|z^{s_t}\|(y^{s_t})^T My^{s_t} + q^T y^{s_t} \geqq q^T y^{s_t}$, since M is copositive and $y^{s_t} \geq 0$. Now taking the limit as t tends to $+\infty$, we get $0 \geqq q^T \bar{y}$.

If, in addition, M is copositive plus, and symmetric, $\bar{y}^T M\bar{y} = 0$, $\bar{y} \geq 0$ implies $M\bar{y} = 0$ by the definition of copositive plus. Also, in this case, if (9.15) has a solution z, multiplying both sides of (9.15) by $\bar{y}^T$ on the left yields (since $\bar{y} \geqq 0$) $0 < \bar{y}^T(Mz+q) = q^T\bar{y} + z^T(M\bar{y}) = q^T\bar{y} \leqq 0$, a contradiction. Similarly, if (9.16) has a solution z in this case, multiplying both sides of (9.16) on the left by $\bar{y} \geq 0$ yields $0 < \bar{y}^T Mz = z^T(M\bar{y}) = 0$, a contradiction.

Hence (9.15) has no solution z in this case. Also the system (9.16) has no solution z in this case.

□

Theorem 9.11 *Suppose either*

(a) M is a symmetric strictly copositive matrix, or

(b) M is a symmetric copositive plus matrix satisfying the condition that either (9.15) or (9.16) has a feasible solution z.

Then the sequence $\{z^r : r = 0, 1, \ldots\}$ generated by the iterative scheme (9.9) is bounded and has an accumulation point which leads to a solution of the LCP (q, M).

Proof. From Theorem 9.9 we know that $f(z^r) \leqq f(z^0)$ for all $r = 1, 2, \ldots$. If the sequence $\{z^r : r = 0, 1, \ldots\}$ is not bounded, it must have a subsequence which diverges, and using it together with the results in Theorem 9.10, we get a contradiction. Hence the sequence $\{z^r : r = 0, 1, \ldots\}$ must be bounded. So it must possess an accumulation point, and by Theorem 9.9, every accumulation point of this sequence leads to a solution of the LCP (q, M).

□

Corollary 9.1 *If M is symmetric, nonnegative and has positive diagonal elements, the sequence $\{z^r : r = 0, 1, \ldots\}$ obtained under (9.9) is bounded, and every accumulation point of it leads to a solution of the LCP (q, M).*

Proof. Follows from Theorem 9.11.

□

Corollary 9.2 *If M is symmetric, copositive plus, and either (9.15) or (9.16) has a feasible solution z, then the LCP (q, M) has a solution. In this case when the complementary pivot method is applied on the LCP (q, M), it cannot terminate in a ray, it terminates with a solution for the problem.*

Proof. Follows from Theorem 9.11 and Theorem 2.1.

□

Exercise

9.2 Suppose that M is symmetric and copositive plus. If $q < 0$ and there exists a z satisfying $Mz + q \geqq 0$, prove that the LCP (q, M) has a solution.

Now we state a theorem due to Ostrowski (Theorem 28.1 in reference [9.17], Theorem 6.3.1 in reference [9.12]) which we will use in proving Theorem 9.13 later on.

Theorem 9.12 *If the sequence $\{x^r : r = 0, 1, \ldots\}$ in $\mathbf{R}^n$ is bounded and limit $\|x^{r+1} - x^r\|$ as r tends to ∞ is zero, and if the set of accumulation points of $\{x^r : r = 0, 1, \ldots\}$ is not a continuum (i. e., a closed set which cannot be written as the union of two nonempty disjoint closed sets), then $\{x^r : r = 0, 1, \ldots\}$ converges to a limit.*

Proof. See references [9.17] mentioned above. □

Theorem 9.13 *Suppose M is symmetric, copositive plus and nondegenerate. Then the sequence $\{z^r : r = 0, 1, \ldots\}$ obtained under (9.9) converges to a solution of the LCP (q, M).*

Proof. In this case the determinant of M is nonzero, so M^{-1} exists. The vector $z = M^{-1}e$ can be verified to be a feasible solution for (9.16), so by Theorem 9.11, the sequence $\{z^r : r = 0, 1, \ldots\}$ of points obtained under the iterative scheme (9.9) for this case is bounded, and has at least one limit point. So the nonincreasing sequence of real numbers $\{f(z^r) : r = 0, 1, \ldots\}$ is also bounded and hence converges. From (9.14) we also conclude that limit $\|z^{r+1} - z^r\|$ as r tends to ∞ is zero. By Theorem 9.9 every accumulation point of $\{z^r : r = 0, 1, \ldots\}$ leads to a solution of the LCP (q, M). But the LCP (q, M) has only a finite number of solutions in this case, since M is nondegenerate (Theorem 3.2). So the sequence $\{z^r : r = 0, 1, \ldots\}$ has only a finite number of limit points in this case. This, together with the fact that limit $\|z^{r+1} - z^r\|$ as r tends to $+\infty$ is zero, implies by Theorem 9.12, that the sequence $\{z^r : r = 0, 1, \ldots\}$ converges to a limit, say $\bar{z}$. By Theorem 9.9, $\bar{z}$ leads to a solution of the LCP (q, M). □

Corollary 9.3 *If M is symmetric and* PD, *the sequence $\{z^r : r = 0, 1, \ldots\}$ produced by the iterative scheme (9.9) converges to a point $\bar{z}$ that leads to a solution of the LCP (q, M).*

Choice of Various Parameters in the Iterative Scheme (9.9)

By setting $K^r = 0$, $E^r = E$ for all r, where E is a diagonal matrix with positive diagonal elements, the iterative scheme (9.9) becomes the following scheme

$$\begin{aligned} z^0 &\geqq 0, \text{ an initial point} \\ z^{r+1} &= \lambda(z^r - \omega E(Mz^r + q))^+ + (1 - \lambda)z^r, \ r = 0, 1, \ldots \end{aligned} \tag{9.17}$$

where $0 < \lambda \leq 1$, $\omega > 0$ are chosen to satisfy th property that the matrix $2(\lambda\omega E)^{-1} - M$ is PD (to meet condition (9.10)). This special scheme is a projected Jacobi over-relaxation scheme (see reference [10.33]).

By setting $K^r = L$ or U, $E^r = E$ where E is a diagonal matrix with positive diagonal entries we obtain the following scheme which is a projected SOR (successive over relaxation) scheme.

$$\begin{aligned} z^0 &\geqq 0, \text{ an initial point} \\ z^{r+1} &= \lambda(z^r - \omega E(Mz^r + q + K^r(z^{r+1} - z^r)))^+ + (1-\lambda)z^r, \ r = 0, 1, \ldots \end{aligned} \tag{9.18}$$

where $0 < \lambda \leqq 1$, $\omega > 0$ satisfying the condition that

$$\lambda\omega < 2/\text{Maximum } \{G_{jj}E_{jj} : j \text{ such that } G_{jj} > 0\} \tag{9.19}$$

(where G is the diagonal part of M, and G_{jj} denotes the j^{th} diagonal element fo G if the set $\{j : j \text{ such that } G_{jj} > 0,\ j = 1 \text{ to } n\}$ is non-empty). This is to meet condition (9.10).

In (9.9), by setting $K^r = L$ and U alternately, we get the following projected symmetric SOR scheme.

$$\begin{aligned} z^0 &\geqq 0, \text{ an initial point.} \\ z^{r+1} &= \lambda(z^r - \omega E(Mz^r + q + L(z^{r+1} - z^r)))^+ + (1-\lambda)z^r, \ r = 0, 2, 4, \ldots \\ &= \lambda(z^r - \omega E(Mz^r + q + U(z^{r+1} - z^r)))^+ + (1-\lambda)z^r, \ r = 1, 3, 5, \ldots \end{aligned} \tag{9.20}$$

where $0 < \lambda \leqq 1$, $\omega > 0$ and E is a diagonal matrix with positive diagonal entries satisfying (9.19).

9.3.1 Application of These Methods to Solve Convex Quadratic Programs

The LCP (1.19) corresponding to the quadratic program (1.11) is associated with a matrix M which is not symmetric, and hence the iterative methods discussed in this section cannot be applied to solve it. Here we show that by treating the sign restrictions on the variables, also as contraints, and writing down the KKT optimality conditions for the resulting problem, we can derive an LCP associated with a symmetric matrix M corresponding to the problem, if the objective function is strictly convex (i. e., if D is PD). We consider the quadratic program (1.11), but include all the sign restrictions under the system of constraints. This leads to a problem in the following form :

$$\begin{aligned} &\text{Minimize} \quad Q(x) = cx + \tfrac{1}{2}x^T Dx \\ &\text{Subject to } Ax \geqq b \end{aligned} \tag{9.21}$$

where A is a given matrix of order $m \times n$; b, c are given vectors, and D is a given symmetric matrix of order n. *We assume that D is* PD. So (9.21) is a convex program with a strictly convex objective function. Associate the Lagrange multiplier u_i to the ith constraint in (9.21), $i = 1$ to m, and let $u = (u_1, \ldots, u_m)^T$. The Lagrangian for this problem is $L(x, u) = cx + \frac{1}{2}x^T Dx - u^T(Ax - b)$. The KKT necessary optimality conditions for this problem are (since D is symmetric)

$$\begin{aligned} \frac{\partial}{\partial x} L(x, u) = c^T + Dx - A^T u &= 0 \\ u &\geqq 0 \\ u^T(Ax - b) &= 0 \\ Ax - b &\geqq 0 \ . \end{aligned} \tag{9.22}$$

Since D is assumed to be PD here, D^{-1} exists. So from the first set of conditions in (9.22), we get $x = D^{-1}(A^T u - c^T)$. Using this we can eliminate x from (9.22). Denoting the slack variables $Ax - b$ by v, this leads to the LCP

$$\begin{aligned} v - (AD^{-1}A^T)u &= -(b + AD^{-1}c^T) \\ v \geqq 0, \quad u &\geqq 0 \\ v^T u &= 0 \ . \end{aligned} \tag{9.23}$$

So if $(\bar{u}, \bar{v})$ is a solution of the LCP (9.23), then $\bar{x} = D^{-1}(A^T\bar{u} - c^T)$ is a KKT point for the quadratic program (9.21). Applying Theorems 1.13, 1.14 to the convex quadratic program (9.21), we conclude that an optimum solution of (9,21) is a KKT point and vice versa. So solving (9.21) is equivalent to solving the LCP (9.23). Since the matrix $AD^{-1}A^T$ is symmetric this is an LCP associated with a symmetric matrix, and can be solved by the iterative methods discussed above. In particular, let L, G, U be respectively the strictly lower triangular part, the diagonal part, and the strictly upper triangular part of the matrix $AD^{-1}A^T$. Generate the sequence $\{u^r : r = 0, 1, \ldots\}$ in $\mathbf{R}^m$ by the following iterative scheme :

$$\begin{aligned} u^0 &\geqq 0 \text{ selected arbitrarily} \\ u^{r+1} &= (u^r - \omega E(AD^{-1}A^T u^r - b - AD^{-1}c^T + K^r(u^{r+1} - u^r)))^+ \end{aligned} \tag{9.24}$$

where E is a diagonal matrix with positive diagonal entries, K^r is either L or U and

$$0 < \omega < 2/ \text{ Maximum } \{G_{jj}E_{jj} : j \text{ such that } G_{jj} > 0\} \tag{9.25}$$

Note that (9.24) corresponds to setting $\lambda = 1$ in (9.9). Also (9.25) is the condtion (9.19) for this case. Also, using (9.24), u^{r+1} is computed from u^r in the specific order $j = 1, 2, \ldots, n$ if $K^r = L$, or in the specific order $j = n, n-1, \ldots, 1$ if $K^r = U$,as discussed earlier. We have the following theorems.

Theorem 9.14 *Each accumulation point $\bar{u}$ of the sequence $\{u^r : r = 0, 1, \ldots\}$ generated by (9.24) satisfies the property that $(\bar{v} = AD^{-1}A^T\bar{u} - (b + AD^{-1}c^T), \bar{u})$ is a solution of the LCP (9.23), and $\bar{x} = D^{-1}(A^T\bar{u} - c^T)$ is the optimum solution of the quadratic program (9.21).*

Proof. Follows by applying Theorem 9.9 to this case.

□

Theorem 9.14 does not, of course, guarantee that the sequence $\{u^r : r = 0, 1, \ldots\}$ generated by (9.24) has an accumulation point. This requires some more conditions on (9.21) as discussed below in Theorem 9.15.

Theorem 9.15 *If the set of feasible solutions of (9.21) has an interior point (i. e., there exists an x satisfying $Ax > b$) and D is symmetric* PD*, then the sequence $\{u^r : r = 0, 1, \ldots\}$ generated under (9.24) is bounded, and has at least one accumulation point. Each accumulation point $\bar{u}$ satisfies the statement in Theorem 9.14.*

Proof. Since $Ax > b$ is feasible, there exists a $\delta > 0$ such that the set of feasible solutions of

$$Ax \geqq b + \delta e \tag{9.26}$$

is nonempty. Fix δ at such a positive value. Since the set of feasible solutions of (9.26) is nonempty, and $Q(x)$ is strictly convex, the problem of minimizing $Q(x)$ subject to (9.26) has an optimum solution and it is unique. Suppose this optimum solution is $\tilde{x}$. The KKT necessary optimality conditions for this problem are

$$\begin{aligned} c^T + Dx - A^T u &= 0 \\ u &\geqq 0 \\ Ax &\geqq b + \delta e \\ u(Ax - b - \delta e) &= 0 \ . \end{aligned} \tag{9.27}$$

So there exists a $\tilde{u} \in \mathbf{R}^m$ such that $\tilde{x}$, $\tilde{u}$ together satisfy (9.27). Hence $(AD^{-1}A^T)\tilde{u} + (-b - AD^{-1}c^T) \geqq \delta e > 0$. This is like condition (9.15) for the LCP (9.23). Using this, this theorem follows from Theorem 9.11.

□

9.3.2 Application to Convex Quadratic Program Subject to General Constraints

The constraints in a quadratic program may be either linear inequalities or equations. Here we discuss how to apply the iterative scheme to solve the quadratic program directly without carrying out any transformations first to transform all the constraints into inqualities. We consider the quadratic program

$$\begin{aligned} &\text{Minimize} && Q(x) = cx + \tfrac{1}{2}x^T Dx \\ &\text{Subject to} && Ax \geqq b \\ & && Fx = d \end{aligned} \tag{9.28}$$

where A, F are given matrices of orders $m \times n$, $k \times n$ respectively; b, d, c are given vectors; and D is a given symmetric positive definite matrix of order n. Associate the Lagrange multiplier u_i, to the i^{th} inequality constraint in (9.20), $i = 1$ to m; and the Lagrange multiplier ξ_t to the t^{th} equality constraint in (9.28), $t = 1$ to k. Let $u = (u_i)$, $\xi = (\xi_t)$. The Lagrangian for this problems is $L(x, u, \xi) = cx + \frac{1}{2}x^T Dx - u^T(Ax - b) - \xi^T(Fx - d)$. Since D is symmetric, the KKT necessary optimality conditions for this problem are :

$$\begin{aligned} \frac{\partial}{\partial x} L(x, u, \xi) = c^T + Dx - A^T u - F^T \xi &= 0 \\ u &\geqq 0 \\ u^T(Ax - b) &= 0 \\ Ax - b &\geqq 0 \\ Fx - d &= 0 \ . \end{aligned} \tag{9.29}$$

From (9.29) we get $x = D^{-1}(A^T u - F^T \xi - c^T)$. Using this we can eliminate x from (9.29). When this is done, we are left with a quadratic program in terms of u and ξ associated with a symmetric matrix, in which the only constraints are $u \geqq 0$. The iterative scheme discussed above, specialized to solve this problem, becomes the following. Let L, G, U be respectively the strict lower triangular part, the diagonal part, and the strict upper triangular part of $\begin{pmatrix} A \\ F \end{pmatrix} D^{-1} \begin{pmatrix} A^T & F^T \end{pmatrix}$. Generate the sequence $\{(u^r, \xi^r) : r = 0, 1, \ldots\}$ by the following scheme

$$\begin{aligned} &(u^0, \xi^0) \text{ selected arbitrarily to satisfy } u^0 \geqq 0. \\ &\begin{pmatrix} u^{r+1} \\ \xi^{r+1} \end{pmatrix} = \begin{pmatrix} u^r \\ \xi^r \end{pmatrix} - \omega E \left[\begin{pmatrix} A \\ F \end{pmatrix} D^{-1} \begin{pmatrix} A^T & F^T \end{pmatrix} \begin{pmatrix} u^r \\ \xi^r \end{pmatrix} \right. \\ &\left. - \begin{pmatrix} A \\ F \end{pmatrix} D^{-1} c^T - \begin{pmatrix} b \\ d \end{pmatrix} + K^r \left(\begin{pmatrix} u^{r+1} \\ \xi^{r+1} \end{pmatrix} - \begin{pmatrix} u^r \\ \xi^r \end{pmatrix} \right) \right]^* \end{aligned} \tag{9.30}$$

where, as before, E is a diagonal matrix with positive diagonal entries, K^r is either L or U, ω is a positive number satisfying (9.25), and

$$\begin{pmatrix} u \\ \xi \end{pmatrix}^* = \begin{pmatrix} u^+ \\ \xi \end{pmatrix}$$

In (9.30), if $K^r = L$, u_j^{r+1} are computed in the order $1, 2, \ldots, m$ first and then ξ^{r+1} is computed. If $K^r = U$, ξ^{r+1} is first computed and then u_j^{r+1} are computed in the order $j = m, m-1, \ldots, 1$. We have the following theorems about this iterative scheme, corresponding to Theorems 9.14, 9.15 discussed earlier.

Theorem 9.16 *Each accumulation point $(\bar{u}, \bar{\xi})$ of $\{(u^r, \xi^r) : r = 0, 1, \ldots\}$ generated by (9.30) satisfies the property that $(\bar{u}, \bar{\xi}, \bar{x} = D^{-1}(A^T \bar{u} - F^T \bar{\xi} - c^T))$, satisfies (9.29) and $\bar{x}$ is the optimum solution of the quadratic program (9.28).*

Proof. Similar to Theorem 9.14.

□

Theorem 9.17 *If there exists an $\hat{x}$ satisfying $A\hat{x} > b$, $F\hat{x} = d$; and the set of rows of F is linearly independent, then the sequence $\{(u^r, \xi^r) : r = 0, 1, \ldots\}$ generated by (9.30) is bounded, and at last one accumulation point.*

Proof. Similar to Theorem 9.15.

□

9.3.3 How to Apply These Iterative Schemes in Practice

In practice we can only carry out the iterative scheme up to a finite number of steps, and obtain only a finite number of elements in the sequence. Usually the iterative scheme can be terminated whenever the current element in the sequence satisfies the constraints in the LCP to a reasonable degree of accuracy, or when the difference between successive elements in the sequence is small.

Exercise

9.3 Consider the LP

$$\begin{array}{ll} \text{Minimize} & \theta(x) = cx \\ \text{Subject to} & Ax \geqq b \end{array} \tag{9.31}$$

where A is a given matrix of order $m \times n$, and b, c are given vectors. Suppose this problem has an optimum solution, and let $\bar{\theta}$ denote the unknown optimum objective value in this problem. Now consider the following quadratic programming pertubation of this LP where ε is a small positive number

$$\begin{array}{ll} \text{Minimize} & \frac{\varepsilon}{2} x^T x + cx \\ \text{Subject to} & Ax \geqq b \end{array} \tag{9.32}$$

i) Prove that if (9.31) has an optimum solution, there exists a real positive number $\bar{\varepsilon}$ such that for each ε in the interval $0 < \varepsilon \leqq \bar{\varepsilon}$, (9.32) has an unique optimum solution $\bar{x}$ which is independent of ε, and which is also an optimum solution of the LP (9.31).

ii) If $\bar{\gamma}$ is the nonnegative optimal Lagrange multiplier associated with the last constraint in the following problem, where $\bar{\theta}$ is the optimum objective value in (9.31), prove that the $\bar{\varepsilon}$ in (i) can be selected to be any value satisfying $0 < \bar{\varepsilon} \leq \frac{1}{\bar{\gamma}}$. If $\bar{\gamma} = 0$, $\bar{\varepsilon}$ can be chosen to be any postive number.

$$\begin{array}{ll} \text{Minimize} & \frac{1}{2} x^T x \\ \text{Subject to} & Ax \geqq b \\ & -cx \geqq -\bar{\theta} \end{array}$$

(O. L. Mangasarian and R. R. Meyer [9.15])

9.4 Sparsity Preserving SOR Methods For Separable Quadratic Programming

The iterative SOR methods discussed in Section 9.3 for quadratic programming require the product of the constraint matrix by its transpose which can cause loss of both sparsity and accuracy. In this section we discuss special sparsity preserving versions of the general SOR algorithms presented in Section 9.3 for the LCP associated with a symmetric matrix, or equivalently for the quadratic program with nonnegativity constraints only; these versions are given in a simple explicit form in terms of the rows of the matrix M, and very large sparse problems can be tackled with them. Then we specialize these algorithms into SOR algorithm for solving separable quadratic programming problems that do not require multiplication of the constraint matrix by its transpose. The algorithms and the results discussed in this section are from O. L. Mangasarian [9.14].

We consider the LCP (9.1) in which $M = (m_{ij})$ is a symmetric matrix. As discussed in Section 9.3, solving (9.1) is equivalent to finding a KKT point for the quadratic programming problem (9.11). The SOR algorithm given here is a type of gradient projection algorithm for (9.11) with ω as the relaxation factor or step size that must satisfy $0 < \omega < 2$, and is based on those discussed in Section 9.3. The algorithm is the following. Choose $z^0 \geqq 0$ as the initial point. For $r = 0, 1, \ldots$ define for $j = 1$ to n.

$$z_j^{r+1} = (z_j^r - \omega\alpha_j(\gamma_j^{r+1} + \sum_{t=j}^{n} m_{jt}z_j^r + q_j))^+ \tag{9.33}$$

where $\alpha_j = \frac{1}{m_{jj}}$ if $m_{jj} > 0$, and $\alpha_j = 1$ if $m_{jj} \leqq 0$; $\gamma_1^{r+1} = 0$, $\gamma_j^{r+1} = \sum_{t=1}^{j-1} m_{jt}z_t^{r+1}$ for $j > 1$.

Convergence Theorems

Theorem 9.18 *Let M be a symmetric matrix. Then the following hold.*

(1) Each accumulation point of the sequence $\{z^r : r = 0, 1, \ldots\}$ generated by the iterative scheme (9.33) leads to a solution of the LCP (9.1).

(2) If M is symmetric and PSD *and the system: $Mz + q > 0$, has a solution z, the sequence $\{z^r : r = 0, 1, \ldots\}$ generated by (9.33) is bounded and has an accumulation point that leads to a solution of (9.1).*

(3) If M is symmetric and PD *the sequence $\{z^r : r = 0, 1 \ldots\}$ generated by (9.33) converges to a point $\bar{z}$ that leads to the unique solution of the LCP (9.1) (i. e., $(\overline{w} = M\bar{z} + q, \bar{z})$ is the solution of the LCP).*

(4) If M is symmetric and PSD *and (9.1) has a nonempty bounded solution set, the sequence $\{z^r : r = 0, 1, \ldots\}$ generated by (9.33) is bounded and has an accumulation point that leads to a solution of (9.1).*

Proof. Part (1) follows from Theorem 9.9. Part (2) follows from Theorem 9.11. Part (3) follows from Corollary 9.3. To prove part (4), notice that if the sequence $\{z^r : r = 0, 1, \ldots\}$ generated by (9.33) is unbounded, by Theorem 9.10, there exists a $\bar{y} \in \mathbf{R}^n$ satisfying: $\bar{y} \geq 0$, $M\bar{y} = 0$, $q^T\bar{y} \leqq 0$. So, if $(\overline{w} = M\bar{z} + q, \bar{z})$ is a solution of (9.1), then $(M(\bar{z} + \lambda\bar{y}) + q, \bar{z} + \lambda\bar{y})$ is also a solution of (9.1) for all $\lambda \geqq 0$ (since $\bar{z} + \lambda\bar{y} \geqq 0$, $M(\bar{z} + \lambda\bar{y}) + q \geqq 0$ and $0 \leqq (\bar{z} + \lambda\bar{y})^T(M(\bar{z} + \lambda\bar{y}) + q) = \lambda q^T\bar{y} \leqq 0$) contradicting the boundedness assuption of the solution set of (9.1).

□

9.4.1 Application to Separable Convex Quadratic Programming

Consider the quadratic program

$$\begin{array}{ll} \text{Minimize} & cx + \frac{1}{2}x^T Dx \\ \text{Subject to} & Ax \geqq b \\ & x \geqq 0 \end{array} \tag{9.34}$$

where A is a given matrix of order $m \times n$ and D is a positive diagonal matrix of order n. Let $u^T \in \mathbf{R}^m$, $v^T \in \mathbf{R}^n$ be the row vectors of Lagrange multipliers associated with the constraints and sign restrictions in (9.34). From the necessary optimality conditions for (9.34) it can be verified that an optimum solution for (9.34) is given by

$$x = D^{-1}(A^T u + v - c^T) \tag{9.35}$$

where (u, v) is an optimum solution of

$$\begin{array}{ll} \text{Minimize} & -b^T u + \frac{1}{2}(A^T u + v - c^T)^T D^{-1}(A^T u + v - c^T) \\ \text{Subject to} & (u, v) \geqq 0 \ . \end{array} \tag{9.36}$$

The problem (9.36) is in the same form as (9.11) and so the iterative algorithm (9.33) can be applied to solve it. It leads to the following iterative scheme. Choose $(u^0, v^0) \geqq 0$, $0 < \omega < 2$. Having (u^r, v^r) define for $i = 1$ to m.

$$\begin{aligned} u_i^{r+1} &= \left(u_i^r - \frac{\omega}{\|A_i.D^{-\frac{1}{2}}\|^2} \left(\left(A_i.D^{-1}(\gamma^{i,r+1} + \sum_{t=i}^{m} (A_t.)^T u_t^r + v^r - c^T) \right) - b_i \right) \right)^+ \\ v^{r+1} &= \left(v^r - \omega(A^T u^{r+1} + v^r - c^T) \right)^+ \end{aligned} \tag{9.37}$$

where $\gamma^{i,r+1} = 0$ for $i = 1$, or $= \sum_{t=1}^{i-1} (A_t.)^T u_i^{r+1}$ for $i > 1$. Notice that the sparsity or any structural properties that the constraint coefficient matrix A may have are taken advantage of in (9.37).

Theorem 9.19 *The following hold.*

(1) Each accumulation point $(\bar{u}, \bar{v})$ of the sequence $\{(u^r, v^r) : r = 0, 1, \ldots\}$ generated by (9.37) solves (9.36) and the corresponding $\bar{x}$ determined by (9.35) solves (9.34).

(2) If $\{x : Ax > b, x > 0\} \neq \emptyset$, the sequence $\{(u^r, v^r) : r = 0, 1, \ldots\}$ generated by (9.37) is bounded and has an accumulation point (u, v) and the corresponding x determined by (9.35) solves (9.34).

Proof. Part (1) follows from Theorem 9.18. To prove part (2), if $\{x : Ax > b, x > 0\} \neq \emptyset$, the perturbed positive definite quadratic program: minimize $cx + \frac{1}{2}x^T Dx$ subject to $Ax \geqq b + e_m\delta$, $x > e_n\delta$, where e_t is the column vector of all 1's in $\mathbf{R}^t$ for any t, has an optimum solution $\tilde{x}$. If $(\tilde{u}, \tilde{v})$ are the corresponding Lagrange multiplier vectors, from the KKT necessary optimality conditions we have

$$\tilde{x} = D^{-1}(A^T\tilde{u} + \tilde{v} - c^T) \geqq e_n\delta > 0$$
$$AD^{-1}(A^T\tilde{u} + \tilde{v} - c^T) - b \geqq e_m\delta > 0 \ .$$

These conditions are equivalent to the condition $Mz + q > 0$ in Theorem 9.18 for the LCP corresponding to problem (9.36). Hence, by Theorem 9.18, the sequence $\{(u^r, v^r) : r = 0, 1, \ldots\}$ generated by (9.37) is bounded, and hence has an accumulation point (u, v). The corresponding x determined from (9.35) solves (9.34) by the result in part (1).

□

In [9.14] O. L. Mangasarian used the iterative scheme (9.37) to develop a sparsity preserving SOR algorithm for solving linear programs. These schemes are also discussed in Section 16.4 [2.26].

Note 9.1 Suppose we have observations on the yield a_t at values of the temperature $t = 1, 2, \ldots, n$; and it is believed that this yield can be approximated very closely by a convex function of t. Let $x(t)$ be a convex function in t, and denote $x(t)$ by x_t for $t = 1, \ldots, n$. The problem of finding the best convex approximation to the yield, usng the least squares formulation, leads to the quadratic programming problem : find $x = (x_1, \ldots, x_n)^T$ to

$$\begin{array}{ll} \text{minimize} & \sum_{i=1}^{n}(x_i - a_i)^2 \\ \text{subject to} & x_{i+1} - x_i \geqq x_i - x_{i-1}, \quad i = 2, \ldots, n-1 \end{array}$$

This leads to the LCP (q, M), where

$$M = \begin{pmatrix} 6 & -4 & 1 & 0 & 0 & 0 & \ldots & 0 \\ -4 & 6 & -4 & 1 & 0 & 0 & \ldots & 0 \\ 1 & -4 & 6 & -4 & 1 & 0 & \ldots & 0 \\ 0 & 1 & -4 & 6 & -4 & 0 & \ldots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 1 & -4 & 6 \end{pmatrix}$$

and $q = (a_2 - a_1, a_3 - a_2, a_4 - a_3, \ldots)^T$.

J. S. Pang has tried to solve this class of LCPs for $n = 100$, using various iterative SOR methods discussed in this section and in Section 9.3 and found that convergence is not obtained even after several thousands of iterations. The matrix M given above is a very specially structured positive definite symmetric matrix, and the pivotal methods discussed in Chapters 2,4 perform very well in solving LCPs associated with this matrix M. An explanation for the poor performance (slow convergence) of SOR iterative methods on LCPs associated with M can be given in terms of the eigenvalues of M. At any rate, this example shows that iterative methods may not perform well on some classes of LCPs. These iterative methods are particularly useful for solving LCPs of very large orders or those which lack special structure, and thus are not easily handled by pivotal methods.

9.5 Iterative Methods for General LCPs

The results in Section 9.3 have been generalized by B. H. Ahn to the case of LCPs in which the matrix M may not be symmetric [9.3]. We discuss his results in this section. We want to solve the LCP (q, M) (9.1), where M is a given matrix of order n, not necessarily symmetric.

Given any matrix $A = (a_{ij})$ we will denote by $|A|$ the matrix $(|a_{ij}|)$. Also if A is a square matrix of order n, the matrix $C = (c_{ij})$ of order n where $c_{ii} = |a_{ii}|$ for $i = 1$ to n; and $c_{ij} = -|a_{ij}|$, $i, j = 1$ to n, $i \neq j$, is known as the **comparison matix** of A. We will now discuss some results on which the algorithm will be based.

Suppose we are given a square matrix A of order n which is not necessarily symmetric. So some of the eigenvalues of A may be complex. The **spectral radius** of A denoted by $\rho(A)$, is the maximum $\{|\lambda_1|, \ldots, |\lambda_n|\}$ where $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A. See Ortega and Rheinboldt [10.33] for results on the spectral radius of A.

Theorem 9.20 *Let $x, y \in \mathbf{R}^n$. Then $(x + y)^+ \leqq x^+ + y^+$, also $x \leqq y$ implies $x^+ \leqq y^+$. Also $(x - y)^+ \geqq x^+ - y^+$.*

Proof. Follows by direct verification. □

Theorem 9.21 *Let $g(z) = (z - \omega E(Mz + q))^+$, $\omega > 0$ and E is a diagonal matrix with positive diagonal entries. $(w = Mz + q, z)$ is a solution of the LCP (q, M) iff $g(z) = z$.*

Proof. Follows from Theorem 9.7 of Section 9.3. □

The Iterative Scheme

Choose $z^0 \geqq 0$ in $\mathbf{R}^n$ arbitrarily. Given z^r, determine z^{r+1} from

$$z^{r+1} = \left(z^r - \omega E\big(Mz^r + q + K(z^{r+1} - z^r)\big)\right)^+, \quad r = 0, 1, \ldots \tag{9.38}$$

where $\omega > 0$, E is a diagonal matrix with positive diagonal entries, and K is either a strictly upper triangular or a strictly lower triangular matrix. This scheme is a special case of (9.9) discussed earlier in Section 9.3. We will now study the convergence properties of the sequence $\{z^r : r = 0, 1, \ldots\}$ when M is not necessarily symmetric. Notice that the convergence properties of this sequence established in Section 9.3 using the descent function $\frac{1}{2}z^T Mz + q^T z$, need the symmetry of M, and hence may not hold when M is not symmetric.

Convergence Properties

Theorem 9.22 *The vectors in the sequence $\{z^r : r = 0, 1, \ldots\}$ obtained using (9.32) satisfy for each $r = 1, 2, \ldots$*

$$|z^{r+1} - z^r| \leqq (I - \omega E|K|)^{-1}|I - \omega E(M - K)|\cdot|z^r - z^{r-1}| \ . \tag{9.39}$$

Proof. From (9.38), we have $z^{r+1} - z^r = (z^r - \omega E(Mz^r + q + K(z^{r+1} - z^r)))^+ - (z^{r-1} - \omega E(Mz^{r-1} + q + K(z^r - z^{r-1})))^+ \leqq ((z^r - z^{r-1}) - \omega EM(z^r - z^{r-1}) - \omega EK(z^{r+1} - z^r) + \omega EK(z^r - z^{r-1}))^+$ from Theorem 9.20. So $(z^{r+1} - z^r)^+ \leqq ((I - \omega E(M - K))(z^r - z^{r-1}))^+ + (-\omega EK(z^{r+1} - z^r))^+$. We can obtain a similar result for $z^r - z^{r+1}$, that is $(z^r - z^{r+1})^+ \leqq ((I - \omega E(M - K))(z^{r-1} - z^r))^+ + (-\omega EK(z^r - z^{r+1}))^+$. Remembering that $|x| = x^+ + (-x)^+$ for any vector $x \in \mathbf{R}^n$, and adding the above two inequalities we get $|z^{r+1} - z^r| \leqq |I - \omega E(M - K)|\cdot|z^r - z^{r-1}| + \omega E|K|\cdot|z^{r+1} - z^r|$. Since K is strictly lower or upper triangular, the matrix $I - \omega E|K|$ is either a lower or upper triangular matrix, is invertible, and has a nonnegative inverse. Using this we get (9.39) from the last inequality.

□

Theorem 9.23 *Suppose the iteration parameters ω, E, K and the underlying matrix satisfy $\rho(Q) = \|Q\| < 1$, where $Q = (I - \omega E|K|)^{-1}(|I - \omega E(M - K)|)$. Then the sequence of points $\{z^r : r = 0, 1, \ldots\}$ generated by (9.38) converges to a point $\bar{z}$ where $(\overline{w} = M\bar{z} + q, \bar{z})$ is a solution of the LCP.*

Proof. Since $\rho(Q) < 1$, by the result in Theorem 9.22 we conclude that limit of $(z^{r+1} - z^r)$ as r tends to ∞, is zero. Also, clearly $Q \geqq 0$. Now $|z^r - z^0| \leqq |z^r - z^{r-1}| + \ldots + |z^1 - z^0| \leqq (Q^r + \ldots + I)|z^1 - z^0| \leqq (I - Q)^{-1}|z^1 - z^0|$, (since $\|Q\| < 1$) = a constant vector independent of r. So the sequence $\{z^r : r = 0, 1, \ldots\}$ is bounded. So it has a subsequence $\{z^{r_t} : t = 1, 2, \ldots\}$ which converges to a limit, $\bar{z}$, say. So $\lim_{t\to\infty} |z^{r_t+1} - \bar{z}| < \lim_{t\to\infty} |z^{r_t+1} - z^{r_t}| + \lim_{t\to\infty} |z^{r_t} - \bar{z}| = 0$, which shows that limit z^{r_t+1} as t tends to ∞ is $\bar{z}$ too. Now by the definition of z^{r_t+1} from equation (9.38), and taking the limit as t tend to $+\infty$, we conclude that $\bar{z} = (\bar{z} - \omega E(M\bar{z} + q))^+$. So by Theorem 9.21, $(\overline{w} = M\bar{z} + q, \bar{z})$ is a solution of the LCP. Also, as in the proof of Theorem 9.22, we can show that $|z^{r+1} - \bar{z}| \leqq Q|z^r - \bar{z}|$ holds for all r. Since $|\rho(Q)| < 1$; we conclude that limit $|z^r - \bar{z}|$ as r tends to $+\infty$ is zero. So the entire sequence $\{z^r : r = 0, 1, \ldots\}$ itself converges in the limit to $\bar{z}$.

□

Theorem 9.24 *Let L, D, U be respectively the strictly lower triangular, diagonal and strictly upper triangular parts respectively of M. Let K be L or U or 0. Let $B = I - \omega E|K|$, $C = |I - \omega E(M - K)|$, $A = B - C$. If A is a P-matrix, then the sequence $\{z^r : r = 0, 1, \ldots\}$ generated by (9.38) converges to a point $\bar{z}$ where $(\overline{w} = M\bar{z} + q, \bar{z})$ is a solution of the LCP (q, M).*

Proof. From the definition of B, we know that B is invertible and $B^{-1} \geqq 0$. Also $C \geqq 0$. So by 2.4.17 of Ortega and Rheinboldt's book [10.33], $\rho(B^{-1}C) < 1$ iff A^{-1} exists and is nonnegative. Since A is a Z-matrix, for it to have a nonnegative inverse, it sufficies if A is a P-matrix. The result follows from these and from Theorem 9.23. □

Theorem 9.25 *If $D - |L + U|$ is a P-matrix, then the sequence $\{z^r : r = 0, 1, \ldots\}$ generated by (9.38) with $K = L$ or U or 0 and $0 < \omega < 1/\max\{M_{jj}E_{jj} : j = 1 \text{ to } n\}$ where M_{jj}, E_{jj} are the j^{th} diagonal entries of the matrices M, E respectively, converges to a solution $\bar{z}$ where $(\overline{w} = M\bar{z} + q, \bar{z})$ is a solution of the LCP.*

Proof. Follows from Theorem 9.23. □

9.6 Iterative Methods for LCPs Based on Matrix Splittings

The iterative scheme and the results discussed in this section are due to J. S. Pang [9.18]. Consider the LCP (q, M), (9.1), of order n. If B, C are square matrices of order n satisfying

$$M = B + C \ , \tag{9.40}$$

(9.40) is said to be a splitting of the matrix M. Let E be a square nonnegative diagonal matrix of order n with diagonal entries $E_{ii} < 1$ for all i. This iterative scheme generates a sequence of points $\{z^r : r = 0, 1, \ldots\}$ by the following: Let B, C, be a splitting of M as in (9.40), $z^0 \in \mathbf{R}^n$ be an arbitrarily selected nonnegative vector. Given z^r, solve the LCP with data (q^r, B) where $q^r = q + (C + BE)z^r$, and let the solution of this LCP be $(u^{r+1} = Bz^{r+1} + q^r, z^{r+1})$. Then z^{r+1} is the next point in the sequence.

For this scheme to be practical, the matrix B should be such that the LCP (p, B) can be solved easily for any $p \in \mathbf{R}^n$. If B is a diagonal matrix with positive diagonal entries, or a triangular matrix with positive diagonal entries this will be the case. We assume that the splitting B, C of M is chosen so that the computation of the LCP (p, B) is easily carried out. Matrix splittings are used extensively in the study of iterative methods for solving systems of linear equations. The results in this section show that they are also useful for contructing iterative methods to solve LCPs. It can be verified that the iterative scheme discussed in Section 9.3 is a special case of the scheme discussed here, obtained by setting, $E = (1 - \lambda)I$ and the splitting B, C given

by $B = K + G/(\lambda\omega^*)$, and $C = (M - K) - G/(\lambda\omega^*)$ where $0 < \lambda < 1$, $\omega^* > 0$, and K is either a strictly lower triangular or a strictly upper triangular matrix and G is a diagonal matrix with positive diagonal entries.

Theorem 9.26 *Suppose the following conditions hold:*

(i) *B satisfies the property that the LCP (p, B) has a solution for all $p \in \mathbf{R}^n$;*

(ii) *$B = U + V + C^T$ with U, V being matrices satisfying conditions mentioned below;*

(III) *there exists a permutation matrix P such that the following matrices have the stated partitioned structure.*

$$P^T V P = \begin{pmatrix} V_{\Gamma\Gamma} & 0 \\ 0 & 0 \end{pmatrix}, \qquad P^T C P = \begin{pmatrix} C_{\Gamma\Gamma} & 0 \\ 0 & 0 \end{pmatrix},$$

$$P^T E P = \begin{pmatrix} E_{\Gamma\Gamma} & 0 \\ 0 & 0 \end{pmatrix}, \qquad P^T U P = \begin{pmatrix} 0 & U_{\Gamma\overline{\Gamma}} \\ -U^T_{\Gamma\overline{\Gamma}} & 0 \end{pmatrix},$$

with $V_{\Gamma\Gamma}$ being symmetric positive definite matrix, where $\mathbf{\Gamma} \subset \{1, \ldots, n\}$, $\overline{\mathbf{\Gamma}} = \{1, \ldots, n\} \setminus \mathbf{\Gamma}$, and $V_{\Gamma\Gamma}$ is the matrix of V_{ij} with $i \in \mathbf{\Gamma}$, $j \in \mathbf{\Gamma}$, etc.

(iv) *the initial vector $z^0 \geqq 0$ satisfies $q_{\overline{\Gamma}} - U^T_{\Gamma\overline{\Gamma}} z^0_\Gamma \geqq 0$.*
Then every accumulation point, $\bar{z}$ of the sequence $\{\bar{z} : r = 0, 1, \ldots\}$ generated by the scheme discussed above, satisfies the property that $(\overline{w} = M\bar{z} + q, \bar{z})$ is a solution of the LCP (q, M). Also if the following additional condition is satisfied:

(v) *the matrix $A_{\Gamma\Gamma} = (V + C + C^T)_{\Gamma\Gamma}$ is copositive plus and there exists vectors y^1_Γ, y^2_Γ such that*

$$q_\Gamma + A_{\Gamma\Gamma} y^1_\Gamma > 0. \tag{9.41}$$

$$y^2_\Gamma \geqq 0, \quad q_{\overline{\Gamma}} - U^T_{\Gamma\overline{\Gamma}} y^2_\Gamma > 0 \tag{9.42}$$

then the sequence $\{z^r : r = 0, 1, \ldots\}$ generated by the above scheme is bounded, and has an accumulation point.

Proof. Define $f(z) = q^T_\Gamma z_\Gamma + \frac{1}{2} z^T M z$. From the choice of z^0, and the iteration formula it is clear that $z^r \geqq 0$ for all r, and that $q_{\overline{\Gamma}} - U^T_{\Gamma\overline{\Gamma}} z^r \geqq 0$ for all $r \geqq 0$. In order to satisfy all these conditions, the matrix M need not be symmetric or PSD, but it must be copositive plus (for condition (iv)), and a principal rearrangement of M is given by

$$\begin{pmatrix} A_{\Gamma\Gamma} & U_{\Gamma\overline{\Gamma}} \\ -U^T_{\Gamma\overline{\Gamma}} & 0 \end{pmatrix} .$$

So $f(z) = q^T_\Gamma z_\Gamma + z^T_\Gamma A_{\Gamma\Gamma} z_\Gamma / 2$. Hence

$$\begin{aligned}
&f(z^{r+1}) - f(z^r) = \\
&= (q_\Gamma + A_{\Gamma\Gamma} z^r_\Gamma)^T (z^{r+1}_\Gamma - z^r_\Gamma) + (z^{r+1}_\Gamma - z^r_\Gamma)^T A_{\Gamma\Gamma} (z^{r+1}_\Gamma - z^r_\Gamma)/2 \\
&= (q_\Gamma + C_{\Gamma\Gamma} z^r_\Gamma + (V + C^T)_{\Gamma\Gamma} z^{r+1}_\Gamma)^T (z^{r+1}_\Gamma - z^r_\Gamma) \\
&\quad - (z^{r+1}_\Gamma - z^r_\Gamma)^T V_{\Gamma\Gamma} (z^{r+1}_\Gamma - z^r_\Gamma)/2 \\
&= (q_\Gamma + C_{\Gamma\Gamma} z^r_\Gamma + (V + C^T)_{\Gamma\Gamma} z^{r+1}_\Gamma + U_{\Gamma\overline{\Gamma}} z^{r+1}_{\overline{\Gamma}})^T (z^{r+1}_\Gamma - z^r_\Gamma) \\
&\quad - (z^{r+1}_{\overline{\Gamma}})^T (U^T_{\Gamma\overline{\Gamma}} z^{r+1}_\Gamma - U^T_{\Gamma\overline{\Gamma}} z^r_\Gamma) - (z^{r+1}_\Gamma - z^r_\Gamma)^T V_{\Gamma\Gamma} (z^{r+1}_\Gamma - z^r_\Gamma)/2 \\
&= (q_\Gamma + C_{\Gamma\Gamma} z^r_\Gamma + (V + C^T)_{\Gamma\Gamma} z^{r+1}_\Gamma + U_{\Gamma\overline{\Gamma}} z^{r+1}_{\overline{\Gamma}})^T ((z^{r+1} - Ez^r) + (E - I) z^r)_\Gamma \\
&\quad + (z^{r+1}_{\overline{\Gamma}})^T ((q_{\overline{\Gamma}} - U^T_{\Gamma\overline{\Gamma}} z^{r+1}_\Gamma) - (q_{\overline{\Gamma}} - U^T_{\Gamma\overline{\Gamma}} z^r_\Gamma)) - (z^{r+1}_\Gamma - z^r_\Gamma)^T V_{\Gamma\Gamma} (z^{r+1}_\Gamma - z^r_\Gamma)/2 \ ,
\end{aligned}$$

because $(u^{r+1} = Bz^{r+1} + q^r, z^{r+1})$ solves the LCP (q^r, B). From this we conclude that

$$f(z^{r+1}) - f(z^r) \leqq -\frac{1}{2}(z_\Gamma^{r+1} - z_\Gamma^r)^T V_{\Gamma\Gamma}(z_\Gamma^{r+1} - z_\Gamma^r) \leqq 0 \ . \tag{9.43}$$

Now let z^* be an accumulation point of the sequence $\{z^r : r = 0, 1, \ldots\}$ and let $\{z^{r_t} : t = 1, 2, \ldots\}$ be a subsequence coverging to z^*. This clearly implies by (9.43) that the sequence $\{f(z^r) : r = 0, 1, \ldots\}$ converges. As in the proof of Theorem 9.9, it can be shown that in this case,

$$\lim_{t\to\infty} z_\Gamma^{r_t - 1} = \lim_{t\to\infty} z_\Gamma^{r_t} = z_\Gamma^* \ . \tag{9.44}$$

Also, for each r_t we have

$$u_\Gamma^{r_t} = q_\Gamma + C_{\Gamma\Gamma} z_\Gamma^{r_t - 1} + B_{\Gamma\Gamma} z_\Gamma^{r_t} \geqq 0, \quad z^{r_t} \geqq E_{\Gamma\Gamma} z_\Gamma^{r_t - 1} \tag{9.45}$$

$$u_{\overline{\Gamma}}^{r_t} = q_{\overline{\Gamma}} + B_{\overline{\Gamma}\Gamma} z_\Gamma^{r_t} \geqq 0, \quad z_{\overline{\Gamma}}^{r_t} \geqq 0 \ . \tag{9.46}$$

$$(u_\Gamma^{r_t})^T (z_\Gamma^{r_t} - E_{\Gamma\Gamma} z_\Gamma^{r_t - 1}) = (u_{\overline{\Gamma}}^{r_t})^T z_{\overline{\Gamma}}^{r_t} = 0 \ . \tag{9.47}$$

Taking the limit as t tends to ∞ and usng (9.44), we conclude that $(w^* = Mz^* + q, z^*)$ solves the LCP (q, M).

Suppose now that condition (iv) holds. We will first show that the sequence $\{z_\Gamma^r : r = 0, 1, \ldots\}$ remains bounded. If not, by the results in Section 9.3, there must exist a $\bar{z}_\Gamma$ satisfying $\bar{z}_\Gamma \geq 0$, $q_\Gamma^T z_\Gamma \leqq 0$, $\bar{z}_\Gamma^T A_{\Gamma\Gamma} \bar{z}_\Gamma = 0$. Since $A_{\Gamma\Gamma}$ is copositive plus, this implies that $A_{\Gamma\Gamma}\bar{z}_\Gamma = 0$. These facts contradict the existence of a solution to the system (9.41). So $\{z_\Gamma^r : r = 0, 1, \ldots\}$ must be bounded.

Now we will prove that the sequence $\{z_{\overline{\Gamma}}^r : r = 0, 1, \ldots\}$ must be bounded too. Suppose not. Then there exists a subsequence $\{z_{\overline{\Gamma}}^{r_t} : t = 1, 2, \ldots\}$ such that $\|z_{\overline{\Gamma}}^{r_t}\|$ diverges to $+\infty$ as t tends to ∞. Let $y_{\overline{\Gamma}}^{r_t} = z_{\overline{\Gamma}}^{r_t} / \|z_{\overline{\Gamma}}^{r_t}\|$. This normalized sequence $\{y_{\overline{\Gamma}}^{r_t} : t = 1, 2, \ldots\}$ is bounded and hence has an accumulation point $y_{\overline{\Gamma}}^*$. Take a subsequence of $\{y_{\overline{\Gamma}}^{r_t} : t = 1, 2, \ldots\}$ which converges to $y_{\overline{\Gamma}}^*$. Denote this subsequence by $\{y_{\overline{\Gamma}}^{r_s} : s = 1, 2, \ldots\}$. Since the sequence $\{z_\Gamma^{r_s} : s = 1, 2, \ldots\}$ is bounded, it has a limit point. By considering a suitable subsequence again, if necessary, we can assume that we finally have a subsequence $\{z_\Gamma^{r_s} : s = 1, 2, \ldots\}$ which converges to z_Γ^*. Dividing (9.45) by $\|z_{\overline{\Gamma}}^{r_s}\|$ and taking the limit as s tends to ∞, we get $B_{\Gamma\overline{\Gamma}} y_{\overline{\Gamma}}^* \geqq 0$. From (9.47) we have $((I - E_{\Gamma\Gamma})z_\Gamma^*)^T B_{\Gamma\overline{\Gamma}} y_{\overline{\Gamma}}^* = 0$, and since $(I - E_{\Gamma\Gamma})$ is a positive diagonal matrix, this implies that $(z_\Gamma^*)^T B_{\Gamma\overline{\Gamma}} y_{\overline{\Gamma}}^* = 0$. Similarly, from (9.46), (9.47), we obtain that $(y_{\overline{\Gamma}}^*)^T (q_{\overline{\Gamma}} + B_{\overline{\Gamma}\Gamma} z_\Gamma^*) = 0$. Since $B_{\Gamma\overline{\Gamma}} = U_{\Gamma\overline{\Gamma}} = -B_{\overline{\Gamma}\Gamma}^T$, it follows that $(y_{\overline{\Gamma}}^*)^T q_{\overline{\Gamma}} = 0$. This together with $B_{\Gamma\overline{\Gamma}} y_{\overline{\Gamma}}^* \geqq 0$ and the fact that $y_{\overline{\Gamma}}^* \geq 0$ contradicts the existence of a solution to (9.42). So $\{z_{\overline{\Gamma}}^r : r = 0, 1, \ldots\}$ is also bounded. Hence the sequence $\{z^r : r = 0, 1, \ldots\}$ is bounded when the additional condition (iv) holds.

□

In [9.18] J. S. Pang, has established the convergence properties of the sequence $\{z^r : r = 0, 1, \ldots\}$ generated by the scheme discussed here, under various other sets of conditions on M, B, C, q.

9.7 Exercises

9.4 Consider the problem of finding $x, y \in \mathbf{R}^n$ satisfying

$$\begin{aligned} c^T + Dx + y &\geqq 0, \quad x \geqq 0, \quad y \geqq 0 \\ b - x &\geqq 0 \\ x^T(c^T + Dx + y) = y^T(b - x) &= 0 \end{aligned} \tag{9.48}$$

where $b > 0$, c, D are given matrices of order $n \times 1$ and $n \times n$ respectively. When D is symmetric, these are the necessary optimality conditions for the quadratic program: minimize $cx + \frac{1}{2}x^T Dx$, subject to $0 \leqq x \leqq b$. A model of type (9.48) arises in the study of multicommodity market equilibrium problems with institutional price controls (here D is not necessarily symmetric).

1) Show that (9.48) is equivalent to the LCP (q, M) where

$$q = \begin{pmatrix} c^T \\ b \end{pmatrix}, \qquad M = \begin{pmatrix} D & I \\ -I & 0 \end{pmatrix} .$$

2) Let $\mathbf{\Delta} = \{x : 0 \leqq x \leqq b\}$ and let $P_\Delta(y)$ denote the nearest point in $\mathbf{\Delta}$ (in terms of the usual Euclidean distance) to y. Give $\bar{x} \in \mathbf{\Delta}$, define the corresponding $\bar{y} = (\bar{y}_i) \in \mathbf{R}^n$ by $\bar{y}_i = 0$ if $\bar{x}_i < b_i$, or $= -D_i.\bar{x} + c_i$ if $\bar{x}_i = b_i$. We say that $\bar{x}$ leads to a solution of (9.48) if $(\bar{x}, \bar{y})$ solves (9.48). Consider the following iterative scheme. Choose $x^0 \in \mathbf{\Delta}$. For $r = 0, 1, \ldots$, given x^r, define

$$x^{r+1} = \lambda P_\Delta(x^r - \omega E(Dx^r + c^T + K(x^{r+1} - x^r))) + (1 - \lambda)x^r \tag{9.49}$$

where $0 < \lambda \leqq 1$, $\omega > 0$, E is a positive diagonal matrix of order n, and K is either the strictly lower or the strictly upper triangular part of D. Using the result in Exercise 7.7, x_j^{r+1} in (9.49) can be determined in the order $j = 1$ to n if K is the strictly lower triangular part of D, or in the order $j = 1$ to n if $\mathbf{K}$ is the strictly upper triangular part of D. In the sequence $\{x^r : r = 0, 1, \ldots\}$ generated by (9.49), $x^r \in \mathbf{\Delta}$ for all r, so, it has at least one accumulation point. If D is symmetric and $\lambda\omega < 2/$ (maximum$\{D_{jj}E_{jj} : j$ such that $D_{jj} > 0\}$), (here D_{jj}, E_{jj} are the j^{th} diagonal entries in the matrices D, E respectively), prove that every accumulation point of the sequence generated by (9.49) leads to a solution of (9.48). In addition, if D is also nondegenerate, prove that the sequence $\{x^r : r = 0, 1, \ldots\}$ generated by (9.49) in fact converges to a point $\bar{x}$ that leads to a solution of (9.48).

3) If D is a Z-matrix, not necessarily symmetric, and $x^0 \in \mathbf{T} = \{x : x \in \mathbf{\Delta}$ and for each i either $x_i = b_i$ or $c_i + D_i.x \geqq 0\}$, (for example, $x^0 = b$ will do) and $\lambda\omega \leqq 1/$ (maximum$\{D_{jj}E_{jj} : j$ such that $D_{jj} > 0\}$), prove that the sequence $\{x^r : r = 0, 1, \ldots\}$ generated by (9.49) is a monotonic sequence that converges to a point $\bar{x}$ leading to a solution of (9.48).

4) A square matrix is said to be a H-matrix if its comparison matrix (which is a Z-matrix by definition) is a P-matrix. If D is a H-matrix, not necessarily symmetric, with positive diagonal elements, prove that the sequence $\{x^r : r = 0, 1, \ldots\}$ generated by (9.49), with $\omega \leqq 1/$ (maximum$\{D_{jj}E_{jj} : j = 1$ to $n\}$) converges to the point $\bar{x}$ that leads to the unique solution of (9.48).

(B. H. Ahn [9.4])

9.5 For each $i = 1$ to m, let $f_i(x)$ be a real valued convex function defined on $\mathbf{R}^n$. Let $\mathbf{K} = \{x : f_i(x) \leqq 0, i = 1 \text{ to } m\}$. Assume that $\mathbf{K} \neq \emptyset$. Let $x^0 \in \mathbf{R}^n$ be an arbitrary initial point. The following iterative method known as the method of successive projection is suggested as a method for finding a point in $\mathbf{K}$. Given x^r, let x^{r+1} be the nearest point in the set $\{x : f_{i_r}(x) \leqq 0\}$ to x^r. The index i_r is choosen by one of the following

Cyclic Order : Here the indices $\{i_r : r = 0, 1, \ldots\}$ are choosen in cyclical order from $\{1, 2, \ldots, m\}$. So $i_0 = 1$, $i_1 = 2$, ..., $i_m = 1$, $i_{m+1} = 2$, and so on.

Most Violated Criterion : Here i_r is the i for which the distance between x^r and the nearest point to x^r in the set $\{x : f_i(x) \leqq 0\}$ is maximum (ties for this maximum are broken arbitrarily).

Prove that the sequence $\{x^r : r = 0, 1, \ldots\}$ converges to a point in $\mathbf{K}$.

(L. M. Bregman [9.5])

9.8 References

9.1 M. Aganagic, "Iterative methods for linear complementarity problems", Tech. Report SOL78-10, Systems Optimization Laboratory, Dept. of Operations Research, Stanford Universty (Sept. 1978).

9.2 M. Aganagic, "Newton's Method for Linear Complementarity Problems", *Mathematical Programming 28*, (1984) 349-362.

9.3 B. H. Ahn, "Computation of Asymmetric Linear Complementarity Problems by Iterative Methods", *Journal of Optimization Theory and Applications*, 33 (1981) 175-185.

9.4 B. H. Ahn, "Iterative Methods for Linear Complementarity Problems With Upper Bounds on Primary Variables", *Mathematical Programming*, 26 (No. 3) (1983) 295-315.

9.5 L. M. Bregman, "The method of successive projection for finding a common point of convex sets", *Soviet Mathematics Doklady*, 162, 3 (1965) 688-692.

9.6 C. W. Cryer, "The Solution of a Quadratic Programming Problem Using Systematic Overrelaxation", *SIAM Journal on Control 9*, (1971) 1421-1423.

9.7 Y. C. Cheng, "Iterative Methods for Solving Linear Complementarity and Linear Programming Problems", Ph.D. dissertation, Department of Computer Science, University of Wisconsin-Madison, Wisconsin (1981).

9.8 I. M. Gelfand, "*Lectures on Linear Algebra*", Kreiger (1968).

9.9 K. M. Hoffman and R. A. Kunze, "*Linar Algebra*", Prentice Hall (1971).

9.10 A. S. Householder, "*Theory of Matrices and Numerical Analysis*", Dover (1975).

9.11 N. W. Kappel and L. T. Watson, "Iterative Algorithms for the Linear Complementarity Problem", *International J. Computer Math.*, 19 (1986) 273–297.

9.12 O. L. Mangasarian, "Solution of Symmetric Linear Complementarity Problems by Iterative Methods", *Journal of Optimization Theory and Applications*, 22 (No. 4) (1977) 465-485.

9.13 O. L. Mangasarian, "Iterative Solutions of Linear Programs", *SIAM Journal on Numerical Analysis*, 18, 4 (August 1981), 606-614.

9.14 O. L. Mangasarian, "Sparsity Preserving SOR Algorithms for Separable Quadratic and Linear Programming", *Computers and Operational Research*, 11, 2 (1984) 105–112.

9.15 O. L. Mangasarian and R. R. Meyer, "Nonlinear Perturbation of Linear Programs", *SIAM Journal on Control and Optimization*, 17 (1979) 743-752.

9.16 J. M. Ortega, "*Numerical Analysis, A Second Course*", Academic Press, New York (1972).

9.17 A. M. Ostrowski, "Solution of Equations and Systems of Equations", Academic Press, Second Edition, New York (1966).

9.18 J. S. Pang, "On the Convergence of a Basic Iterative Method for the Implicit Complementarity Problem", *Journal of Optimization Theory and Applications*, 37 (1982) 149-162.

9.19 J. S. Pang and D. Chan, "Iterative Methods for Variational and Complementarity Problems", *Mathematical Programming*, 24 (1984) 284-313.

9.20 L. B. Rall, "Computational Solution of Nonlinear Operator Equations", John Wiley & Sons, Inc., 1969.

9.21 W. Rudin, "*Principles of Mathematical Analysis*", McGraw-Hill, New York (1976).

9.22 W. M. G. Van Bokhoven, "A Class of Linear Complementarity Problems is Solvable in Polynomial Time", Department of Electrical Engineering, University of Technology, P.O. Box 513, 5600 MB Eindhoven, Netherlands (1980).

9.23 R. S. Varga, "*Matrix iterative analysis*", Prentice-Hall Inc., New Jersey (1962).

9.24 D. M. Young, "*Iterative Solution of Large Linear Systems*", Academic Press, New York (1971).

Chapter 10

SURVEY OF DESCENT BASED METHODS FOR UNCONSTRAINED AND LINEARLY CONSTRAINED MINIMIZATION

Nonlinear Programming Problems

Eventhough the title "Nonlinear Programming" may convey the impression that the subject includes all optimization problems other than linear programming problems, it is not usually the case. Optimization problems involving discrete valued variables (i. e., those which are restricted to assume values from specified discrete sets, such as 0-1 variables) are not usually considered under nonlinear programming, they are called **discrete**, or **mixed-discrete optimization problems** and studied separately. There are good reasons for this. To solve discrete optimization problems we normally need very special techniques (typically of some enumerative type) different from those needed to tackle continous variable optimization problems. So, the term nonlinear program usually refers to an optimization problem in which the variables are continuous variables, and the problem is of the following general form:

$$\begin{array}{lll} \text{minimize} & \theta(x) & \\ \text{subject to} & h_i(x) = 0, & i = 1 \text{ to } m \\ & g_p(x) \geqq 0, & p = 1 \text{ to } t \end{array} \qquad \text{(P)}$$

where $\theta(x)$, $h_i(x)$, $g_p(x)$ are all real valued continuous functions of $x = (x_1, \ldots, x_n) \in \mathbf{R}^n$.

Suppose some of these functions are not differentiable at some points x. Assume that gradients exist for each function almost everywhere, but are not continuous. Then problem (P) is known as a **non-smooth** or **non-differentiable optimization problem**. On such a problem, the usual gradient-based methods and results may fail, and

special attention must be given to the surfaces of non-differentiability, it becomes very important to consider generalized gradients to handle such problems.

If all the functions $\theta(x)$, $h_i(x)$, $g_p(x)$ are continuously differentiable, problem (P) is known as a **smooth nonlinear program**. In this book we only study smooth nonlinear programs. However, some of the techniques that we discuss may convert a smooth NLP into a special type of nonsmooth NLP, and then solve it. As an example, the simplicial method discussed in Section 2.7.6 to solve the smooth NLP: minimize $\theta(x)$, subject to $g_i(x) \leqq 0$, converts it into the NLP: minimize $\theta(x)$, subject to $s(x) \leqq 0$, where $s(x) = \max \cdot \{g_1(x), g_2(x), \ldots, g_m(x)\}$. This modified problem is a nonsmooth optimization problem, since $s(x)$ may not be differentiable at some points x. However, because of the special nature of $s(x)$, we know that $\partial s(x) =$ convex hull of $\{\nabla g_i(x) : i$ such that $g_i(x) = s(x)\}$, and hence for any given x, it is easy to find at least one point in $\partial s(x)$, and the special simplicial algorithms discussed in Section 2.7, are able to solve this modified problem using only this information.

Consider the NLP (P) and assume that all the functions are continuously differentiable. The constraints in (P) are either equality constraints, or inequality constraints. (P) is the general form of the problem, and in a particular instance of (P), there may or may not be such constraints. This problem is said to be:

an **unconstrained minimization problem**, if there are no constraints on the variables, in the statement of the problem,

a **linear programming problem**, if all the functions $\theta(x)$, $h_i(x)$, $g_p(x)$ are affine functions,

a **quadratic programming problem**, if $\theta(x)$ is a quadratic function, and all $h_i(x)$ and $g_p(x)$ are affine functions,

an **equality constrained problem**, if there are no inequality constraints on the variables,

a **linearly constrained** NLP, if all the constraint functions $h_i(x)$, $g_p(x)$ are affine functions,

a **convex programming problem** if $\theta(x)$ is a convex function, all $h_i(x)$ are affine functions, and all $g_p(x)$ are concave functions,

a **nonconvex programming problem**, if it is not a convex programming problem as defined above.

In this chapter, we provide a brief survey of some commonly used algorithms for smooth NLPs, those in the areas of unconstrained and linearly constrained NLPs, which constitute alternate methods to those discussed so far for solving quadratic programs.

10.1 A FORMULATION EXAMPLE FOR A LINEARLY CONSTRAINED NONLINEAR PROGRAM

We begin this chapter with a practical example due to C. H. White, of a nonlinear model in which the constraints are linear. It arose in the boiler shop of a company which has five (5) boilers operating in parallel for generating steam. Data on the boilers is given below.

Tableau 10.1

Boiler i	Boiler load range limits					
	lower l_i	upper k_i	a_{0i}	a_{1i}	a_{2i}	a_{3i}
1	10 units	60	56.49	1.67	−.041	.00030
2	10	60	71.37	0.61	−.016	.00011
3	15	120	23.88	2.05	−.024	.00009
4	12.5	112.5	17.14	2.73	−.035	.00014
5	15	135	72.38	0.34	−.003	.00001

The unit measures the rate at which steam is produced per unit time. If the ith boiler is kept on, it must be operated within its load range limits l_i, k_i. The boiler's energy efficiency defined as a percentage is 100 × (energy content of output steam)/(energy content in the input fuel). It tends to increase as the load moves up from the minimum allowable operating load, and then peaks and drops as the load approaches the upper limit. Data was collected on the boiler efficiencies at different operating load levels, and the plots indicated that boiler efficiency can be approximated very well by a cubic polynomial of the operating load. Let $y(\xi)$ = efficiency of a boiler when it is operating at load ξ units. We approximate $y(\xi)$ by $f(\xi) = a_0 + a_1\xi + a_2\xi^{\mathbf{2}} + a_3\xi^{\mathbf{3}}$, where a_0, a_1, a_2, a_3 are parameters to be estimated from data. The problem of determining the best values of the parameters that give the closest fit between observed efficiency and the cubic polynomial, is known as the **parameter estimation problem** or the **curve fitting problem**. Suppose we have r observations on a boiler, at load levels ξ_t, $t = 1$ to r yielding observed efficiencies of y_t, $t = 1$ to r respectively. To derive the closest fit we need to construct a measure of deviation of the functional value $f(\xi)$ from the observed $y(\xi)$ over the range of values of ξ used in the experiment, depending on the parameter vector $a = (a_0, a_1, a_2, a_3)$. Three different measures are in common use.

They are

$$L_2(a) = \sum_{t=1}^{r} (y_t - a_0 - \sum_{s=1}^{3} a_s \xi_t^{\mathbf{s}})^{\mathbf{2}}$$

$$L_1(a) = \sum_{t=1}^{r} |y_t - a_0 - \sum_{s=1}^{3} a_s \xi_t^{\mathbf{s}}|$$

$$L_\infty(a) = \text{Maximum } \{|y_t - a_0 - \sum_{s=1}^{3} a_s \xi_t^{\mathbf{s}}| : t = 1 \text{ to } r\}.$$

Since the $L_2(a)$ measure is a sum of squares, the technique which chooses the parameter vector a to minimize $L_2(a)$ is called the **least squares approach** or the **method of least squares**. If $\hat{a} = (\hat{a}_0, \hat{a}_1, \hat{a}_2, \hat{a}_3)$ is the best vector of parameter values obtained under this method, the function $\hat{a}_0 + \hat{a}_1\xi + \hat{a}_2\xi^{\mathbf{2}} + \hat{a}_3\xi^{\mathbf{3}}$ is called the **least squares approximation** for $y(\xi)$.

If the parameter vector a is determined so as to minimize the measure $L_\infty(a)$, the resulting function $f(\xi)$ is known as the **Tschebycheff approximation** for $y(\xi)$.

If all the parameters appear linearly in the functional form $f(\xi)$ (as in this boiler efficiency example) the problem of minimizing either the L_1- or L_∞-measures can both be posed as linear programs and solved by the efficient simplex method. However, if the parameters appear nonlinearly in the functional form, the least squares method is preferred for parameter estimation.

If the measure of deviation is too large even at the best parameter values, it is necessary to review the choice of the functional form and modify it. Besides, it is possible that no simple function provides a good approximation for all possible values of load. It is only necessary to find a good functional representation of the efficiency in the neighborhood of the optimum load values, if some reliable practical knowledge is available on the likely location of this optimum.

Thus, even the process of constructing a mathematical model for the problem might itself need the application of optimization algorithms for parameter estimation.

The Basic Difference Between Linear and Nonlinear Models

To construct a linear programming model involving n nonnegative variables subject to m constraints, we need to estimate the $(m+1)(n+1)-1$ coefficients of the variables in the constraints and the objective function, these are the data elements in the model. Real life LP applications routinely involve models with $n = 100,000$ or more, and m as large as 6000. A large scale LP model is usually of this size.

To construct a nonlinear model, we have to determine the functional form of the objective and each constraint function, and obtain the best values for any parameters in each. For this reason, practical nonlinear models tend to have fewer variables than linear models. Depending on how complicated the functions involved are, a nonlinear model with about 200 variables could usually be considered as a large scale model.

Boiler Example, Continued

For the boiler problem, estimates of the best parameter values in the functional form for the efficiency of each boiler are given in Tableau 10.1.

At a point of time, the Company's steam requirements are 350 units per unit time. The problem is to determine how this total load of 350 units should be shared across the five (5) parallel boilers so as to minimize the total fuel cost. It may be possible to get a lower overall cost by shutting down one or more of the boilers and meeting the demand using only the remaining boilers. For example, here it can be verified that the total load of 350 units can be met using boilers 3, 4, and 5 only. Thus the problem of determining the most efficient plan to meet a load of exactly 350 units, leads to a mixed integer nonlinear programming problem in which there are five zero-one variables to determine which of the five boilers are shut down and which are kept operating, and the operating load level for the boilers that are kept operating. In this plant however, it is known that the Company's steam requirements vary with time. When the demand for steam goes up, if a boiler is kept operating, it is a relatively easy matter to increase the boiler's steam output by turning a few valves. On the other hand turning on a shut down boiler is an expensive operation. In order to be able to meet the varying steam requirements over time, it was determined that all the five boilers should be kept operating. Under this condition, since $x_i/f_i(x_i)$ is a measure of the energy cost of obtaining a load of x_i units from boiler i, we are lead to the following nonlinear model:

$$\begin{aligned} &\text{minimize} \quad \sum_{i=1}^{5} x_i/f_i(x_i) \\ &\text{subject to} \quad \sum_{i=1}^{5} x_i = 350 \\ &l_i \leqq x_i \leqq k_i, \ i = 1 \text{ to } 5 \end{aligned}$$

which is a linearly constrained nonlinear program.

Exercise

10.1 Using the $0-1$ variables y_i defined by

$$\begin{aligned} y_i &= 1 \quad \text{if the } i\text{th boiler is kept operating} \\ &= 0 \quad \text{otherwise} \end{aligned}$$

formulate the problem of determining the most efficient plan for producing exactly 350 units of steam per unit time as a mixed integer NLP.

10.2 TYPES OF SOLUTIONS FOR A NONLINEAR PROGRAM

Consider a NLP in which a function $\theta(x)$ is required to be optimized subject to some constraints on the variables $x = (x_1, \ldots, x_n)^T$. Let $\mathbf{K}$ denote the set of feasible solutions for this problem. For this problem a feasible solution $\overline{x} \in \mathbf{K}$ is said to be a

local minimum, if there exists an $\varepsilon > 0$ such that $\theta(x) \geqq \theta(\overline{x})$ for all $x \in \mathbf{K} \cap \{x : \|x - \overline{x}\| < \varepsilon\}$,

strong local minimum, if there exists an $\varepsilon > 0$ such that $\theta(x) > \theta(\overline{x})$ for all $x \in \mathbf{K} \cap \{x : \|x - \overline{x}\| < \varepsilon\}$, $x \neq \overline{x}$,

weak local minimum, if it is a local minimum, but not a strong one,

global minimum, if $\theta(x) \geqq \theta(\overline{x})$ for all $x \in \mathbf{K}$,

local maximum, if there exists an $\varepsilon > 0$ such that $\theta(x) \leqq \theta(\overline{x})$ for all $x \in \mathbf{K} \cap \{x : \|x - \overline{x}\| < \varepsilon\}$,

strong local maximum, if there exists an $\varepsilon > 0$ such that $\theta(x) < \theta(\overline{x})$ for all $x \in \mathbf{K} \cap \{x : \|x - \overline{x}\| < \varepsilon\}$, $x \neq \overline{x}$,

weak local maximum, if it is a local maximum, but not a strong one,

global maximum, if $\theta(x) \leqq \theta(\overline{x})$ for all $x \in \mathbf{K}$,

stationary point, if some necessary optimality conditions for the problem are satisfied at the point $\overline{x}$.

These concepts are illustrated in Figure 10.1 for the one dimensional problem: optimize $\theta(x)$ subject to $x \in \mathbf{R}^1$, $a \leqq x \leqq b$. $\theta(x)$ is plotted in Figure 10.1.

The points a, x^5, x^7, x^{10}, x^{12} are strong local minima; x^0, x^4, x^6, x^{11}, b are strong local maxima; x^{12} is the global minimum; x^6 is the global maximum; in this problem. At the point x^3 the derivative of $\theta(x)$ is zero, and so it is a stationary point (satisfies the necessary optimality condition $\frac{d\theta(x)}{dx} = 0$) even though it is neither a local minimum or maximum. In each of the intervals $x^1 \leqq x \leqq x^2$, and $x^8 \leqq x \leqq x^9$, $\theta(x)$ is a constant. x^1, x^2 are weak local minima; and x^8, x^9 are weak local maxima. Every point x satisfying $x^1 < x < x^2$, $x^8 < x < x^9$ is both a weak local minimum and a weak local maximum.

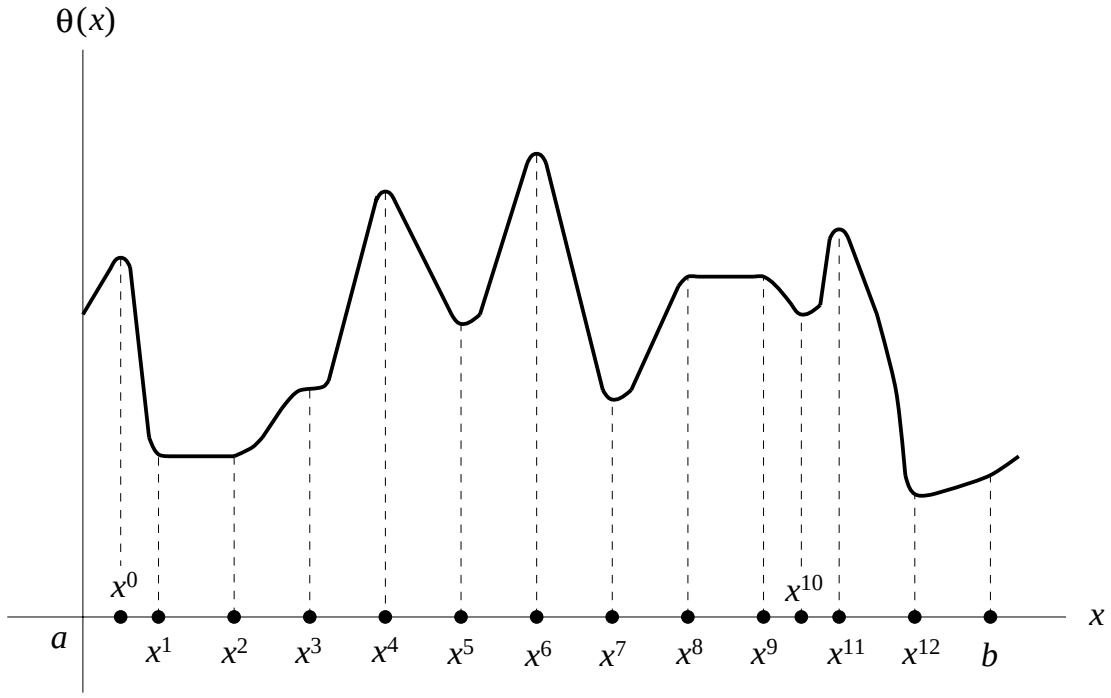


Figure 10.1

10.3 TYPES OF NONLINEAR PROGRAMS, WHAT CAN AND CANNOT BE DONE EFFICIENTLY BY EXISTING METHODS

Every local minimum is a global minimum for the problem of minimizing a convex objective function on a convex set. Likewise, every local maximum is a global maximum for the problem of maximizing a concave function on a convex set. Problems of this type are considered to be nice problems in nonlinear programming, they are called convex programming problems. The other class of NLPs in which a nonconvex objective function is required to be minimized, or in which the set of feasible solutions is not convex, are called nonconvex programming problems.

In general, it is very hard to find the global minimum, or even to check whether a given feasible solution is a global minimum in a nonconvex programming problem. Efforts have been made to find global minima by enumerating all local minima, but these methods tend to be very inefficient. The enormity of this task can be appreciated when we realize that some of the most difficult problems in mathematics that have remained unresolved for centuries, can be posed as nonconvex programming problems.

As an example, consider **Fermat's last Theorem** (unresolved since 1637 AD, see [10.34]) which states that the equation: $x^{\mathbf{n}} + y^{\mathbf{n}} - z^{\mathbf{n}} = 0$, has no solution in integers in the region $x \geqq 1,\ y \geqq 1,\ z \geqq 1,\ n \geqq 3$. Consider the following NLP, where α is some positive parameter, π denotes the irrational number which is the length of the circumference of the circle with unit diameter in $\mathbf{R}^2$, and $\cos\theta$ denotes the cosine function of the angle θ measured in radians.

$$\begin{array}{ll} \text{minimize} & (x^{\mathbf{n}} + y^{\mathbf{n}} - z^{\mathbf{n}})^{\mathbf{2}} + \alpha((-1 + \cos(2\pi x))^{\mathbf{2}} + (-1 + \cos(2\pi y))^{\mathbf{2}} + \\ & (-1 + \cos(2\pi z))^{\mathbf{2}} + (-1 + \cos(2\pi n))^{\mathbf{2}}) \\ \text{subject to} & x, y, z \geqq 1,\ n \geqq 3. \end{array} \tag{10.1}$$

(10.1) is a linearly constrained NLP. It can be verified that Fermat's last Theorem is false iff the optimum objective value in (10.1) is 0 and attained, since any feasible solution (x, y, z, n) to (10.1) which makes the objective value zero provides a counterexample to Fermat's last Theorem. (10.1) is a nonconvex programming problem in which every integer feasible solution is a local minimum. The objective function in (10.1) is a sum of several penalty terms. The number of distinct local minima can be very large even in nonconvex programming problems that do not have such penalty terms in the objective function. As an example, consider the concave minimization problem

$$\begin{array}{ll} \text{minimize} & \theta(x) = -\sum_{j=1}^{n} (x_j - (1/2))^{\mathbf{2}} \\ \text{subject to} & 0 \leqq x_j \leqq 1,\ j = 1 \text{ to } n. \end{array} \tag{10.2}$$

Each of the $2^{\mathbf{n}}$ extreme points of the set of feasible solutions of (10.2) is a local minimum. Unfortunately, there are no techniques known for determining how many local minima a general nonconvex programming problem has, other than plain enumeration. In nonconvex programming problems, since in general it is very difficult to guarantee that a global minimum will be obtained, the best thing that we can expect from an algorithm is that it leads to a point satisfying a necessary condition for being a local mimimum, and many of the descent type methods discussed in this chapter do that. In these methods, the terminal solution obtained may depend on the initial point with which the method is initiated. Usually, by running the algorithm with different initial points, several local minima may be obtained, and the best among them might be a reasonably good solution for the problem.

Starting the algorithm with an initial point, suppose a local minimum $\overline{x}$ is obtained for a nonconvex programming problem. A technique often used to move to a different local minimum is to add a penalty term like $\alpha/(\|x - \overline{x}\|)^{\mathbf{p}}$ where $\alpha > 0$ and $p \geqq 2$, to the objective function, and use the algorithm again on the augmented problem. As x approaches $\overline{x}$, the penalty term $\alpha/(\|x - \overline{x}\|)^{\mathbf{p}}$ blows up to ∞, and this guarantees that the algorithm moves to a point different from $\overline{x}$. But this may not be a satisfactory approach to enumerate the local minima in a nonconvex program, because of the numerical difficulties created by the addition of the penalty terms to avoid previously

obtained local minima. Also, the augmented problem may have new local minima which are not local minima of the original problem.

Because of this, if someone can establish the global minimum in a class of nonconvex programming problems, it is considered to be a mathematical breakthrough and becomes a major international headline item. An example of this is the recent breakthrough on establishing the minimum value of the permanent of a doubly stochastic matrix of order n. Given a square matrix $A = (a_{ij})$ of order n, its permanent is defined by

$$f(A) = \sum([(a_{1p_1}) \dots (a_{np_n})] : \text{ sum over all the } n! \text{ permutations } (p_1, \dots, p_n) \text{ of } \{1, \dots, n\}).$$

A doubly stochastic matrix of order n is a nonnegative square matrix $X = (x_{ij})$ of order n, whose row sums and column sums are all equal to 1. The problem of minimizing the permanent of doubly stochastic matrix of order n is therefore the NLP: find a square matrix $X = (x_{ij})$ of order n to

$$\begin{aligned} &\text{minimize} && f(X) \\ &\text{subject to} && \sum_{j=1}^{n} x_{ij} = 1, \ i = 1 \text{ to } n \\ &&& \sum_{i=1}^{n} x_{ij} = 1, \ j = 1 \text{ to } n \\ &&& x_{ij} \geqq 0, \ i, j = 1 \text{ to } n. \end{aligned}$$

The objective function in this NLP is nonconvex, hence, this is a nonconvex programming problem. In 1926 B. L. vanderWaerden [10.40] conjectured that the global optimum for this problem is the doubly stochastic matrix $(\overline{x}_{ij})$ in which $\overline{x}_{ij} = 1/n$ for all i, j; with an optimum objective value of $n!/n^{\mathbf{n}}$. This conjecture resisted the attacks of many of the world's greatest mathematicians, but was finally resolved in the affirmative by G. P. Egorychev in 1980, see references [10.10, 10.11, 10.20].

10.4 CAN WE AT LEAST COMPUTE A LOCAL MINIMUM EFFICIENTLY?

In convex programming problems, any point satisfying any of the well known necessary optimality conditions such as the KKT conditions, is a local minimum and therefore it is also a global minimum for the problem. To solve a convex programming problem, any algorithm that is guaranteed to find a KKT point, if one exists, is thus adequate. Most of the algorithms for solving NLP's discussed in this book can be shown to converge to a KKT point, if one exists, and so these algorithms compute local, and thus global minima when applied on convex programming problems.

In a nonconvex program, given a feasible solution x satisfying the usual necessary optimality conditions, it may or may not even be a local minimum. If x does not satisfy the sufficient optimality condition given in Appendix 4 for being a local minimum, it may be very hard to verify whether it is a local minimum. As an example, consider the problem discussed in Section 2.9.3

$$\begin{array}{ll} \text{minimize} & x^T D x \\ \text{subject to} & x \geqq 0 \end{array}$$

where D is a given square matrix of order n. When D is not PSD, this NLP is the **simplest nonconvex** NLP.

A sufficient condition for 0 to be a local minimum for this problem is that D be PSD. If D is not PSD, 0 is a local minimum for this problem iff the matrix D is copositive, no efficient methods are known at the moment for doing this. The method discussed in Section 2.9.1 for testing copositiveness is a finite enumeration method, but it may not be practically useful when n is large. As discussed in Section 2.9.3, the problem of checking whether 0 is a local minimum for this problem is a hard problem.

On nonconvex programs involving inequality constraints, existing algorithms can at best guarantee convergence to a KKT point in general. If the KKT point obtained does not satisfy some known sufficient condition for being a local minimum, it is then hard to check whether it is actually a local minimum. However, as mentioned in Section 2.7.6, if the algorithm is based on a descent process (i. e., in a minimization problem, if the algorithm is designed to obtain a sequence of points with decreasing objective values) one can be reasonably confident that the solution obtained is likely to be a local minimum.

10.5 PRECISION IN COMPUTATION

In linear or in convex quadratic programming problems, if all the data are rational numbers, and if an optimum solution exists, there exists an optimum solution which is a rational vector that can be computed exactly with finite precision arithmetic using algorithms like the simplex algorithm or the complementary pivot method discussed earlier. However, in general nonlinear programming, even when the constraints are linear, and all the data in the model is rational, there may be optimum solutions, but no rational optimum solution. For example consider the simple one dimensional optimization problem: find $x \in \mathbf{R}^1$ that minimizes $f(x) = -2x + (x^{\mathbf{3}}/3)$ subject to $x \geqq 0$. The unique optimum solution of this problem is $\overline{x} = \sqrt{2}$, an irrational number, so we can never compute the exact optimum solution of this problem on digital computers that operate with finite precision arithmetic.

Hence, when dealing with general nonlinear programs, emphasis is placed on getting an approximate optimum solution. In practical implementations, nonlinear algorithms are usually terminated when optimality conditions are satisfied to a reasonable degree of approximation, or when it is evident that the algorithm has obtained an interval of sufficiently small length containing the true optimum solution.

10.6 RATES OF CONVERGENCE

The algorithms discussed in this chapter are iterative in nature. They generate a sequence of points $\{x^r : r = 0, 1, 2, \ldots\}$ beginning with an initial point x^0. Under some conditions on the problem being solved, for most of these methods, it is usually possible to prove that the sequence converges in the limit to a point x^* which is a point satisfying the necessary optimality conditions for a local minimum. Even when this convergence is mathematically proven, the method is useful for solving practical problems only if x^r converges rapidly to x^* as r increases. Here we discuss how this rate of convergence is measured mathematically.

Finite Termination Property: The sequence is said to have this property, if there exists a finite value N such that $x^N = x^*$ and the method terminates.

Quadratic Termination Property: The method is said to have this property if the sequence generated terminates in a known finite number of iterations when applied to a strictly convex quadratic function minimization problem.

Suppose the method does not have either of the above properties. Then it generates the truly infinite sequence $\{x^r : r = 0, 1, 2, \ldots\}$. Assume that the sequence converges to x^*, that $x^r \neq x^*$ for any r. The measure of the rate of convergence of this sequence, tries to assess the improvement that occurs in each step, that is, in effect it measures how close x^{r+1} is to x^* compared to the closeness of x^r to x^*, as r goes to ∞. The converging sequence $\{x^r\}$ is said to converge with order k (or to have an asymptotic convergence rate k) if k is the largest number such that $\text{limit}_{r\to\infty}(\|x^{r+1} - x^*\|/\|x^r - x^*\|^{\mathbf{k}}) < \infty$. When $k = 1$, the sequence is said to have **linear** (or **first order**, or **geometric**) convergence rate, if $\text{limit}_{r\to\infty}(\|x^{r+1} - x^*\|/\|x^r - x^*\|) = \gamma < 1$. In this case, the quantity γ is called the convergence ratio of the sequence. If in fact $\gamma = 0$ in this case, the sequence is said to have superlinear convergence rate.

As an example consider the sequence of real numbers $\{\alpha^r : r = 0, 1, \ldots\}$ where $0 < \alpha < 1$. The sequence converges to zero linearly. On the other hand the sequence of real numbers $\{x^r = (1/r) : r = 1, 2, \ldots\}$ converges to zero with $k = 1$, but its rate of convergence is not linear, since $\text{limit}_{r\to\infty}(\|x^{r+1}\|/\|x^r\|) = \text{limit}_{r\to\infty}((r/(r+1)) = 1$ which is not strictly less than one.

If $k = 2$, the sequence $\{x^r\}$ is said to have **quadratic** (or **second order**) **convergence rate**. Quadratic convergence is rapid, since it implies that once the sequence reaches a small neighborhood of x^*, the error in a step decreases as the square of the error in the previous step (i. e. , the number of digits to which x^r agrees with x^* begin to double after each step, after a certain number of steps).

Summary of Later Sections

In the following sections we discuss various descent methods in common use for solving linearly constrained NLPs. These algorithms typically use some unconstrained minimization algorithms and algorithms for solving nonlinear programs in a single variable

(the so-called line minimization algorithms) as subroutines. So we survey these algorithms first.

10.7 SURVEY OF SOME LINE MINIMIZATION ALGORITHMS

The line minimization problem is the problem of minimizing a real valued function $f(\lambda)$ of one variable λ, either over the whole real line, or over the half-line $\lambda \geqq l$ for a specified number l, or over a specified finite interval $[l, u] = \{\lambda : l \leqq \lambda \leqq u\}$. Assuming that $f(\lambda)$ is continuously differentiable, the global minimum for $f(\lambda)$ in the interval $l \leqq \lambda \leqq u$ is the point λ^* in this interval which gives the minimum value for $f(\lambda)$ among those λ satisfying $\frac{df(\lambda)}{d\lambda} = 0$, and the points l, u, if these are finite. In fact if $f(\lambda)$ is concave and l, u are finite, the global minimum for $f(\lambda)$ in the interval $l \leqq \lambda \leqq u$ is either l or u, whichever gives a smaller value for $f(\lambda)$. See Figure 10.2.

In the interval $[a, b]$ if $f'(a) > 0$, a is a local minimum for $f(\lambda)$; and if $f'(b) < 0$, b is a local minimum for $f(\lambda)$.

When $f(\lambda)$ is a general function, a **bracket** is defined to be an interval in the feasible region which contains the minimum. When the derivative $f'(\lambda) = \frac{df(\lambda)}{d\lambda}$ is not available, a bracket usually refers to an interval $[\lambda_1, \lambda_3]$ in the feasible region, satisfying the property that we have a λ_2 satisfying $\lambda_1 < \lambda_2 < \lambda_3$ and $f(\lambda_2) \leqq$ minimum $\{f(\lambda_1), f(\lambda_3)\}$. If the derivative $f(\lambda)$ is available, a bracket usually refers to an interval $[\lambda_1, \lambda_2]$ with $\lambda_1 < \lambda_2$, satisfying the property that $f'(\lambda_1) < 0$ and $f'(\lambda_2) > 0$.

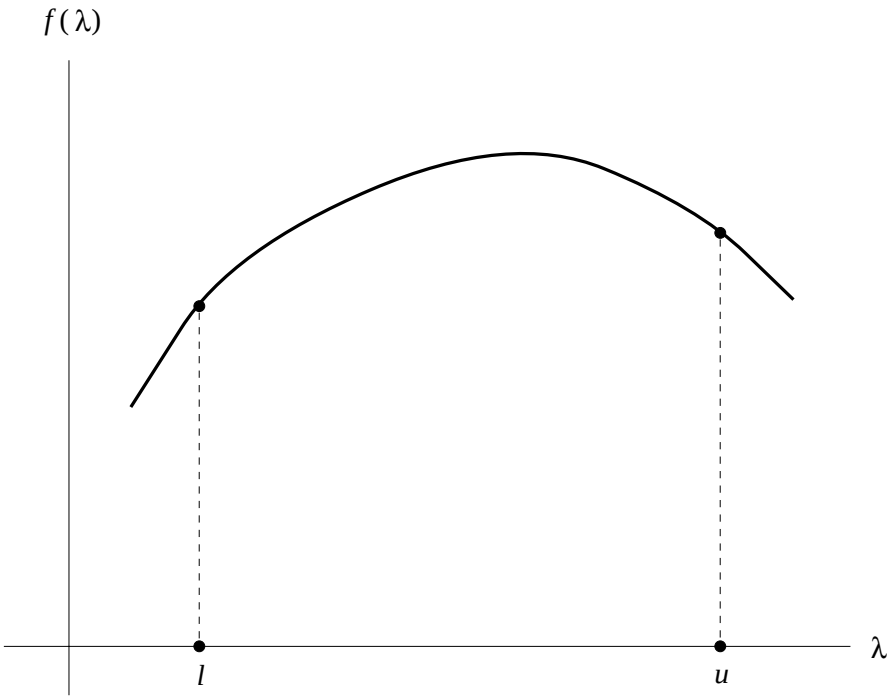


Figure 10.2 The global minimum for one dimensional concave minimization problem is a boundary point (l here).

How to Select an Initial Bracket?

First consider the problem in which we are required to minimize $f(\lambda)$ over the entire real line. Begin with an initial point λ_0 and choose a positive step lenth Δ. Compute $f(\lambda_0)$ and $f(\lambda_1)$, where $\lambda_1 = \lambda_0 + \Delta$. If $f(\lambda_1) < f(\lambda_0)$, the direction of increasing λ is the right direction to pursue; otherwise, replace Δ by $-\Delta$ to reverse the direction and go through the procedure discussed next. Define $\lambda_r = \lambda_{r-1} + 2^{\mathbf{r-1}}\Delta$ for $r = 2, 3, \ldots$ as long as they keep on decreasing, until either the upper bound on λ is reached or a value k for r is found such that $f(\lambda_{k+1}) > f(\lambda_k)$. In this case we have λ_{k-1}, λ_k, λ_{k+1} satisfying $f(\lambda_k) < f(\lambda_{k-1})$, $f(\lambda_{k+1}) > f(\lambda_k)$. Among the four points λ_{k-1}, λ_k, $(\lambda_k + \lambda_{k+1})/2$, and λ_{k+1}, drop either λ_{k-1} or λ_{k+1}, whichever is farther from the point in the pair $\{\lambda_k, (\lambda_k + \lambda_{k+1})/2\}$ that yields the smallest value to $f(\lambda)$. Let the remaining points be called λ_a, λ_b, λ_c, where $\lambda_a < \lambda_b < \lambda_c$. These points are equi-distant, and $f(\lambda_b) \leqq f(\lambda_c)$, $f(\lambda_b) \leqq f(\lambda_a)$. So this interval λ_a to λ_c brackets the minimum.

If the problem is to minimize $f(\lambda)$ over $\lambda \geqq l$ or $u \geqq \lambda \geqq l$, it is reasonable to expect that $f(\lambda)$ decreases as λ increases through l (i. e., the derivative $f'(l) < 0$, otherwise l is itself a local minimum for the problem). So in these problems, we can get a bracket by beginning with $\lambda_0 = l$ and applying the above procedure.

10.7.1 The Golden Section Search Method

The function $f(\lambda)$ is said to be a unimodal function in the interval $a \leqq \lambda \leqq b$ if it has a unique local minimum in the interval. See Figures 10.3, 10.4.

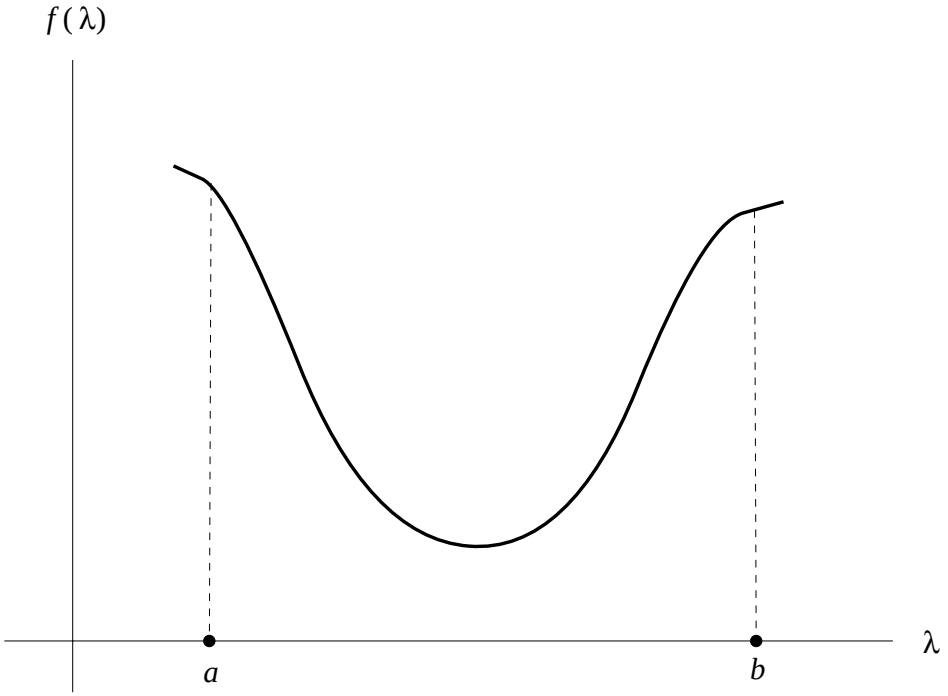


Figure 10.3 A unimodal function in the interval $[a, b]$.

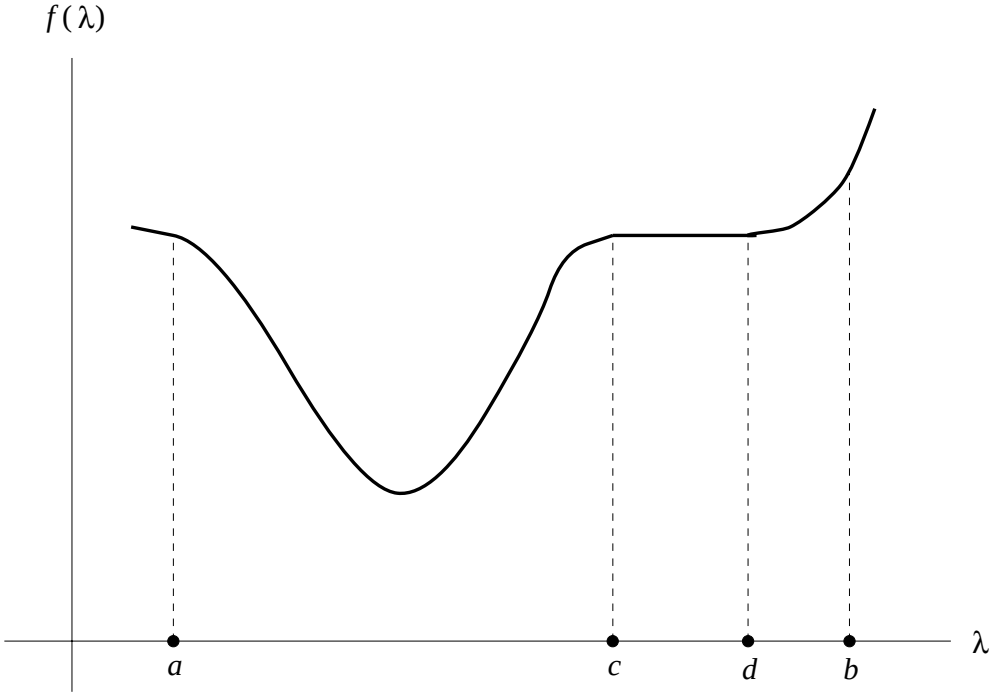


Figure 10.4 This function is constant in the interval $c < \lambda < d$, so every point in this interval is a local minimum. So this function is not unimodal in the interval $[a, b]$.

In many practical applications, it is reasonable to assume that the interval has been narrowed down using prior knowledge of the problem such that the objective function has a single minimum in the interval. A unimodal function $f(\lambda)$ in the interval $a \leqq \lambda \leqq b$ satisfies the property that there exists a unique λ_* in the interval (this λ_* is the minimum) such that given any λ_1, λ_2 in the interval with $\lambda_1 < \lambda_2$, if $\lambda_2 < \lambda_*$ we have $f(\lambda_1) > f(\lambda_2)$; and if $\lambda_* < \lambda_1$, we have $f(\lambda_1) < f(\lambda_2)$. The golden section search method is a method for minimizing a unimodal function in an interval by sectioning (i. e., interval reduction) using only function values evaluated at selected points.

The number $\tau \equiv 2/(1+\sqrt{5}) \simeq .618$ is known as the golden ratio. Let $[\alpha, \beta]$ be the current interval in which the minimum is known to lie. If function value has not been evaluated at any interior point in this interval, let $\lambda_1 = \alpha + .382(\beta - \alpha)$, $\lambda_2 = \alpha + .618(\beta - \alpha)$, evaluate $f(\lambda_1)$, $f(\lambda_2)$ (depending on what happened in the previous step, it is possible that the function value at one of these points λ_1 or λ_2 has already been computed in the previous steps). If $f(\lambda_1) < f(\lambda_2)$, the minimum is contained in the interval $[\alpha, \lambda_2]$. If $f(\lambda_1) > f(\lambda_2)$, the minimum is contained in the interval $[\lambda_1, \beta]$. If $f(\lambda_1) = f(\lambda_2)$, the minimum is contained in the interval $[\lambda_1, \lambda_2]$. Repeat this process with the new interval.

There is a reduction in the length of the interval of uncertainty (i. e., the bracket length) by a factor of .618 or more in each step. The length of the interval of uncertainty converges linearly to zero. When the length of the interval of uncertainty has become less than a specified tolerance, ε, any point in the final interval could be taken as an approximation for the minimum.

10.7.2 The Method of Bisection

This method can be used if $f(\lambda)$ is continuously differentiable and the derivative $f'(\lambda)$ can be computed. It starts with an initial bracket for the minimum $[a, b]$ satisfying $f'(a) < 0$ and $f'(b) > 0$. Evaluate $f'((a+b)/2)$. If $f'((a+b)/2) = 0$, the point $(a+b)/2$ satisfies the first order necessary condition for a local minimum. If $f'((a+b)/2) > 0$, take $[a, (a+b)/2]$ as the new bracket and continue. If $f'((a+b)/2) < 0$, take $[(a+b)/2, b]$ as the new bracket and continue.

Since the bracket is cut in half each time, the length of this interval converges to zero linearly. When its length has become less than a specified tolerance ε, any point in the final interval could be taken as an approximation to the minimum. One disadvantage of this method is that it relies totally on the values of the derivative $f'(\lambda)$ and does not use the values of the function $f(\lambda)$ being minimized.

10.7.3 Newton's Method

This is a second order gradient method that can be used if $f(\lambda)$ is twice continuously differentiable and the second derivative $f''(\lambda)$ can be computed easily either through

a subroutine or by using a finite difference approximation, and $f(\lambda)$ is required to be minimized over the entire real line. The method is the application of the Newton-Raphson method to find a solution of the equation: $f'(\lambda) = 0$. The method generates a sequence $\{\lambda_r : r = 0, 1, \ldots\}$ beginning with an initial point λ_0. Given λ_r, the second order Taylor series approximation for $f(\lambda)$ at λ_r is $f(\lambda_r) + f'(\lambda_r)(\lambda - \lambda_r) + (1/2)f''(\lambda_r)(\lambda - \lambda_r)^{\mathbf{2}}$. If $f''(\lambda_r) > 0$, this has a minimum at

$$\lambda_{r+1} = \lambda_r - f'(\lambda_r)/f''(\lambda_r). \tag{10.3}$$

Equation (10.3) gives the iterative scheme for Newton's method. The method is not suitable to be used if $f''(\lambda)$ turns out to be $\leqq 0$ at any point encountered during the algorithm. It is quite suitable if an initial point λ_0 in the vincinity of a local minimum is known. In the vincinity of a minimum, the second derivative $f''(\lambda)$ is of constant sign (nonnegative) and the first derivative $f'(\lambda)$ changes sign from a negative to a positive value. If $f(\lambda)$ is a quadratic function with a minimum, this method finds the minimum in one step. In general, any twice continuously differentiable function has a Taylor series expansion around a point, the first three terms of this series (which form a quadratic function) are dominant when the point is in the vincinity of the minimum. The method has rapid convergence (quadratically) once the vincinity of the minimum is reached. A result on the convergence rate of this method follows as a corollary of Theorem 10.1, where a convergence rate result for Newton's method applied to find the unconstrained minimum of a real valued function $\theta(x)$ over $x \in \mathbf{R}^n$ is proved. See references [10.9, 10.13, 10.26, 10.33] for results on the convergence and rates of convergence of Newton's method.

10.7.4 Modified Newton's Method

Several modifications have been proposed for Newton's method to handle cases where a good initial point is not available to initiate Newton's method, or when a point satisfying $f''(\lambda) \leqq 0$ is encountered during the method, and to handle the problem in which the feasible range is a specified interval and not the entire real line. We discuss one such modification here. We consider the problem of minimizing a twice continuously differentiable function $f(\lambda)$ in the interval $[a, c] = \{\lambda : a \leqq \lambda \leqq c\}$ and we have a piont b satisfying $a < b < c$ and $f(b) <$ minimum $\{f(a), f(c)\}$. This method generates a sequence of points $\{\lambda_r : r = 0, 1, \ldots\}$ satisfying the property that the entire sequence lies in the interval $[a, c]$ and that $f(\lambda_{r+1}) < f(\lambda_r)$ for all r. Initiate the method with $\lambda_0 = b$, and select a constant α satisfying $0 < \alpha < 1$. The quantity α is called the **attenuation factor**.

Given λ_r, the point obtained by moving in the direction of $f'(\lambda_r)$ a step of length β is $\lambda_r - \beta f'(\lambda_r)$. From the Taylor series, $f(\lambda_r - \beta f'(\lambda_r)) = f(\lambda_r) - \beta(f'(\lambda_r))^{\mathbf{2}} +$ error term, where the error term tends to zero faster than β. So, if $\beta > 0$, we make improvement in the objective value by this move. Notice that Newton's method takes

$\beta = 1/f''(\lambda_r)$ to get the next point in the sequence. In this method you do the following.

(a) If $f''(\lambda_r) > 0$ compute $y_r = \lambda_r - f'(\lambda_r)/f''(\lambda_r)$. If $y_r \in [a, c]$ and $f(\lambda_r) - f(y_r) \geqq (\alpha/2)(f'(\lambda_r))^{\mathbf{2}}/f''(\lambda_r)$, define $\lambda_{r+1} = y_r$. If $y_r \in [a, c]$ but $f(\lambda_r) - f(y_r) < (\alpha/2)(f'(\lambda_r))^{\mathbf{2}}/f''(\lambda_r)$, use a first order Armijo step size procedure which requires the determination of the smallest nonnegative integer s satisfying

$$(\lambda_r - f'(\lambda_r)/2^{\mathbf{s}}) \in [a, c], \text{ and}$$
$$f(\lambda_r) - f(\lambda_r - f'(\lambda_r)/2^{\mathbf{s}}) > (\alpha/2^{\mathbf{s}})(f'(\lambda_r))^{\mathbf{2}}$$

and then define $\lambda_{r+1} = \lambda_r - f'(\lambda_r)/2^{\mathbf{s}}$. The motivation for this step size procedure is explained in Section 10.8.1.

(b) If $f''(\lambda_r) \leqq 0$, define $\delta = -1$ if $f'(\lambda_r) \geqq 0$, $+1$ if $f'(\lambda_r) < 0$ and use the second order Armijo step size procedure. This requires the determination of the smallest nonnegative integer s satisfying

$$(\lambda_r - (f'(\lambda_r)/2^{\mathbf{s}}) + (\delta/2^{\mathbf{s/2}})) \in [a, c], \text{ and}$$
$$f(\lambda_r) - f(\lambda_r - (f'(\lambda_r)/2^{\mathbf{s}}) + \delta/2^{\mathbf{s/2}}) \geqq \alpha(((f'(\lambda_r))^{\mathbf{2}}/2^{\mathbf{s}}) - f''(\lambda_r)/2^{\mathbf{s+1}}).$$

For a finite s satisfying these conditions to exist, it is sufficient that $f''(\lambda_r) < 0$ if $f'(\lambda_r) = 0$. Then define $\lambda_{r+1} = \lambda_r - (f'(\lambda_r)/2^{\mathbf{s}}) + \delta/2^{\mathbf{s/2}}$.

Under certain conditions it can be shown that this method has second-order convergence. See references [10.1, 10.26, 10.27].

10.7.5 Secant Method

In Newton's method or modified Newton's method discussed above, we need to compute the value of the second derivative $f''(\lambda_r)$. This may be hard. In the secant method we replace $f''(\lambda_r)$ by its finite difference approximation $(f'(\lambda_r) - f'(\lambda_{r-1}))/(\lambda_r - \lambda_{r-1})$. This is the only change in the Secant method from Newton's or modified Newton's method. The secant method is initiated with two initial points λ_0, λ_1 in the feasible region satisfying $\lambda_0 < \lambda_1$ and $f'(\lambda_0) < 0$, $f'(\lambda_1) > 0$.

10.7.6 The Method of False Position

In the secant method we always use $f'(\lambda_r)$ and $f'(\lambda_{r-1})$ to get a finite difference approximation for $f''(\lambda_r)$ for each r. Even though initially $f'(\lambda_0)$, $f'(\lambda_1)$ are of opposite signs, after some steps it may happen that $f'(\lambda_r)$ and $f'(\lambda_{r-1})$ have the same sign, and this could make the iterates diverge when minimizing over the real line. In this method we make sure that $f''(\lambda)$ is always approximated using the values of $f'(\lambda)$ of opposite signs at two different values of λ. For some r, suppose $f''(\lambda_r)$ was approximated using $f'(\lambda_r)$ and $f'(\lambda_s)$ for an $s \leqq r - 1$. Compute λ_{r+1} using this approximation as under the secant method, and compute $f'(\lambda_{r+1})$. Determine which of $f'(\lambda_t)$ for $t = r$ or s has a sign opposite to that of $f'(\lambda_{r+1})$. Then approximate $f''(\lambda_{r+1})$ by $f'(\lambda_{r+1} - f'(\lambda_t))/(\lambda_{r+1} - \lambda_t)$, and continue in the same way.

10.7.7 Univariate Minimization by Polynomial Approximation Methods

The essential feature of these methods is to approximate the original function $f(\lambda)$ by a simpler function $P(\lambda)$ (normally a second or third degree polynomial) by curve fitting, and then using the minimum of $P(\lambda)$ to approximate that of $f(\lambda)$. These methods are also called polynomial interpolation methods. If the minimum is known to lie in a small enough interval, the application of these methods usually produces very satisfactory results.

Quadratic Interpolation

This method needs an interval of the form $\lambda_1 < \lambda_2 < \lambda_3$ with $f(\lambda_2) < \min\{f(\lambda_1), f(\lambda_3)\}$, a bracket for the minimum, as discussed earlier. λ_2, the initial best point, is the initial point in the sequence. It constructs a quadratic approximation $P(\lambda) = a\lambda^{\mathbf{2}} + b\lambda + c$ which coincides with $f(\lambda)$ at $\lambda = \lambda_1, \lambda_2, \lambda_3$. By the properties mentioned above, $P(\lambda)$ determines a parabola. The three independent pieces of information (value of $P(\lambda)$ = value of $f(\lambda)$ at $\lambda = \lambda_1, \lambda_2, \lambda_3$) are used to determine a, b, c in $P(\lambda)$ uniquely. Since $P(\lambda)$ is a parabola (by the condition imposed), the minimum of $P(\lambda)$ lies in the interval $[\lambda_1, \lambda_3]$ at the point λ satisfying $\frac{dP(\lambda)}{d\lambda} = 0$. It can be verified that this point is

$$\lambda_* = \frac{(\lambda_2^{\mathbf{2}} - \lambda_3^{\mathbf{2}})f(\lambda_1) + (\lambda_3^{\mathbf{2}} - \lambda_1^{\mathbf{2}})f(\lambda_2) + (\lambda_1^{\mathbf{2}} - \lambda_2^{\mathbf{2}})f(\lambda_3)}{2[(\lambda_2 - \lambda_3)f(\lambda_1) + (\lambda_3 - \lambda_1)f(\lambda_2) + (\lambda_1 - \lambda_2)f(\lambda_3)]}$$

λ_* is a minimum for $P(\lambda)$ if

$$\frac{(\lambda_2 - \lambda_3)f(\lambda_1) + (\lambda_3 - \lambda_1)f(\lambda_2) + (\lambda_1 - \lambda_2)f(\lambda_3)}{(\lambda_1 - \lambda_2)(\lambda_2 - \lambda_3)(\lambda_3 - \lambda_1)} < 0$$

a condition which will hold because of the properties satisfied by λ_1, λ_2, λ_3.

It is possible for λ_* to be equal to λ_2 even though this point is far away from a local minimum of $f(\lambda)$. See Figure 10.5. If this happens, the quadratic interpolation has failed to generate a new trial point.

If $|\lambda_* - \lambda_2|$ is not too small, we can replace one of the points in λ_1, λ_2, λ_3 by λ_* so that the new set of three points again satisfies the conditions for a bracket for the minimum of $f(\lambda)$. The best point among these three is the next point in the sequence, and the procedure is repeated with the new bracket. If λ_* and λ_2 are too close (even if they are not equal) repeating the procedure with such close values could lead to numerical problems in the next step. In this case, we select a small distance δ, and take the new point to be either $\lambda_* + \delta$ or $\lambda_* - \delta$ whichever leads to the smallest length new bracket.

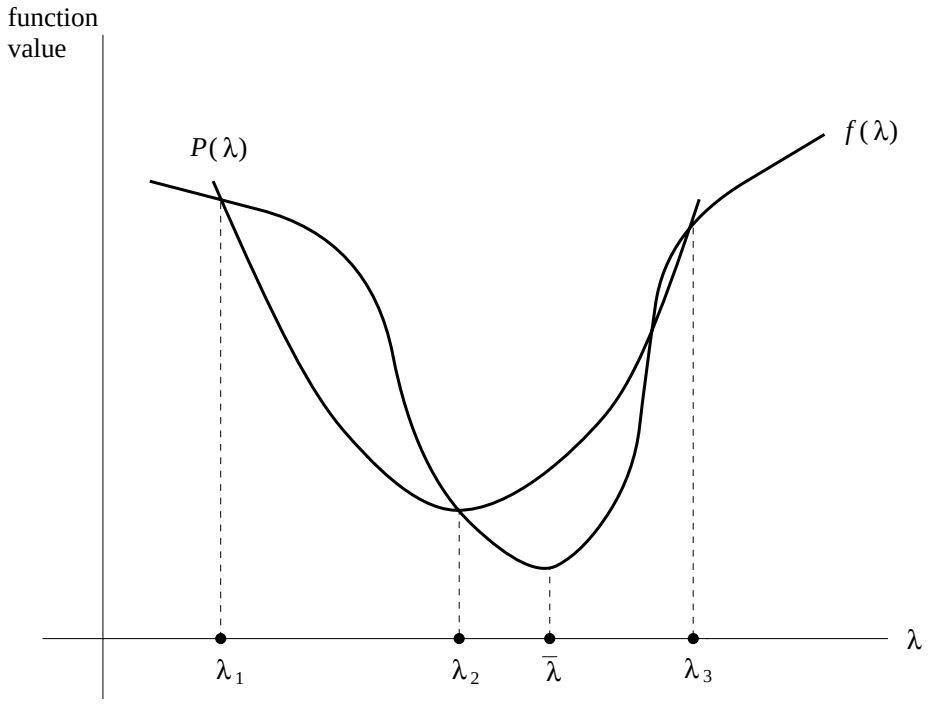


Figure 10.5 The minimum of $f(\lambda)$ in the bracket $[\lambda_1, \lambda_3]$ is at $\overline{\lambda}$. But the minimum of Quadratic approximation, λ_* is the same as λ_2.

Note 10.1 Newton's method is a quadratic approximation method. Given the current point λ_r at which the second derivative $f''(\lambda_r) > 0$, Newton's method constructs a quadratic function $P(\lambda)$ satisfying the three properties $P(\lambda_r) = f(\lambda_r)$, $P'(\lambda_r) = f'(\lambda_r)$ and $P''(\lambda_r) = f''(\lambda_r)$. It can be verified that the function $P(\lambda)$ is just the second order Taylor series approximation to $f(\lambda)$ around this point λ_r, and that the next point in the sequence λ_{r+1} is the minimum of this quadratic approximation $P(\lambda)$.

Cubic Interpolation Method

This method can be used when $f(\lambda)$ is differentiable and the derivative $f'(\lambda)$ can be computed either numerically using a finite difference approximation or computed directly using a subroutine for evaluating it. The method needs a bracket $[\lambda_1, \lambda_2]$ satisfying the property that $f'(\lambda_1) < 0$, $f'(\lambda_2) > 0$. A cubic function $P_3(\lambda) = a\lambda^{\mathbf{3}} + b\lambda^{\mathbf{2}} + c\lambda + d$ can be fitted such that it agrees in value with $f(\lambda)$ at λ_1 and λ_2 and its derivative has the same value as $f'(\lambda)$ at λ_1 and λ_2. From the bracket conditions the minimum of this cubic function occurs inside the bracket at the point λ_* satisfying

$\frac{d}{d\lambda}(P_3(\lambda)) = 0$. It can be verified that

$$\lambda_* = \lambda_1 + (\lambda_2 - \lambda_1)\Big(1 - \frac{f'(\lambda_2) + \nu - \eta}{f'(\lambda_2) - f'(\lambda_1) + 2\nu}\Big)$$

where

$$\eta = \frac{3(f(\lambda_1) - f(\lambda_2))}{\lambda_2 - \lambda_1} + f'(\lambda_1) + f'(\lambda_2)$$
$$\nu = (\eta^{\mathbf{2}} - f'(\lambda_1)f'(\lambda_2))^{\mathbf{1/2}}.$$

If $|f'(\lambda_*)|$ is small, λ_* can be accepted as a good approximation for the minimum. Otherwise, if $f'(\lambda_*) > 0$, repeat the process with $[\lambda_1, \lambda_*]$ as the new bracket. If $f'(\lambda_*) < 0$, repeat the process with $[\lambda_*, \lambda_2]$ as the new bracket.

It can be shown that these polynomial approximation methods have superlinear or better convergence rate under certain conditions. See [10.13, 10.17, A8]. It is possible to develop algorithms based on a combination of sectioning and polynomial interpolation steps.

Difficulty in Computing Derivatives During Line Minimization Steps Encountered in Solving NLPs Involving Several Variables

Let $\theta(x)$ be a continuously differentiable real valued function defined on $\mathbf{R}^n$. Consider the NLP in which $\theta(x)$ is to be minimized, possibly subject to some constraints. Many algorithms for solving such a problem make repeated use of line minimization algorithms to solve problems of the form: given a point $x^0 \in \mathbf{R}^n$ and a search direction $y \in \mathbf{R}^n$, $y \neq 0$, find the step length λ that minimizes $\theta(x^0 + \lambda y)$ subject to $\lambda \geqq 0$.

In this problem, since x^0 and y are given vectors, $\theta(x^0 + \lambda y) = f(\lambda)$ is purely a function of the step length parameter λ. If the problem of minimizing $f(\lambda)$ in $\lambda \geqq 0$ needs the derivative $f'(\lambda)$ for some given value of λ, we use

$$f'(\lambda) = \frac{d}{d\lambda}(\theta(x^0 + \lambda y)) = (\nabla\theta(x^0 + \lambda y))y$$

where $\nabla\theta(x^0 + \lambda y)$ is the row vector of partial derivatives of $\theta(x)$ evaluated at $x = x^0 + \lambda y$. So, the computation of $f'(\lambda)$ needs the evaluation of each of the partial derivatives of $\theta(x)$ at the point $x^0 + \lambda y$, which in the worst case takes n function evaluations (the work would be less if, for example, we know from the structure of $\theta(x)$ that some of these partial derivatives are zero). Thus, evaluating $f(\lambda) = \theta(x^0 + \lambda y)$ needs only one function evaluation; while evaluating $f'(\lambda)$ needs n function evaluations, considerably more work. In the same manner, evaluation of the second derivative $f''(\lambda)$ for any λ, needs $n^{\mathbf{2}}$ function evaluations in the worst case. These facts should be considered in choosing an algorithm for line minimization, to be used as a subroutine in algorithms for NLPs involving many variables. Since evaluating derivatives ($f'(\lambda)$ or $f''(\lambda)$) requires a

lot more function evaluations, typically line minimization algorithms based on function values only, are to be preferred as far as possible.

When $f(\overline{\lambda}) = \theta(x^0 + \overline{\lambda}y)$, the formula $f'(\overline{\lambda}) = (\nabla\theta(x^0 + \overline{\lambda}y))y$ is an analytical formula for the exact derivative of $f(\lambda)$ at $\overline{\lambda}$, and the value of $f'(\overline{\lambda})$ computed using this formula is known as the analytically computed derivative. Since the analytical computation of the derivative is so expensive, it may be appropriate to use an approximation for it. Let ε be a small positive number, it is called the finite difference interval. Then $f'(\overline{\lambda})$ can be approximated by any of the three following quantities

$$\frac{f(\overline{\lambda}) - f(\overline{\lambda} - \varepsilon)}{\varepsilon}$$

$$\text{or} \quad \frac{f(\overline{\lambda} + \varepsilon) - f(\overline{\lambda})}{\varepsilon}$$

$$\text{or} \quad \frac{f(\overline{\lambda} + \varepsilon) - f(\overline{\lambda} - \varepsilon)}{2\varepsilon}.$$

The topmost quantity is called the backward-difference approximation, the middle quantity is known as the forward-difference approximation, and the bottom quantity is known as the central-difference approximation, to $f'(\overline{\lambda})$. If the value of $f(\overline{\lambda})$ is already known, the computation of the forward or backward-difference approximation to $f'(\overline{\lambda})$ needs one more function evaluation, whereas the computation of the central-difference approximation needs two more function evaluations. If ε is small compared to $|f'(\overline{\lambda})|$ and the magnitude of $|f''(\lambda)|$ in the neighborhood of $\overline{\lambda}$, the error in approximation will be small, because $f(\overline{\lambda} + \varepsilon) = f(\overline{\lambda}) + \varepsilon f' + \frac{\varepsilon^2}{2} f''(\overline{\lambda} + \gamma)$ for some $0 \leqq \gamma \leqq \varepsilon$, by Taylor's theorem. Thus with a suitable choice of the finite difference interval, these finite difference approximations provide a reasonable approximation to the derivative, with much less computational effort than that involved in using the analytical formula. Because of this, many professional software packages for NLP algorithms use finite difference approximations to the derivatives.

Even the partial derivatives of $\theta(x)$ can be approximated by finite difference approximations. Let I be the unit matrix of order n. Then

$$\frac{\theta(x) - \theta(x - \varepsilon I_{.j})}{\varepsilon}$$

$$\text{or} \quad \frac{\theta(x + \varepsilon I_{.j}) - \theta(x)}{\varepsilon}$$

$$\text{or} \quad \frac{\theta(x + \varepsilon I_{.j}) - \theta(x - \varepsilon I_{.j})}{2\varepsilon}$$

where ε is the suitable finite difference interval, are the backward, forward and central-difference approximations for the partial derivative $\frac{\partial\theta(x)}{\partial x_j}$, respectively.

10.7.8 Practical Termination Conditions for Line Minimization Algorithms

In practice, line minimization algorithms discussed above are terminated either when the bracket length is small, or when a point λ satisfying $|f'(\lambda)| < \varepsilon$ for some specified tolerance ε is obtained, or when the improvement in objective value between two consecutive points obtained in the method is small, or when the difference between two consecutive points obtained under the method is small. At termination, if we have a bracket for the minimum, a final interpolation step can be carried out to provide the approximate location of the minimum in the bracket.

10.7.9 Line Minimization Algorithms Based on Piecewise Linear and Quadratic Approximations

In this section we discuss new line minimization algorithms based upon a combination of piecewise linear (or polyhedral) and quadratic approximations, due to C. Lemarechal and R. Mifflin [10.23, 10.28, 10.29]. These algorithms are rapidly convergent, and seem best suited as line search subroutines in higher dimensional optimization algorithms.

Let $f(\lambda) : \mathbf{R}^1 \to \mathbf{R}^1$ be the real valued function defined on $\mathbf{R}^1$ which is required to be minimized over $\lambda \in \mathbf{R}^1$. At any given λ, the limit (if it exists) of $\frac{f(\lambda+\varepsilon)-f(\lambda)}{\varepsilon}$ as $\varepsilon \to 0$ through positive values is known as the right derivative of $f(\lambda)$ at λ and denoted by $f'_+(\lambda)$, the limit of the same quantity as $\varepsilon \to 0$ through negative values is known as the left derivative of $f(\lambda)$ at λ and is denoted by $f'_-(\lambda)$. If $f(\lambda)$ is differentiable at λ, then $f'_-(\lambda) = f'_+(\lambda) = f'(\lambda)$. If $f(\lambda)$ is convex, these $f'_-(\lambda)$ and $f'_+(\lambda)$ exist and they satisfy

$$\text{if } \lambda < \gamma, \text{ then } f'_-(\lambda) \leqq f'_+(\lambda) \leqq f'_-(\gamma) \leqq f'_+(\gamma).$$

When $f(\lambda)$ is convex, the subdifferential $\partial f(\lambda)$ is the line segment $[f'_-(\lambda), f'_+(\lambda)]$, and a necessary and sufficient condition for a λ_* to be the minimizer point for $f(\lambda)$ is:

$$f'_-(\lambda_*) \leqq 0 \leqq f'_+(\lambda_*).$$

For the moment, let $g(\lambda)$ denote the derivative $f'(\lambda)$ if it exists; or a number from $\partial f(\lambda)$, that is, a subgradient of $f(\lambda)$ at λ, otherwise. Given two points $\overline{\lambda}$ and $\overline{\gamma}$ satisfying the properties that $f(\overline{\lambda}) \leqq f(\overline{\gamma})$ and $g(\overline{\lambda})g(\overline{\gamma}) < 0$, the interval between $\overline{\lambda}$ and $\overline{\gamma}$ is a bracket for the minimum.

Polyhedral Approximation

The affine functions $f(\overline{\lambda}) + g(\overline{\lambda})(\lambda - \overline{\lambda})$, $f(\overline{\gamma}) + g(\overline{\gamma})(\lambda - \overline{\gamma})$, are the linearizations of $f(\lambda)$ at $\lambda = \overline{\lambda}, \overline{\gamma}$ respectively. The pointwise supremum function $P(\lambda) = \max.\{f(\overline{\lambda}) + g(\overline{\lambda})(\lambda-\overline{\lambda}), f(\overline{\gamma})+g(\overline{\gamma})(\lambda-\overline{\gamma})\}$ provides a piecewise linear or polyhedral approximation

for $f(\lambda)$ in the interval between $\overline{\lambda}$ and $\overline{\gamma}$. If $f(\lambda)$ is convex, this piecewise linear function underestimates $f(\lambda)$ at each point in the interval, see Figure 10.6. The point where this piecewise linear function attains its minimum is the point that equalizes the two expressions inside the max., it is $\overline{\lambda} + d^P$, where

$$d^P = \frac{f(\overline{\lambda}) - f(\overline{\gamma}) - g(\overline{\gamma})(\overline{\lambda} - \overline{\gamma})}{g(\overline{\gamma}) - g(\overline{\lambda})}.$$

This d^P provides the polyhedral approximation step from the point $\overline{\lambda}$ for the line minimization problem. If $f(\lambda)$ is convex, the numerator in d^P is $\geqq 0$ and $\overline{\lambda} + d^P$ lies in the interval between $\overline{\lambda}$ and $\overline{\gamma}$.

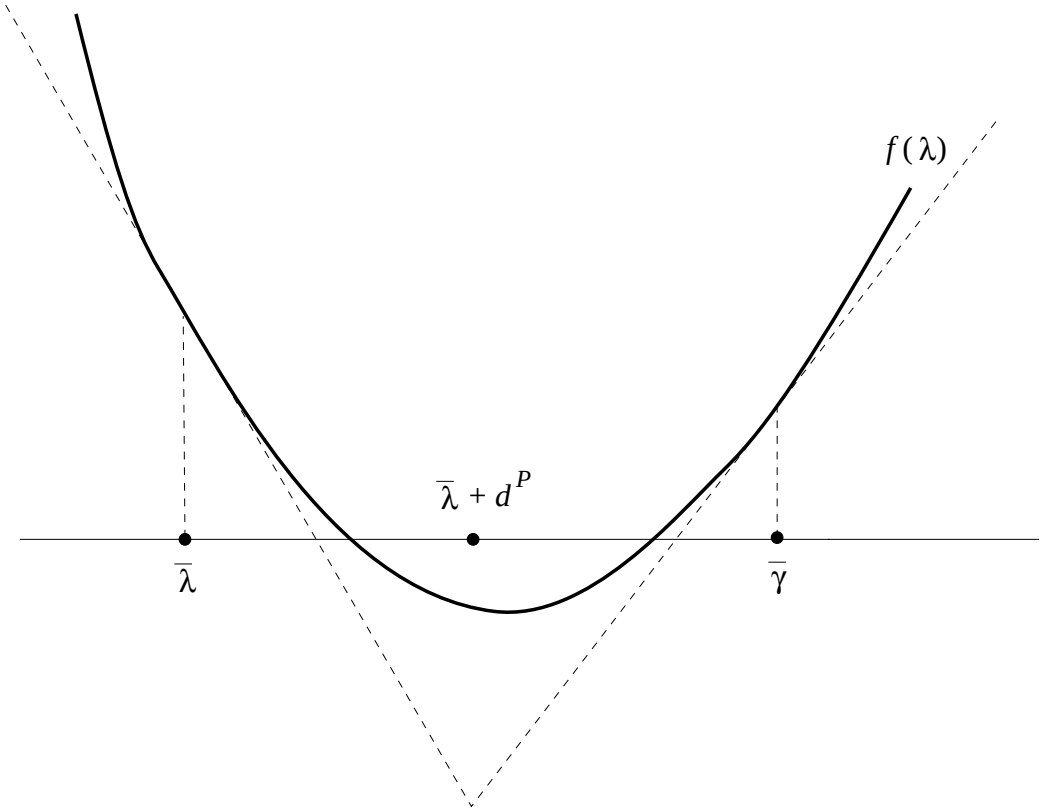


Figure 10.6 A Polyhedral approximation (the dashed lines) for $f(\lambda)$, and the point $\overline{\lambda} + d^P$ where it attains its minimum.

Quadratic Approximation

A quadratic approximation for $f(\lambda)$ at $\lambda = \overline{\lambda}$ is of the form

$$Q(\lambda) = f(\overline{\lambda}) + g(\overline{\lambda})(\lambda - \overline{\lambda}) + \frac{1}{2}(\lambda - \overline{\lambda})^2 G(\overline{\lambda})$$

where $G(\overline{\lambda})$ approximates the second derivative of $f(\lambda)$ and is determined in a one-sided secant manner, that is,

$$G(\overline{\lambda}) = \frac{g(\overline{\lambda}) - g(t)}{\overline{\lambda} - t}$$

where t is a point such that $\overline{\lambda}$ is in the interval between t and the minimizer of $f(\lambda)$. If $f(\lambda)$ is convex, $G(\lambda)$ is ≥ 0. If $G(\overline{\lambda}) > 0$, the minimum of $Q(\lambda)$ is attained at $\overline{\lambda} + d^Q$ where

$$d^Q = \frac{-g(\overline{\lambda})}{G(\overline{\lambda})}.$$

If $G(\lambda) \leqq 0$, $|d^Q| = +\infty$. d^Q is the quadratic approximation step from $\overline{\lambda}$ for the line minimization problem.

The algorithm uses a step that is the shorter of the quadratic approximation and the polyhedral approximation steps. Some modifications are made to these steps if the functions are not convex, to guarantee convergence to at least a stationary point.

These methods generate two sequences $\{\lambda_r\}$, $\{\gamma_r\}$ where for each r, λ_r and γ_r are on opposite sides of the minimizing point λ_*. The sequence $\{f(\lambda_r)\}$ will be non-increasing, and $|\lambda_r - \gamma_r|$ is a decreasing sequence, since at least one of the two points λ_r, γ_r changes in each step.

We describe different versions of the algorithm in various numbered subsections in the following, for ease of cross referencing.

1 Line Minimization of a Convex Function

Assume that $f(\lambda)$ is convex and that it is required to find the point λ_* that minimizes $f(\lambda)$ over $\lambda \in \mathbf{R}^1$. In this subsection, $g(\lambda)$ denotes $f'(\lambda)$ if $f(\lambda)$ is differentiable at λ, or a subgradient of $f(\lambda)$ at λ otherwise (i. e., a point from $\partial f(\lambda)$, the interval between $f'_-(\lambda)$ and $f'_+(\lambda)$). The method initially needs two points λ_1 and γ_1 satisfying

$$f(\lambda_1) \leqq f(\gamma_1) \text{ and } g(\lambda_1)g(\gamma_1) < 0.$$

A pair of points like this can be generated by some initialization procedure. In this case λ_* is in the interval between λ_1 and γ_1. Choose $G(\lambda_1) = (g(\gamma_1) - g(\lambda_1))/(\gamma_1 - \lambda_1)$. We will now describe the general step.

Step r. At the beginning of this step we have λ_r, γ_r satisfying

$$f(\lambda_r) \leqq f(\gamma_r) \text{ and } g(\lambda_r)g(\gamma_r) < 0$$

and we also have $G(\lambda_r)$. Compute

$$d_r^P = \frac{f(\lambda_r) - (f(\gamma_r) + g(\gamma_r)(\lambda_r - \gamma_r))}{g(\gamma_r) - g(\lambda_r)}$$

$$d_r^Q = \frac{-g(\lambda_r)}{G(\lambda_r)}$$

where $|d_r^Q| = +\infty$ if $G(\lambda_r) = 0$. Now determine

$$d_r = (\text{sign of } (-g(\lambda_r)))(\min.\{|d_r^P|, |d_r^Q|\})$$
$$\nu_r = \lambda_r + d_r.$$

Terminate with the conclusion that ν_r is the minimizer of $f(\lambda)$ if either $d_r = 0$ or $g(\nu_r) = 0$.

Otherwise, update the quantities for the next step as given below. If $f(\nu_r) \geqq f(\lambda_r)$, then set $\lambda_{r+1} = \lambda_r$, $\gamma_{r+1} = \nu_r$, $G(\lambda_{r+1}) = G(\lambda_r)$. In this case there is no move in the λ_r-sequence.

If $f(\nu_r) < f(\lambda_r)$, then set $\lambda_{r+1} = \nu_r$, and

$$\text{if } g(\lambda_r)g(\nu_r) > 0, \text{ then set } \gamma_{r+1} = \gamma_r$$
$$G(\lambda_{r+1}) = \frac{g(\nu_r) - g(\lambda_r)}{d_r}$$
$$\text{if } g(\lambda_r)g(\nu_r) < 0, \text{ then set } \gamma_{r+1} = \lambda_r$$
$$G(\lambda_{r+1}) = \frac{g(\nu_r) - g(\gamma_r)}{\nu_r - \gamma_r}.$$

Under rather general assumptions, it has been proved in [10.23] that if this algorithm does not terminate in a finite number of steps, then $f(\lambda_r) \to f(\lambda_*)$ as $r \to \infty$; and that the sequence $\{\lambda_r\}$ itself converges superlinearly to λ_*, a minimizer of $f(\lambda)$.

2 Constrained Line Minimization With Convex Functions

Need For a Constraint in Line Minimization

Let $\theta(x)$ be a real valued function defined on $\mathbf{R}^n$. In algorithms for the unconstrained minimization of $\theta(x)$, we start at a point $\overline{x} \in \mathbf{R}^n$, develop a search direction $y \in \mathbf{R}^n$, $y \neq 0$, which is a descent direction at $\overline{x}$; and then have to solve the line minimization problem of minimizing $f(\lambda) = \theta(\overline{x} + \lambda y)$ over $\lambda \geqq 0$. It has been shown that such algorithms will have desirable convergence properties if the step length $\lambda \geqq 0$ is chosen so as to satisfy

$$f(\lambda) - f(0) \leqq \omega\lambda$$

where ω is a negative number that is a positive fraction of an estimate of the directional derivative of $\theta(x)$ at $\overline{x}$ in the direction y. To satisfy this condition, we define $c(\lambda) = f(\lambda) - f(0) - \omega\lambda$, and solve the constrained line minimization problem

$$\begin{array}{ll} \text{minimize} & f(\lambda) \\ \text{subject to} & c(\lambda) \leqq 0. \end{array}$$

For another application of constrained line minimization, consider the general NLP

$$\begin{array}{ll} \text{minimize} & \theta(x) \\ \text{subject to} & h_i(x) \leqq 0,\ i = 1 \text{ to } m. \end{array}$$

Algorithms for solving these problems usually begin with an initial feasible point $\overline{x}$, find a descent search direction y at $\overline{x}$, and do a line minimization in that direction. Define $c(\lambda) = \max .\{h_i(\overline{x} + \lambda y) : i = 1 \text{ to } m\}$. The problem of finding the best feasible point in this search direction, leads to the constrained line minimization problem: minimize $f(\lambda)$, subject to $c(\lambda) \leqq 0$.

The Constrained Line Minimization Problem

Here we consider the constrained line minimization problem

$$\begin{array}{ll} \text{minimize} & f(\lambda) \\ \text{subject to} & c(\lambda) \leqq 0 \end{array}$$

where both functions $f(\lambda)$ and $c(\lambda)$ are convex. Let

$$\mathbf{S} = \{\lambda : c(\lambda) \leqq 0\}.$$

Since $c(\lambda)$ is convex, $\mathbf{S}$ is an interval, but it may be hard to determine $\mathbf{S}$ explicitly if $c(\lambda)$ is nonlinear. However, we assume that $\mathbf{S}$ has a nonempty interior and that a feasible point (i. e., $\lambda \in \mathbf{S}$) may be found, for example, by finding an unconstrained minimum of $c(\lambda)$.

Here we discuss a modification of the algorithm of Subsection 1 due to R. Mifflin [10.28] for solving this constrained problem.

The method generates two sequences $\{\lambda_r\}$, $\{\gamma_r\}$, where for each r, λ_r is feasible and λ_r, γ_r are on opposite sides of any constrained minimization point λ_*; γ_r is either infeasible (i. e., $c(\gamma_r) > 0$) or $f(\gamma_r) \geqq f(\lambda_r)$. The sequence $\{f(\lambda_r)\}$ is non-increasing with r. In this subsection we define

$$\begin{array}{ll} g(\lambda) \in \partial f(\lambda) & \text{if } c(\lambda) \leqq 0 \text{ (i. e., } \lambda \in \mathbf{S}) \\ g(\lambda) \in \partial\, c(\lambda) & \text{if } c(\lambda) > 0 \text{ (i. e., } \lambda \notin \mathbf{S}). \end{array}$$

We therefore have

$$\begin{aligned} c(\lambda) &\geqq c(\gamma) + g(\gamma)\,(\lambda - \gamma), \text{ for all } \lambda, \text{ and } \gamma \notin \mathbf{S} \\ f(\lambda) &\geqq f(\gamma) + g(\gamma)(\lambda - \gamma), \text{ for all } \lambda, \text{ and } \gamma \in \mathbf{S}. \end{aligned}$$

For $\overline{\lambda}$ feasible, as in Subsection 1, we define $G(\overline{\lambda}) = (g(\overline{\lambda}) - g(t))/(\overline{\lambda} - t)$ where t is feasible and $\overline{\lambda}$ is between t and λ_*. The quadratic approximation step at a feasible point $\overline{\lambda}$ is defined as before, using $G(\overline{\lambda})$.

In Step r of the algorithm, if both the points λ_r and γ_r are feasible, the polyhedral approximation step is defined exactly as under Subsection 1.

Given λ_r, γ_r, if γ_r is infeasible; then $g(\gamma_r)$ is a subgradient of the constraint function $c(\lambda)$, and is not related to the objective function $f(\lambda)$. Thus, in this case, the polyhedral approximation step given λ_r, γ_r is not well defined as in Subsection 1. One aim for this step could be to move the γ-sequence towards feasibility. Taking this step to be $\hat{d}$, where $\lambda_r + \hat{d} = \gamma_r - (c(\gamma_r)/g(\gamma_r))$ would correspond to a Newton-Raphson step for solving $c(\lambda) = 0$ based upon linearization of $c(\lambda)$ at γ_r. On the other hand, in order to make a move not just towards feasibility, but towards a minimizing feasible point, we could take the step to be $\tilde{d}$ where $\lambda_r + \tilde{d}$ is the point at which the linearization of $f(\lambda)$ at λ_r, and the linearization of $c(\lambda)$ at γ_r become equal. This leads to $\tilde{d} = (-c(\gamma_r) - g(\gamma_r)(\lambda_r - \gamma_r))/(g(\gamma_r) - g(\lambda_r))$. In order to achieve fast convergence, the actual polyhedral approximation step in this case, from the feasible point λ_r, is taken to be a compromise between $\hat{d}$ and $\tilde{d}$ given by

$$d_r^P = \frac{P(\lambda_r, \gamma_r)}{g(\gamma_r) - b_r g(\lambda_r)}$$

where $P(\lambda_r, \gamma_r) = -c(\gamma_r) - g(\gamma_r)(\lambda_r - \gamma_r)$ and $b_r = P(\lambda_r, \gamma_r)$. We are now ready to describe the algorithm.

The algorithm needs an initial pair of points λ_1, γ_1 such that λ_1 is feasible (i. e., $\lambda_1 \in \mathbf{S}$), and either $c(\gamma_1) > 0$ or $f(\gamma_1) \geqq f(\lambda_1)$; and $g(\lambda_1)g(\gamma_1) < 0$. This implies that a constrained minimizing point lies between λ_1 and γ_1. Also choose $G(\lambda_1) \geqq 0$. We will now describe the general step in the algorithm.

Step r. Let λ_r, γ_r be the points at the beginning of this step. Define

$$\begin{aligned}
P(\lambda_r, \gamma_r) &= -c(\gamma_r) - g(\gamma_r)(\lambda_r - \gamma_r), \text{ and } b_r = P(\lambda_r, \gamma_r), \text{ if } c(\gamma_r) > 0 \\
P(\lambda_r, \gamma_r) &= \quad f(\lambda_r) - f(\gamma_r) - g(\gamma_r)(\lambda_r - \gamma_r), \text{ and } b_r = 1, \text{ if } c(\gamma_r) \leqq 0 \\
d_r^P &= \frac{P(\lambda_r, \gamma_r)}{g(\gamma_r) - b_r g(\lambda_r)} \\
d_r^Q &= \frac{-g(\lambda_r)}{G(\lambda_r)}, \text{ if } G(\lambda_r) > 0 \\
|d_r^Q| &= +\infty, \text{ if } G(\lambda_r) \leqq 0 \\
d_r &= (\text{sign of } (-g(\lambda_r)))(\min.\{|d_r^P|, |d_r^Q|\}) \\
\nu_r &= \lambda_r + d_r.
\end{aligned}$$

Terminate with the conclusion that ν_r is the optimum solution of the problem if either $d_r = 0$ or $g(\nu_r) = 0$.

Otherwise, update the quantities for the next step as given below. If $c(\nu_r) > 0$ or $f(\nu_r) \geqq f(\lambda_r)$, set $\lambda_{r+1} = \lambda_r$, $\gamma_{r+1} = \nu_r$, $G(\lambda_{r+1}) = G(\lambda_r)$.

If $c(\nu_r) \leqq 0$ and $f(\nu_r) < f(\lambda_r)$, then set $\lambda_{r+1} = \nu_r$, and

$$\text{if } g(\lambda_r)g(\nu_r) > 0, \text{ then set } \gamma_{r+1} = \gamma_r, G(\lambda_{r+1}) = \frac{g(\nu_r) - g(\lambda_r)}{d_r}$$

$$\text{if } g(\lambda_r)g(\nu_r) < 0, \text{ then set } \gamma_{r+1} = \lambda_r, G(\lambda_{r+1}) = \frac{g(\nu_r) - g(\gamma_r)}{\nu_r - \gamma_r}.$$

Under rather general conditions, R. Mifflin [10.28] has proved that if the algorithm does not terminate finitely, then $f(\lambda_r)$ converges to the minimum value of $f(\lambda)$ over $\mathbf{S}$, and that the sequence $\{\lambda_r\}$ itself converges to an optimum solution of the problem, λ_*, with $|\lambda_r - \lambda_*||\gamma_r - \lambda_r|$ converging to zero superlinarly.

3 General Constrained Line Minimization

Let $f(\lambda)$, $c(\lambda)$ be real valued functions defined on $\mathbf{R}^1$, not necessarily convex. Here we consider the constrained line minimization problem

$$\begin{array}{ll} \text{minimize} & f(\lambda) \\ \text{subject to} & c(\lambda) \leqq 0. \end{array}$$

The set $\mathbf{S} = \{\lambda : c(\lambda) \leqq 0\}$ is the feasible set. Since $c(\lambda)$ is not assumed to be convex, $\mathbf{S}$ may consist of a collection of disjoint intervals.

Let $F(\lambda)$ denote either $f(\lambda)$ or $c(\lambda)$. If $F(\lambda)$ is continuoulsy differentiable at λ, we let $\partial F(\lambda)$ be the singleton set $\{\frac{dF(\lambda)}{d\lambda}\}$, as in Appendix 3. If $F(\lambda)$ is not differentiable at λ, $\partial F(\lambda)$ denotes the set of subgradients or generalized gradients, it is the convex hull of all limits of sequences of the form $\{\frac{dF(\lambda_k)}{d\lambda} : \{\lambda_k\} \to \lambda$ and $F(\lambda)$ is differentiable at each $\lambda_k\}$. With this definition $\partial F(\lambda)$ agrees with the subdifferential set when $F(\lambda)$ is convex. Also if $F(\lambda)$ is not given explicitly, but is defined implicitly as the pointwise supremum, say, as $F(\lambda) = \max\{F_1(\lambda), \ldots, F_t(\lambda)\}$ where each $F_i(\lambda)$ is continuously differentiable, then $\partial F(\lambda)$ will be the convex hull of $\{\frac{dF_i(\lambda)}{d\lambda}$: over all i such that $F_i(\lambda) = F(\lambda)\}$. The algorithm discussed in this subsection needs a subroutine which can evaluate $F(\lambda)$ for any λ, and another subroutine to obtain a number $g(\lambda) \in \partial F(\lambda)$.

Stationary Points

A point $\lambda_* \in \mathbf{S}$ is a stationary point for this constrained line minimization problem if

$$\begin{array}{l} \text{either } c(\lambda_*) < 0, \text{ and } 0 \in \partial f(\lambda_*) \\ \quad \text{or } c(\lambda_*) = 0, \text{ and } 0 \in \text{ convex hull of } \partial f(\lambda_*) \cup \partial c(\lambda_*) \end{array}$$

because these are the necessary optimality conditions for this problem. See Figure 10.7.

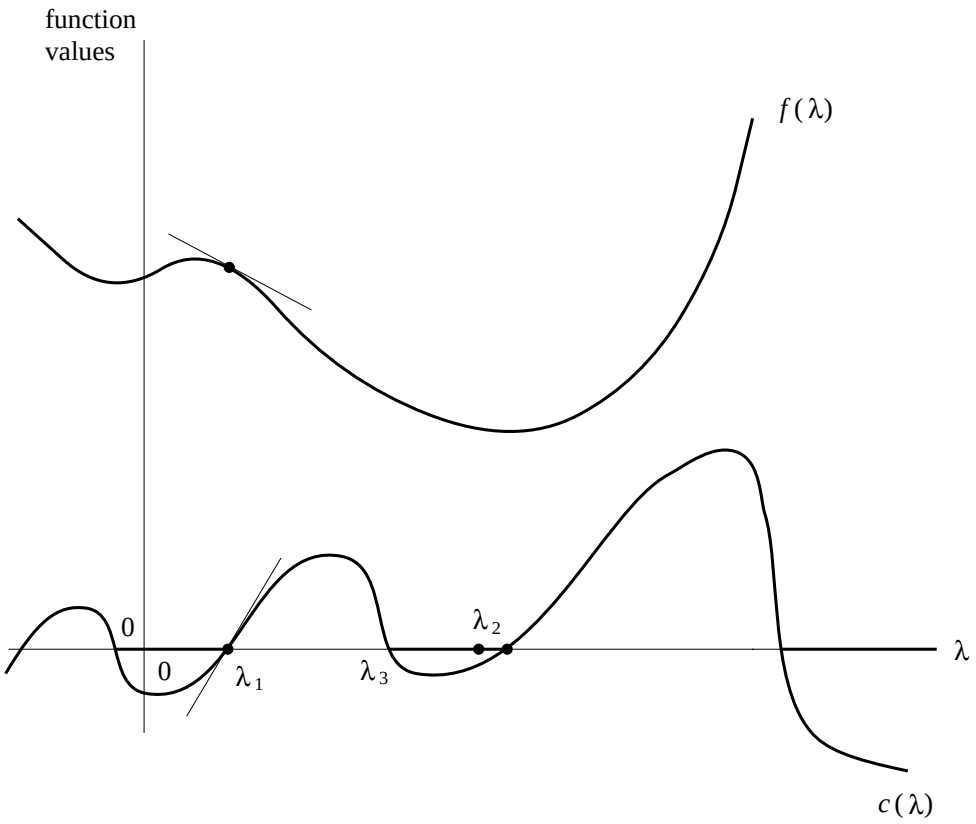


Figure 10.7 The feasible set **S** consists of the thick portion of the λ-axis. λ_1 is a stationary point since $c(\lambda_1) = 0$; $\frac{dc(\lambda_1)}{d\lambda_1} > 0$, $\frac{df(\lambda_1)}{d\lambda} < 0$ and so 0 is in the convex hull of $\frac{dc(\lambda_1)}{d\lambda}$, and $\frac{df(\lambda_1)}{d\lambda}$. λ_2 is another stationary point, but λ_3 is not.

In the algorithm discussed in this subsection, we need

$$g(\lambda) \in \partial f(\lambda) \text{ if } \lambda \in \mathbf{S}$$
$$g(\lambda) \in \partial c(\lambda) \text{ if } \lambda \notin \mathbf{S}.$$

The algorithm generates two sequences of points $\{\lambda_r\}$, $\{\gamma_r\}$ with λ_r feasible for all r and $f(\lambda_r)$ non-increasing. For each r we will have

$$c(\lambda_r) \leqq 0 \text{ and } g(\lambda_r)(\gamma_r - \lambda_r) < 0$$
$$\text{and either } c(\gamma_r) > 0, \text{ or } c(\gamma_r) \leqq 0 \text{ and } f(\gamma_r) \geqq f(\lambda_r).$$

These conditions imply that there exists a stationary point between λ_r and γ_r. The algorithm needs a pair of initial points λ_1, γ_1 satisfying the above conditions, these can be obtained by a suitable initialization routine. The sequence of points $\{\lambda_r\}$ obtained in the algorithm converges to a stationary point λ_* and $|\lambda_r - \lambda_*|$. $|\gamma_r - \lambda_*|$ converges to zero superlinearly.

The Quadratic Approximation Step

As before, $G(\lambda_r)$ is an approximation to the second derivative of $f(\lambda)$ at λ_r, and it is determined in a one-sided secant-manner, that is, when $\lambda_r \neq \lambda_1$,

$$G(\lambda_r) = \frac{g(\lambda_r) - g(t_r)}{\lambda_r - t_r}$$

where t_r is a feasible λ_j or γ_j for some $j < r$ and is on the opposite side of λ_r from γ_r. If $f(\lambda)$ is convex, then we will have $G(\lambda_r) \geqq 0$ for all r. But due to nonconvexity we may get some $G(\lambda_r) \leqq 0$. So, the quadratic approximation step is defined here by

$$d_r^Q = \frac{-g(\lambda_r)}{\max.\{G(\lambda_r), 0\}}$$

with the understanding that $|d_r^Q| = +\infty$ if $G(\lambda_r) \leqq 0$.

The Polyhedral Approximation Step

Consider the case when both λ_r and γ_r are feasible first. In this case, if $g(\lambda_r)$, $g(\gamma_r)$ have opposite signs, we define the polyhedral approximation step by

$$d_r^P = \frac{P(\lambda_r, \gamma_r)}{g(\gamma_r) - g(\lambda_r)}$$

where $P(\lambda_r, \gamma_r) = f(\lambda_r) - f(\gamma_r) - g(\gamma_r)(\lambda_r - \gamma_r)$, as before. If $P(\lambda_r, \gamma_r) \geqq 0$ (which will be the case when $f(\lambda)$ is convex) then $\lambda_r + d_r^P$ will be between λ_r and γ_r. Due to nonconvexity it may happen that $g(\lambda_r)$ and $g(\gamma_r)$ do not have opposite signs and/or $P(\lambda_r, \gamma_r)$ is negative. In this case, the polyhedral approximation step needs to be modified as follows. See Figure 10.8.

Let H_r be a secant estimate of $f''(\lambda)$ near γ_r, that is when $\gamma_r \neq \gamma_1$,

$$H_r = \frac{g(\gamma_r) - g(u_r)}{\gamma_r - u_r}$$

where u_r is a feasible λ_j or γ_j for some $j < r$ on the opposite side of γ_r from λ_r. In this case a quadratic approximation to $f(\lambda)$ around γ_r is

$$q(\lambda) = f(\gamma_r) + g(\gamma_r)(\lambda - \gamma_r) + \frac{1}{2}H_r(\lambda - \gamma_r)^{\mathbf{2}}.$$

A linear approximation for $q(\lambda_r + d)$ based at λ_r is

$$f(\gamma_r) + g(\gamma_r)(\lambda_r - \gamma_r) + \frac{1}{2}H_r(\lambda_r - \gamma_r)^{\mathbf{2}} + [g(\gamma_r) + H_r(\lambda_r - \gamma_r)]d.$$

We can take d_r^P to be the value of d which equalizes this to $f(\lambda_r) + dg(\lambda_r)$. This leads to $d = (f(\lambda_r) - f(\gamma_r) - [g(\gamma_r) + h](\lambda_r - \gamma_r))/(g(\gamma_r) + 2h - g(\lambda_r))$, where $h = \frac{1}{2}H_r(\lambda_r - \gamma_r)$. Since this needs to be carried out only under negative curvature, we define a negative curvature correction

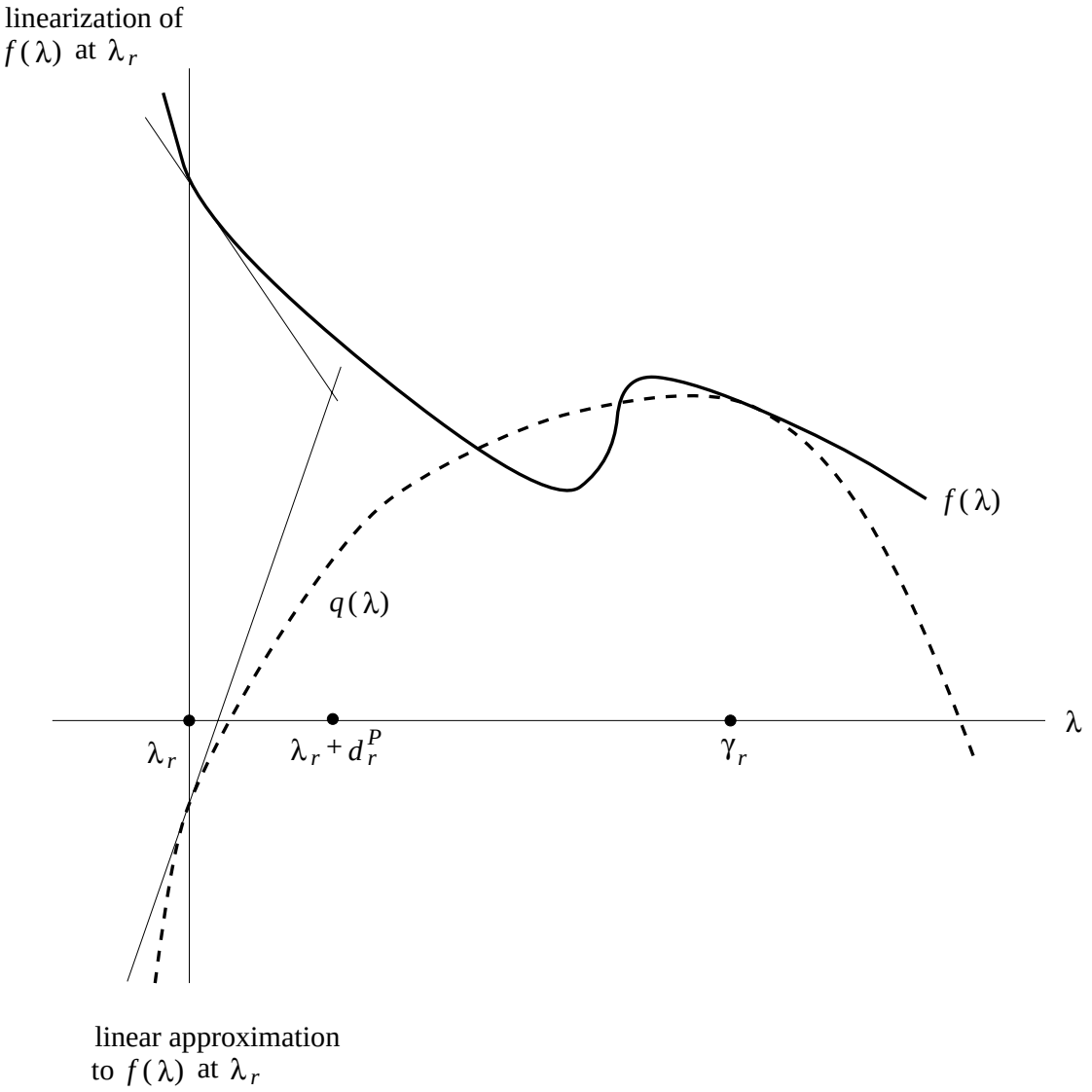


Figure 10.8 $q(\lambda)$ (the dashed curve) is the quadratic approximation to $f(\lambda)$ based at $\lambda = \gamma_r$. The point $\lambda_r + d_r^P$ is the point where the linearizations of $f(\lambda)$ and $q(\lambda)$ based at λ_r, become equal.

$$h_r = \frac{1}{2}(\lambda_r - \gamma_r)\min\{H_r, 0\}$$

and let

$$P_r = f(\lambda_r) - f(\gamma_r) - (g(\gamma_r) + h_r)(\lambda_r - \gamma_r)$$

$$d_r^P \begin{cases} = 0, & \text{if } P_r \leqq 0 \\ = P_r/(g(\gamma_r) + 2h_r - g(\lambda_r)), & \text{if } P_r > 0. \end{cases}$$

Now consider the case when $\gamma_r \notin \mathbf{S}$. In this case we make a similar quadratic approximation to $c(\lambda)$, and using it estimate the point $\lambda_r + d$ where $c(\lambda_r + d)$ would be zero. In this case H_r is an estimate of $c''(\gamma_r)$, and h_r is defined as above. Using again the

compromise as done in Subsection 2 for fast convergence, in this case we are lead to the following polyhedral approximation step.

$$P_r = -c(\gamma_r) - (g(\gamma_r) + h_r)(\lambda_r - \gamma_r)$$

$$d_r^P \begin{cases} = 0, & \text{if } P_r \leqq 0 \\ = P_r/(g(\gamma_r) + 2h_r - P_r g(\lambda_r)), & \text{if } P_r > 0. \end{cases}$$

To handle this general problem, we also define a positive safeguard parameter ζ such that

$$\zeta|\gamma_r - \lambda_r| < \frac{1}{2}$$

so that

$$\zeta(\gamma_r - \lambda_r)^2 < \frac{1}{2}|\gamma_r - \lambda_r| < |\gamma_r - \lambda_r| - \zeta|\gamma_r - \lambda_r|^2.$$

In the algorithm, the step d_r^P is modified into d_r^ζ so that $|d_r^\zeta|$ is between the lower and upper bounds in the above inequality. This guarantees that $\lambda_r + d_r$ is away from λ_r and γ_r. If the problem functions are convex, then $G(\lambda_r) \geqq 0$, $H_r \geqq 0$, $h_r = 0$ and $P_r \geqq 0$, and if $\zeta = 0$, the algorithm discussed below will be the same as the one discussed in Subsection 2. Now we describe the algorithm.

The algorithm needs an initial pair of points λ_1 and γ_1 satisfying the conditions mentioned above. Choose the safeguard parameter $\zeta > 0$ such that $\zeta(\gamma_1 - \lambda_1) < \frac{1}{2}$, and choose the initial curvature estimates $G(\lambda_1)$ and H_1. We will now describe the general step in the algorithm.

Step r. Let λ_r, γ_r be the points at the beginning of this step. Let $G(\lambda_r)$, H_r be the curvature estimates. Set

$$h_r = \frac{1}{2}(\lambda_r - \gamma_r)\min\{H_r, 0\}.$$

$$\left.\begin{aligned} P_r &= -c(\gamma_r) - (g(\gamma_r) + h_r)(\lambda_r - \gamma_r) \\ \text{and } \delta_r &= g(\gamma_r) + 2h_r - P_r g(\lambda_r) \end{aligned}\right\} \quad \text{if } c(\gamma_r) > 0.$$

$$\left.\begin{aligned} P_r &= f(\lambda_r) - f(\gamma_r) - (g(\gamma_r) + h_r)(\lambda_r - \gamma_r) \\ \delta_r &= g(\gamma_r) + 2h_r - g(\lambda_r) \end{aligned}\right\} \quad \text{if } c(\gamma_r) \leqq 0.$$

$$\begin{aligned} d_r^P &= 0 && \text{if } P_r \leqq 0 \\ &= P_r/\delta_r && \text{if } P_r > 0 \\ \zeta_r &= \zeta|\gamma_r - \lambda_r|^2 \\ \alpha_r &= +1 && \text{if } -g(\lambda_r) > 0 \\ &= -1 && \text{if } -g(\lambda_r) < 0. \end{aligned}$$

$$\begin{aligned} d_r^\zeta &= \alpha_r \zeta_r, && \text{if } |d_r^P| < \zeta_r \\ &= \alpha_r |d_r^P|, && \text{if } \zeta_r \leqq |d_r^P| \leqq |\gamma_r - \lambda_r| - \zeta_r \\ &= \alpha_r(|\gamma_r - \lambda_r| - \zeta_r), && \text{if } |d_r^P| > |\gamma_r - \lambda_r| - \zeta_r \end{aligned}$$

$$d_r^Q = -g(\lambda_r)/\max\{G(\lambda_r), 0\}$$
$$d_r = \alpha_r(\min\{|d_r^Q|, |d^\zeta|\})$$
$$\nu_r = \lambda_r + d_r.$$

If $c(\nu_r) \leqq 0$, $f(\nu_r) < f(\lambda_r)$ and $g(\nu_r) = 0$, terminate with the conclusion that ν_r is a stationary point.

Otherwise update the quantities for the next step as given below.

If $c(\nu_r) > 0$, or $f(\nu_r) \geqq f(\lambda_r)$, then set $\lambda_{r+1} = \lambda_r$, $\gamma_{r+1} = \nu_r$, $G(\lambda_{r+1}) = G(\lambda_r)$, and $H_{r+1} = [g(\nu_r) - g(\gamma_r)]/(\nu_r - \gamma_r)$.

If $c(\nu_r) \leqq 0$, $f(\nu_r) < f(\lambda_r)$, $g(\nu_r) \neq 0$, and $g(\lambda_r)g(\nu_r) > 0$, then set $\lambda_{r+1} = \nu_r$, $\gamma_{r+1} = \gamma_r$, $G(\lambda_{r+1}) = (g(\nu_r) - g(\lambda_r))/d_r$, and $H_{r+1} = H_r$.

If $c(\nu_r) \leqq 0$, $f(\nu_r) < f(\lambda_r)$, $g(\nu_r) \neq 0$, and $g(\lambda_r)g(\nu_r) < 0$, then set $\lambda_{r+1} = \nu_r$, $\gamma_{r+1} = \lambda_r$, $G(\lambda_{r+1}) = [g(\nu_r) - g(\gamma_r)]/(\nu_r - \gamma_r)$, and $H_{r+1} = G(\lambda_r)$.

Under rather general conditions on the functions, R. Mifflin [10.29] has proved that if the algorithm does not terminate finitely, then $\{\lambda_r\}$ converges to a stationary point of $f(\lambda)$ on S.

To start the algorithm from a feasible λ_1 when a suitable γ_1 is not known, one can use a safeguarded quadratic step of the form

$$-g(\lambda_j)/\max[G(\lambda_j), a_j],\ j = 1, 2$$

where $\{a_j\}$ is a bounded positive sequence chosen so that it converges to zero if $\{g(\lambda_j)\}$ converges to zero.

10.8 SURVEY OF DESCENT METHODS FOR UNCONSTRAINED MINIMIZATION IN $\mathbf{R}^n$

In this section we consider methods for solving the problem

$$\begin{array}{ll} \text{minimize} & \theta(x) \\ \text{over} & x \in \mathbf{R}^n \end{array} \tag{10.4}$$

where $\theta(x)$ is a real valued continuously differentiable function defined over $\mathbf{R}^n$. The methods discussed in this section make use of the first and sometimes the second order partial derivatives of $\theta(x)$ when they exist, or approximations for these constructed from the information accumulated over the iterations. The methods are iterative, they generate a sequence of points $\{x^0, x^1, x^2, \ldots\} \subset \mathbf{R}^n$ beginning with an initial point x^0, and satisfy the property that $\theta(x^r)$ monotonically decreases as r increases.

In this section $\nabla\theta(x^r)$ denotes the row vector of partial derivatives of $\theta(x)$ at the point x^r (the gradient vector of $\theta(x)$ at x^r). When the second order partial derivatives exist, we denote the $n \times n$ Hessian matrix of $\theta(x)$ at the point x^r by $H(\theta(x^r)) =$

$(\frac{\partial^2 \theta(x^r)}{\partial x_i \partial x_j})$. In the methods discussed in this section, each iteration or step consists of three parts. The $(k+1)$th step begins with the point x^k (x^k is the point obtained at the end of step k if $k > 0$, x^0 is the initial point with which the method is initiated) and consists of the following parts

i) compute the search direction at x^k, denoted by y^k. $y^k \in \mathbf{R}^n$, $y^k \neq 0$,
ii) compute the step length in the search direction, $\alpha_k > 0$,
iii) compute the new point $x^{k+1} = x^k + \alpha_k y^k$ and check whether termination criteria are satisfied. If the termination criteria are satisfied, x^{k+1} is accepted as the solution of (10.4). Otherwise, continue the method by going to the next step.

In order to guarantee that $\theta(x^r)$ decreases monotonically, we require the search directions to be **descent directions**. The point $y \in \mathbf{R}^n$, $y \neq 0$ is said to be a descent direction for $\theta(x)$ at the point x^k if there exists a $\overline{\lambda} > 0$ for which

$$\theta(x^k + \lambda y) < \theta(x^k), \quad \text{for all } 0 < \lambda \leqq \overline{\lambda}. \tag{10.5}$$

Since $\theta(x)$ is differentiable at x^k, (10.5) implies that the limit of $(\theta(x^k + \lambda y) - \theta(x^k))/\lambda$ as λ approaches zero through positive values, is $\leqq 0$, that is $(\nabla\theta(x^k))y \leqq 0$. Conversely, it can be verified that any y satisfying

$$(\nabla\theta(x^k))y < 0 \tag{10.6}$$

is a descent direction at x^k. The condition (10.6) is a sufficient condition for y to be a descent direction at x^k. We define a descent direction for $\theta(x)$ at x^k, to be a $y \in \mathbf{R}^n$, $y \neq 0$, satisfying (10.6). Similarly the point $y \in \mathbf{R}^n$, $y \neq 0$ is said to be a **nonascent direction** for $\theta(x)$ at x^k if,

$$(\nabla\theta(x^k))^T y \leqq 0. \tag{10.7}$$

When $\theta(x)$ is twice continuously differentiable, the point $y \in \mathbf{R}^n$, $y \neq 0$ is said to be a **direction of nonpositive curvature** for $\theta(x)$ at x^k if,

$$y^T H(\theta(x^k))y \leqq 0 \tag{10.8}$$

and a **direction of negative curvature** if,

$$y^T H(\theta(x^k))y < 0. \tag{10.9}$$

10.8.1 How to Determine the Step Length?

Let x^k be the current point and suppose the search direction y^k, which is a descent direction, has been selected. Since x^k, y^k are given points, $\theta(x^k + \lambda y^k)$ is now a function of λ only, and it can be verified that its derivative with respect to λ is $(\nabla\theta(x^k + \lambda y^k))y^k$. The descent step length can be determined to minimize $\theta(x^k + \lambda y^k)$ over $\lambda \geqq 0$. This operation is a line search operation. Step lengths determined to minimize $\theta(x^k + \lambda y^k)$

over $\lambda \geqq 0$ are referred to as **optimal step lengths** and algorithms using them are called **optimal step descent techniques**. Since y^k is a descent direction for $\theta(x)$ at x^k, the optimal step length λ_k is > 0 and

$$\frac{d\theta}{d\lambda}(x^k + \lambda_k y^k) = (\nabla\theta(x^k + \lambda_k y^k))y^k = 0. \tag{10.10}$$

So if optimal step lengths are used, the gradient direction at the termination of a line search step is orthogonal to the descent direction.

In practice, it may not be efficient to use optimal step lengths in every iteration. Algorithms which allow for termination of line searches when conditions for an approximate minimum on the line are satisfied, are said to use **partial** or **inexact line searches**. When using inexact line searches, it is necessary to make sure that the line search achieves a sufficient decrease in objective value, to guarantee convergence. A practical criterion requires that the step length λ be determined to make $|(\nabla\theta(x^k + \lambda y^k))y^k|$ sufficiently small. Stated in terms of the decrease in the magnitude of the derivative of $\theta(x^k + \lambda y^k)$ with respect to λ from that at $\lambda = 0$, another criterion requires that the step length λ be chosen to satisfy

$$|(\nabla\theta(x^k + \lambda y^k))y^k| \leqq \eta|(\nabla\theta(x^k))y^k| \tag{10.11}$$

where η is a parameter satisfying $0 \leqq \eta < 1$. If $\eta = 0$ in (10.11), exact line searches are required, and when η is small, the line search procedure needs to be close to optimal.

Step Length Criterion to Achieve Sufficient Rate of Decrease

A fundamental requirement of step size procedures used in descent methods is that there be a sufficient decrease in the objective value in each step. There are many ways of specifying what a "sufficient-decrease" is. For example, consider the line minimization problem of minimizing $\theta(x^k + \lambda y^k)$ over $\lambda \geqq 0$, where y^k is a descent direction for $\theta(x)$ at x^k. The quantity, $(\nabla\theta(x^k))y^k$, the directional derivative of $\theta(x)$ in the search direction y^k, is a measure of the rate of decrease in $\theta(x)$ at x^k in the direction y^k. Select a number α, $0 < \alpha < 1$, known as the **attenuation factor**. One sufficient decrease criterion requires that over the step length taken, the function value must decrease per unit step, at least a fraction α of the rate of decrease in $\theta(x)$ at x^k in the direction y^k. That is, that the step length λ chosen satisfy

$$\theta(x^k) - \theta(x^k + \lambda y^k) \geqq \lambda\alpha|(\nabla\theta(x^k))y^k|. \tag{10.12}$$

To depict this pictorially, we plot λ on the horizontal axis, and function values on the vertical axis in the two dimensional cartesian plane in Figure 10.9. The curve in Figure 10.9 is $\theta(x^k + \lambda y^k)$ plotted against λ. The straight lines in Figure 10.9 are plots of

$$l_\alpha(\lambda) = \theta(x^k) - \lambda\alpha|(\nabla\theta(x^k))y^k|$$

against λ.

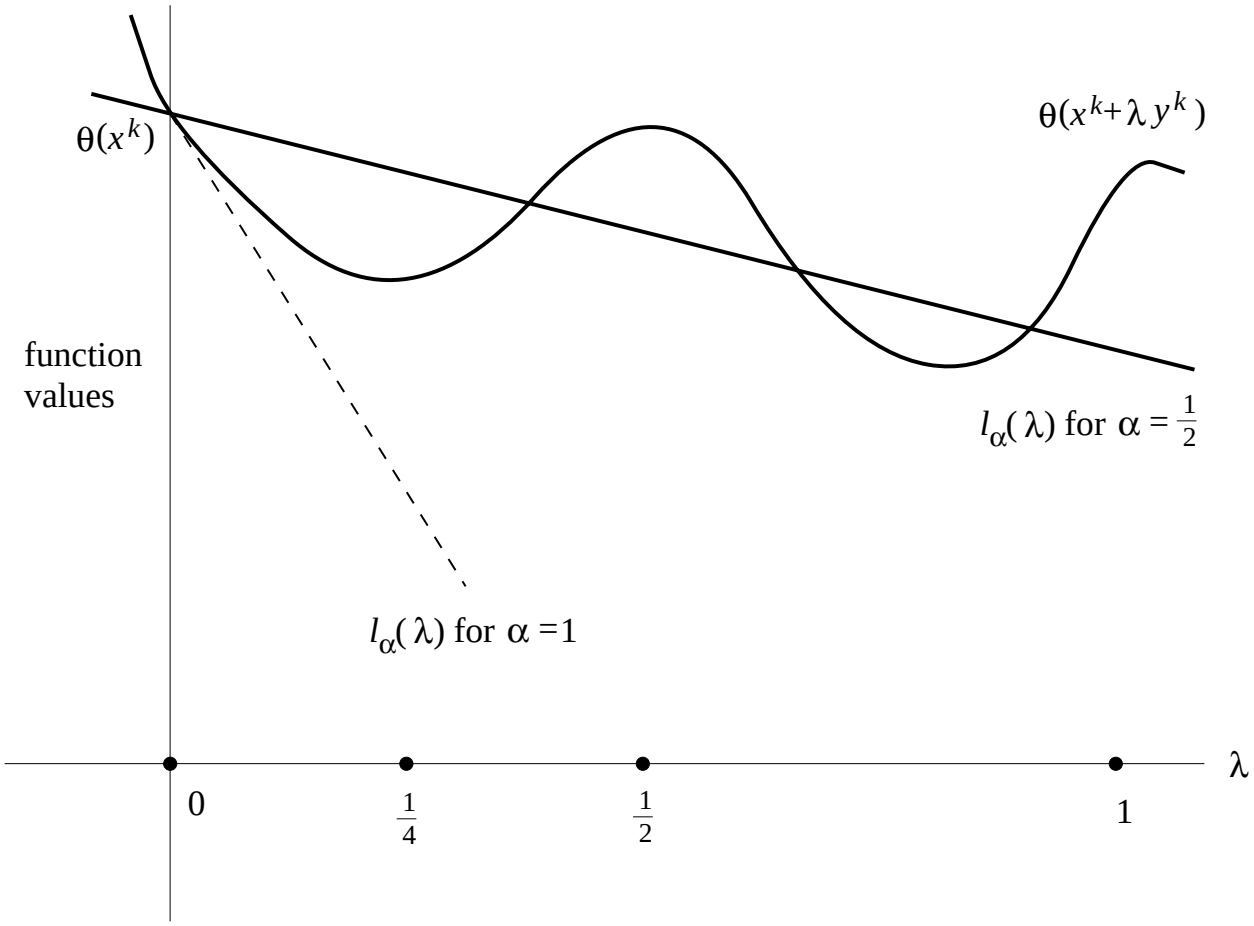


Figure 10.9 The dashed line is $l_\alpha(\lambda)$ for $\alpha = 1$. The continuous straight line is $l_\alpha(\lambda)$ for $\alpha = \frac{1}{2}$. (10.12) requires that for step length λ chosen, $\theta(x^k + \lambda y^k) \leqq l_\alpha(\lambda)$.

The sufficient decrease condition (10.12) states that the step size λ chosen, should satisfy

$$\theta(x^k + \lambda y^k) \leqq l_\alpha(\lambda) = \theta(x^k) - \lambda\alpha|(\nabla\theta(x^k))y^k|.$$

This inequality is called Armijo inequality.

Other Step Length Criteria

Many theoretical convergence proofs for descent algorithms assume that the step length used is the first local minimum along the line in the direction $\lambda \geqq 0$.

The First Order Armijo Step Size Procedure

This procedure was introduced by L. Armijo [10.1]. Let y^k be the descent direction for $\theta(x)$ at the current point x^k. Let $0 < \alpha < 1$ be a predetermined constant. This procedure finds $s =$ smallest nonnegative integer satisfying

$$\theta(x^k) - \theta\left(x^k + \frac{y^k}{2^s}\right) \geqq -\frac{\alpha}{2^s}(\nabla\theta(x^k))y^k \tag{10.13}$$

and chooses the step length to be $1/2^s$. Since y^k is a descent direction, a finite s satisfying (10.13) exists.

As an example, consider the problem depicted in Figure 10.9. Let the attenuation factor $\alpha = \frac{1}{2}$. In Figure 10.9, we verify that $\lambda = 1$ violates the Armijo inequality (10.13), since $\theta(x^k + y^k) > l_\alpha(1)$ for $\alpha = \frac{1}{2}$. Even $\lambda = \frac{1}{2}$ violates the Armijo inequality (10.13) since $\theta(x^k + \frac{1}{2}y^k) > l_\alpha(\frac{1}{2})$ for $\alpha = \frac{1}{2}$. $\lambda = \frac{1}{4}$ satisfies Armijo inequality (10.13) because $\theta(x^k + \frac{1}{4}y^k) < l_\alpha(\frac{1}{4})$ for $\alpha = \frac{1}{2}$ in Figure 10.9. So the step length chosen by this procedure in this problem is $\lambda_k = \frac{1}{4}$.

It can be verified that there is always a positive integer s satisfying (10.13). So, the step length indicated in this procedure is well defined and unique.

One thing that should be noted here is that the step length chosen by this procedure depends on the scaling of y^k. Replacing y^k by βy^k where $\beta > 0$; does not change the search direction, or the line search problem; but it could change the step length chosen by this procedure and the final point obtained in the line search by this procedure. The direction y^k is usually selected by a descent direction selection subroutine, using the values of the function $\theta(x)$ or its gradient vectors or hessian matrices evaluated at previous points, and this procedure takes the output of that subroutine as it is.

Second Order Armijo Step Size Procedure

This procedure is useful when using second order methods like Newton's method discussed below. Let x^k be the current point. Here we will have two directions of nonascent, y^k and h^k. If $\nabla\theta(x^k) \neq 0$, y^k should be a descent direction satisfying (10.6). If $\nabla\theta(x^k) = 0$, then h^k is a direction of negative curvature satisfying (10.9). Let $0 < \alpha < 1$ be a predetermined constant. Let s be the smallest nonnegative integer s satisfying

$$\theta(x^k) - \theta\left(x^k + \frac{y^k}{2^s} + \frac{h^k}{2^{s/2}}\right) \geqq -\frac{\alpha}{2^s}((\nabla\theta(x^k))y^k + \frac{1}{2}(h^k)^T H(\theta(x^k))h^k) \tag{10.14}$$

and take the next point to be $x^{k+1} = x^k + \frac{y^k}{2^s} + \frac{h^k}{2^{s/2}}$. The conditions mentioned above guarantee that a finite s satisfying (10.14) exists. It can be proved that if $\theta(x)$ is twice continuously differentiable, and a descent algorithm using this second order Armijo procedure is carried out, then every limit point $\overline{x}$ of the sequences of points $\{x^r\}$ generated by this method is a point $\overline{x}$ satisfying $\nabla\theta(\overline{x}) = 0$ and $H(\theta(\overline{x}))$ is PSD (the second order necessary optimality condition for $\overline{x}$ to be a local minimum of $\theta(x)$). See [10.1, 10.26, 10.27].

10.8.2 The Various Methods

Now we present various descent methods for (10.4). Since each method has the same structure (each step consisting of three parts (i), (ii), (iii) described under Section 10.8), we will briefly describe how the search direction is chosen in each step, what step size procedures can be used, and a summary of convergence results.

10.8.3 The Method of Steepest Descent

In this method, the search direction in the $(k+1)$th step is chosen to be the steepest descent direction at the current point x^k. The steepest descent direction at x^k is clearly the direction $d \in \mathbf{R}^n$ which minimizes

$$\lim_{\lambda \to 0^+} \frac{\theta(x^k + \lambda d) - \theta(x^k)}{\lambda} = (\nabla\theta(x^k))d$$

subject to $\|d\| = 1$. In $\mathbf{R}^n$, $\|d\|$, the distance between d and 0 can be measured by the general distance function $f(d) = \sqrt{d^T A d}$ where A is a PD symmetric matrix of order n. If $A = I$, $f(d)$ becomes the usual Euclidean distance. The matrix A is known as the metric matrix in the distance function $f(d)$. With respect to this metric matrix A, the steepest descent direction at x^k is therefore the d which minimizes $(\nabla\theta(x^k))d$ subject to $\|d\| = d^T A d = 1$. It can be verified that this direction is given by $-(\nabla\theta(x^k))A^{-1}$ if $\nabla\theta(x^k) \neq 0$.

The steepest descent method, dating back to Cauchy (1847) takes the metric matrix to be I in each step, and thus uses the search direction to be $y^k = -(\nabla\theta(x^k))^T$ when x^k is the current point.

It can be shown that the steepest descent method converges when applied with any of the step length procedures discussed in Section 10.8.1. Every limit point $\overline{x}$ of the sequence $\{x^r\}$ generated satisfies the necessary optimality condition $\nabla\theta(\overline{x}) = 0$. The convergence rate for the algorithm is linear [10.13, 10.17, 10.27, 10.33]. In practice, the method has been observed to be notoriously slow and unreliable due to round-off effects.

10.8.4 Newton's Method

This is a second derivative method that can be used only if $\theta(x)$ is twice continuously differentiable. At the current point x^k, $\theta(x)$ is approximated by the quadratic function $\theta(x^k)+(\nabla\theta(x^k))(x-x^k)+\frac{1}{2}(x-x^k)^T H(\theta(x^k))(x-x^k)$, containing the first three terms of the Taylor series expansion for $\theta(x)$ around x^k. The first order necessary condition for the minimum $y = x - x^k$ of this quadratic approximation is that y satisfies

$$H(\theta(x^k))y = -(\nabla\theta(x^k))^T. \tag{10.15}$$

A direction y that satisfies (10.15) is known as the **Newton direction** for $\theta(x)$ and x^k. Assuming the $H(\theta(x^k))$ is PD (since this matrix is nonsingular, the solution of (10.15) is unique), the unique minimum of the quadratic approximation at x^k is

$$x^{k+1} = x^k - (H(\theta(x^k)))^{-1}(\nabla\theta(x^k))^T \tag{10.16}$$

the iterative scheme given by (10.16) is the traditional Newton's method. When $H(\theta(x^k))$ is PD, $(H(\theta(x^k)))^{-1}(\nabla\theta(x^k))^T$ is the steepest descent direction at x^k using $H(\theta(x^k))$ as the metric matrix, and the formula (10.16) is based on using a constant step length of $+1$ in this direction. When $\theta(x)$ satisfies the property that $H(\theta(x))$ is PD for all x (in this case $\theta(x)$ is strictly convex) Newton's method uses the steepest descent direction with the metric matrix $H(\theta(x^k))$ in the step in which x^k is the current point, and since the metric matrix changes in each step, it is called a **variable metric method** in this case.

As an illustration of convergence proofs we provide below a theorem on the convergence of Newton's method.

Theorem 10.1 *Suppose $\theta(x)$ is twice continuously differentiable. Let $H(x) = (h_{ij}(x)) = H(\theta(x))$. So $h_{ij}(x) = \frac{\partial^2\theta(x)}{\partial x_i \partial x_j}$. Suppose each of the functions $h_{ij}(x)$ satisfies the Lipschitz condition, that is, there exists a positive number α satisfying $|h_{ij}(\xi) - h_{ij}(\eta)| \leqq \alpha\|\xi - \eta\|$ for all $\xi, \eta \in \mathbf{R}^n$. Let $\overline{x}$ be a point satisfying $\nabla\theta(\overline{x}) = 0$, $H(\theta(x))$ is PD. If the initial point x^0 is sufficiently close to $\overline{x}$, the sequence of points $\{x^r : r = 0, 1, \ldots\}$ obtained by Newton's method converges to $\overline{x}$ at a second order rate.*

Proof. By Taylor expansion of $\nabla\theta(x)$ around x^r we have $(\nabla\theta(x^r+\xi))^T = (\nabla\theta(x^r))^T + H(x^r)\xi + f(\xi)$ where $|f(\xi)| \leqq \beta\|\xi\|^{\mathbf{2}}$ for some positive number β, when ξ is sufficiently close to zero. Assuming that x^r is sufficiently close to $\overline{x}$, and substituting $\xi = \overline{x} - x^r$, we get $0 = (\nabla\theta(\overline{x}))^T = (\nabla\theta(x^r))^T + H(x^r)(\overline{x} - x^r) + f(\overline{x} - x^r)$. By the continuity of $H(x)$, and the hypothesis, when x^r is sufficiently close to $\overline{x}$, $H(x^r)$ is also PD and so $(H(x^r))^{-1}$ exists. Multiplying the above equation on both sides by $(H(x^r))^{-1}$ we get (since $x^{r+1} = x^r - (H(x^r))^{-1}(\nabla\theta(x^r))^T$ in Newton's method) $0 = -(x^{r+1} - x^r) + (\overline{x} - x^r) + (H(x^r))^{-1}f(\overline{x} - x^r) = (\overline{x} - x^{r+1}) + (H(x^r))^{-1}f(\overline{x} - x^r)$. Since $H(\overline{x})$ is PD, when x^r is sufficiently close to $\overline{x}$, there exists a constant δ such that $\|H(x^r)^{-1}\| \leqq \delta$. So from the above equation we conclude that

$$\|(\overline{x} - x^{r+1})\| = \| - (H(x^r))^{-1}f(\overline{x} - x^r)\| \leqq \delta\|f(\overline{x} - x^r)\| \leqq \beta\delta\|\overline{x} - x^r\|^{\mathbf{2}}.$$

Using this inequality for $r = 0$, we conclude that there exists an $\varepsilon > 0$ sufficiently small, such that $\|\overline{x} - x^0\| < \varepsilon$ implies $\|\overline{x} - x^1\| < \nu\|\overline{x} - x^0\| < \nu\varepsilon$ where $\nu < 1$. Repeating this argument we conclude that $\|\overline{x} - x^r\| \to 0$ as $r \to \infty$, that is, the sequence $\{x^r\}$ converges to $\overline{x}$. That the convergence is of second order follows from the above inequality.

□

10.8.5 Modified Newton's Methods

When $H(\theta(x^k))$ is PD, $-(H(\theta(x^k)))^{-1}(\nabla\theta(x^k))^T$ is a descent direction for $\theta(x)$ at x^k, but there is no guarantee that $\theta(x^{k+1}) \leqq \theta(x^k)$ when x^{k+1} is determined by (10.16), because the step length is a constant, 1, independent of the data. The sequence can be made into a descent sequence by modifying Newton's method into Newton's method with line search, in which the direction of search is $\pm y^k$ satisfying (10.15), the sign determined (when $H(\theta(x^k))$ is not PD) so as to ensure that the direction is a descent direction, and any of the step length procedures discussed earlier are used for the line search.

The major difficulty with Newton's method arises when $H(\theta(x^k))$ is not PD. If $H(\theta(x^k))$ is singular, (10.15) may not have a solution, and even if it has a solution, when $H(\theta(x^k))$ is not PD, solutions of (10.15) are not necessarily descent directions, and methods based on using them may not converge. In the case when $H(\theta(x^k))$ is not even PSD, it is possible to modify Newton's method by using directions of negative curvature together with step size procedures such as the second order Armijo step.

One modification suggested to guarantee that the search directions are descent directions is to replace $H(\theta(x^k))$ in (10.15) by $\alpha^k Q^k + H(\theta(x^k))$ where Q^k is either I or a positive diagonal matrix and α^k is a positive number to ensure that the resulting matrix is PD, and then solve the modified equation to give the search direction to be used at x^k.

For other modified versions of Newton's method see [10.9, 10.13, 10.26, A8].

One main difficulty in using Newton's method (or modified Newton methods) is that the Hessian matrix has to be evaluated in each step. If subroutines for directly computing each element of the Hessian matrix are not available, they can be approximated by finite differences of the gradient vector. For this, select a positive number α, the finite difference interval. To approximate the Hessian at the point x^k, compute

$$H_{.i} = \frac{1}{\alpha}(\nabla\theta(x^k + \alpha I_{.i}) - \nabla\theta(x^k))^T.$$

Let H be the matrix with columns $H_{.i}$, $i = 1$ to n. Then $(H + H^T)/2$ can be used as an approximation for $H(\theta(x^k))$ in executing Newton's or the appropriate modified Newton's method. With this change, the method is usually called a **discrete** (or **finite difference**) **Newton** or **modified Newton method**. These methods are very worthwhile when the Hessian matrix has a known sparsity pattern.

10.8.6 Quasi Newton Methods

Newton's method is difficult to implement because of the computational burden involved in calculating the Hessian matrix in each step (even if we decide to use a finite difference approximation for it). The Quasi-Newton methods try to build up information on the Hessian through various steps of the descent method using the computed

values of $\nabla\theta(x)$ and $\theta(x)$. In these methods $(H(\theta(x^k)))^{-1}$ is approximated by a symmetric positive definite matrix, D_k, which is updated in each iteration. Thus in these methods, the $(k+1)$th step consists of the following.

(a) Initiate this step with the point x^k obtained in the previous step (if $k = 0$, initiate this step with x^0, some initial point with which the method is started).
(b) Compute the search direction at x^k, denoted by $y^k = -D_k(\nabla\theta(x^k))^T$.
(c) Compute step length in the search direction, $\alpha_k > 0$, by doing a line search, leading to the new point $x^{k+1} = x^k + \alpha_k y^k$.
(d) Check whether termination criteria (see Section 10.8.8) are satisfied by the new point x^{k+1}, in which case accept x^{k+1} as the solution of (10.4) and terminate. Otherwise update D_k giving D_{k+1} and go to the next step.

The methods start out with an initial solution x^0, and a symmetric positive definite matrix D_0 (usually $D_0 = I$). D_k is an approximation to the inverse Hessian at a local minimum to which the sequence of points generated is presumed to converge. Different algorithms use different formula for updating D_k from step to step. The advantages are that these methods only need the computation of the gradient vector $\nabla\theta(x)$ at one point in each step. When the matrices D_k are all PD, the search directions $y_k = -D_k(\nabla\theta(x^k))^T$ are descent directions. In some quasi-Newton methods D_k may not always be PD, but the important methods do maintain this property. When D_k is PD, the search direction y^k is the steepest descent direction at x^k using D_k^{-1} as the metric matrix, and since this metric matrix changes from iteration to iteration, these methods are also known as **variable metric methods**.

The updating formula which gives D_{k+1} as a function of D_k attempts to take into account the second derivative information obtained during the $(k+1)$th step. The formula is derived to ensure that D_k becomes a good approximation of $(H(\theta(x^k)))^{-1}$ as the method progresses. This is done through the use of an equation known as the **quasi-Newton condition**, which we will now derive. By taking the Taylor series expansion of $\nabla\theta(x)$ around the point x^k and neglecting higher order terms, we get

$$(\nabla\theta(x^{k+1}))^T \simeq (\nabla\theta(x^k))^T + H(\theta(x^k))(x^{k+1} - x^k).$$

So, if $H(\theta(x^k))$ is invertible, we have

$$(H(\theta(x^k)))^{-1}(\nabla\theta(x^{k+1}) - \nabla\theta(x^k))^T \simeq (x^{k+1} - x^k). \tag{10.17}$$

Since the quantities x^{k+1} and $\nabla\theta(x^{k+1})$ are not available until the $k+1$th step is completed, we cannot expect the matrix D_k to satisfy (10.17) in place of $(H(\theta(x^k)))^{-1}$, but we could require D_{k+1} to satisfy

$$D_{k+1}(\nabla\theta(x^{k+1}) - \nabla\theta(x^k))^T = (x^{k+1} - x^k). \tag{10.18}$$

This condition is the quasi-Newton condition, and the updating formulae for the matrices D_k in quasi-Newton methods are usually formulated so that this condition holds for all k. If the updating formulae are such that this condition is satisfied, and satisfies certain other prior conditions (sometimes it is also required that

$D_{k+1}(\nabla\theta(x^{j+1}) - \nabla\theta(x^j))^T = (x^{j+1} - x^j)$ hold for all $j \leqq k$) it can be shown that when the algorithm is applied to minimize $\frac{1}{2}x^TAx + cx$ where A is PD and symmetric, using exact line searches, then the search directions generated are conjugate directions (see Section 10.8.7 for the definition of conjugate directions), that $D_n = A^{-1}$, and that the method terminates after at most n steps with the minimum.

The three basic considerations in constructing updating formulae for D_k in quasi-Newton methods are (i) the quasi-Newton condition (10.18), (ii) hereditary symmetry (i. e., if D_k is symmetric, the updating formula should guarantee that D_{k+1} is also symmetric), and (iii) hereditary positive definiteness. Not all the quasi-Newton methods satisfy all these properties. In some of them, these properties may only hold if the line searches are carried out to a high degree of precision in each iteration.

The updating formula usually has the form $D_{k+1} = D_k + C_k$ where C_k is a matrix known as the correction term. Usually C_k has rank 1 or 2, and depending on its rank, the methods are classified either as rank-one or rank-two methods.

Now we will present the updating formulas used by some important quasi-Newton methods. The remaining details are the same as discussed above, for each method. For $k \geqq 1$, we define

$$\begin{aligned} \xi^k &= x^k - x^{k-1} \\ \eta^k &= (\nabla\theta(x^k) - \nabla\theta(x^{k-1}))^T \end{aligned} \tag{10.19}$$

The Davidon-Fletcher-Powell (DFP) Method

Here the updating formula is

$$D_{k+1} = D_k + \frac{\xi^{k+1}(\xi^{k+1})^T}{(\xi^{k+1})^T\eta^{k+1}} - \frac{(D_k\eta^{k+1})(D_k\eta^{k+1})^T}{(\eta^{k+1})^TD_k\eta^{k+1}}$$

where ξ^{k+1}, η^{k+1} are column vectors defined as in (10.19). The method has the hereditary symmetry property. It also has the hereditary PD property if $(\xi^{k+1})^T\eta^{k+1} > 0$ for all k. Notice that this condition will hold if the search direction y^k is a descent direction and the line search is carried out optimally or to a local minimum. The method has superlinear rate of convergence. When applied to minimize a strictly convex quadratic function $\frac{1}{2}x^TAx + cx$ with exact line searches, the method preserves the condition $D_{k+1}\eta^{j+1} = \xi^{j+1}$ for all $j \leqq k$, for all k; it generates conjugate search directions and terminates after n steps with $D_{n+1} = A^{-1}$ and the optimum solution. See [10.9, 10.13, 10.37, A3] for proofs of these results.

The Broyden-Fletcher-Goldfarb-Shanno (BFGS) Method

Here the updating formula is

$$\begin{aligned} D_{k+1} =& D_k + \left(1 + \frac{(\eta^{k+1})^TD_k\eta^{k+1}}{(\xi^{k+1})^T\eta^{k+1}}\right)\left(\frac{\xi^{k+1}(\xi^{k+1})^T}{(\xi^{k+1})^T\eta^{k+1}}\right) - \\ & \frac{\xi^{k+1}(\eta^{k+1})^TD_k + D_k\eta^{k+1}(\xi^{k+1})^T}{(\xi^{k+1})^T\eta^{k+1}} \end{aligned}$$

where ξ^{k+1}, η^{k+1} are column vectors defined as in (10.19). The method has the hereditary symmetry and hereditary PD properties, satisfies the quasi-Newton conditions ($D_{k+1}\eta^{j+1} = \xi^{j+1}$ for all $j \leqq k$), and has the quadratic termination property. At present this is considered the best quasi-Newton method. The method has been shown to converge even with inexact line searches (using several of the line search termination criteria discussed in Section 10.7.8).

Resetting in Quasi-Newton Methods

In quasi-Newton methods, the steps can continue until termination. However, in some implementations the method is reset by setting the matrix D_k to some positive definite matrix (usually the same as D_0, or I) after every n steps. If implemented this way, the method goes through cycles. Each cycle begins with the point obtained at the end of the last step in the previous cycle (the initial cycle begins with the initial point x^0 with which the method is initiated) and the initial step of each cycle begins with the matrix D_0 (usually I or some other PD symmetric matrix) and the cycle consists of n steps.

Also in each step one should check that the search direction y^k satisfies $(\nabla\theta(x^k))y^k < 0$, as otherwise the direction is not a descent direction. Usually the method is also reset whenever this descent condition is violated.

See references [10.2, 10.8, 10.9, 10.13, 10.26, 10.27, 10.33, A3] for a discussion of various other quasi-Newton methods, their best computer implementations, and the convergence results established about them.

10.8.7 Conjugate Direction Methods

These are a class of methods that use only first order derivatives, which obtain search directions without the need for storing or updating a square matrix of order n. Conjugate direction methods were developed with the aim of solving strictly convex quadratic programming problems with an effort of at most n line searches. For this, the search directions have to be chosen to satisfy the conjugancy property. Let $f(x) = cx + \frac{1}{2}x^T Ax$ where A is a PD symmetric matrix of order n. Consider the linear transformation $x = Pz$ where P is a nonsingular square matrix of order n. This transforms $f(x)$ into $F(z) = cPz + \frac{1}{2}z^T P^T APz$. $F(z)$ can be minimized with an effort of at most n line searches in the z-space if it is separable, that is, if $P^T AP = Q$ is a diagonal matrix with positive diagonal entries $Q_{11}, \ldots, Q_{nn}$, and

$$(P_{.i})^T AP_{.j} = 0 \quad \text{for each } i \neq j. \tag{10.20}$$

In this case $F(z)$ is equal to $\sum_{j=1}^{n} F_j(z_j)$ where $F_j(z_j)$ involves only one variable, and hence minimizing $F(z)$ over $z \in \mathbf{R}^n$ can be achieved by n one dimensional problems of minimizing $F_j(z_j)$ over $z_j \in \mathbf{R}^1$ for each $j = 1$ to n separately, that is, n line

searches. The set of nonzero vectors $\{P_{.1}, \ldots, P_{.n}\}$ is said to be **conjugate with respect to the PD symmetric matrix** A if (10.20) holds. Let $\nu = (\nu_1, \ldots, \nu_n) = cP$. So $F(z) = \sum_{j=1}^{n} \nu_j z_j + \frac{1}{2} \sum_{j=1}^{n} Q_{jj} z_j^2$. Hence, the point which minimizes $F(z)$ is $\overline{z} = (\overline{z}_j) = (-\nu_j / Q_{jj})$ and so the point which minimizes $f(x)$ in the x-space is $\overline{x} = (\overline{x}_j) = P\overline{z}$. Since $F(z)$ is separable, we can visualize the minimum of $F(z)$ as being obtained by starting at an arbitrary point z_0 in the z-space and doing n line searches exactly, once in each direction $I_{.j}$, $j = 1$ to n (the alternating variables method). Let z^j be the point obtained at the end of the jth line search in this scheme. So z^{j+1} is the minimizer of $F(z^j + \alpha I_{.j+1})$ over $\alpha \in \mathbf{R}^1$, $j = 0$ to $n-1$. Then $z^n = \overline{z}$. If $x^j = Pz^j$, $j = 0$ to n, it can be verified that x^{j+1} is the minimizer of $f(x^j + \alpha P_{.j+1})$ over $\alpha \in \mathbf{R}^1$, $j = 0$ to $n-1$ and that $x^n = \overline{x}$, the point which minimizes $f(x)$. The following properties can be verified to hold

1. the conjugacy condition (10.20) implies that $\{P_{.1}, \ldots, P_{.n}\}$ is linearly independent.
2. $(\nabla f(x^{k+1}))P_{.j} = 0$, for $j = 1$ to k.
3. Let α_j be the minimizer of $f(x^j + \alpha P_{.j+1})$ over $\alpha \in \mathbf{R}^1$, for $j = 0$ to $n-1$. Then $x^{j+1} = x^j + \alpha_j P_{.j+1}$. So $(\nabla f(x^{j+1}) - \nabla f(x^j))^T = A(x^{j+1} - x^j) = \alpha_j A P_{.j+1}$. So $(\nabla f(x^{i+1}) - \nabla f(x^i))P_{.j} = 0$ for $i \neq j$.

The conjugate gradient methods for minimizing $f(x)$ construct the conjugate directions one after the other using information collected from earlier line searches. Each direction will be a descent direction at the point which is the current point in the step in which this direction is generated. We now describe these methods.

Step 1 is initiated with an arbitrary initial point x^0. The search direction in step 1 is the steepest descent one, $y^0 = -(\nabla f(x^0))^T$. Do a line search to minimize $f(x^0 + \alpha y^0)$, $\alpha \geqq 0$.

The general $(k+1)$th step for $k \geqq 1$ begins with the point x^k obtained at the end of the line search in the kth step. The search direction in this step is

$$y^k = -(\nabla f(x^k))^T + \beta_k y^{k-1}$$

where β_k is a scalar. The various conjugate gradient algorithms use different formula for β_k. They are

$$\beta_k = \|\nabla f(x^k)\|^2 / \|\nabla f(x^{k-1})\|^2 \tag{10.21}$$

in Fletcher and Reeves method [10.13].

$$= (\nabla f(x^k) - \nabla f(x^{k-1}))(\nabla f(x^k))^T / \|\nabla f(x^{k-1})\|^2 \tag{10.22}$$

in Polak and Ribiere and Polyak's method [10.13, 10.17, 10.37].

$$= -\|\nabla f(x^k)\|^2 / (\nabla f(x^{k-1}))y^k \tag{10.23}$$

in conjugate descent method [10.13].

It can be verified that $(\nabla f(x^k))y^k = -\|\nabla f(x^k)\|^2$ if the line search in the previous step is carried out exactly, and in this case y^k is therefore a descent direction at x^k. Now do a line search to minimize $f(x^k + \alpha y^k)$, $\alpha \geqq 0$. If α_k is the optimum step length,

$x^{k+1} = x^k + \alpha_k y^k$. If $\nabla f(x^{k+1}) = 0$, x^{k+1} minimizes $f(x)$, terminate. Otherwise, go to the next step.

The method terminates after at most n steps. It can be verified that the search directions generated are conjugate with respect to the Hessian matrix A, and they are all descent directions if the line search is carried out exactly in each step. Since $f(x)$ is quadratic, it can be verified that β_k obtained in (10.21) or (10.22) or (10.23) are exactly the same if all the line searches are carried out exactly.

To solve the problem of minimizing $\theta(x)$, which is in general not quadratic, we apply the method exactly as above, replacing $f(x)$ by $\theta(x)$ wherever it appears. In this general problem, the search directions generated will be descent directions as long as line searches are carried out exactly in each step. In this general problem, the values for β_k obtained from (10.21), (10.22), (10.23) may be different. In numerical experiments the method using (10.22) seemed to perform better, particularly when n is large. The application of the method can be continued until some termination condition is satisfied (see Section 10.8.8). In practical implementations to minimize general non-quadratic functions $\theta(x)$, the method is usually restarted (or reset) after every n steps. If this is done, the method goes through several cycles. Each cycle consists of n steps. Step 1 of each cycle begins with the point obtained at the end of the previous cycle (or x^0, the initial point, for the first cycle) and uses the negative gradient search direction. In the general non-quadratic case, if inexact line searches are used, the directions generated, y^k, may not be descent directions (that is, $(\nabla\theta(x^k))^T y^k$ may not be < 0). The method based on updating using (10.23) (the conjugate descent method) produces descent directions even when line searches are not very exact. If the search direction in a step is not descent, we can carry out the line search in that step over the entire line (instead of the half-line with step length $\alpha \geqq 0$ as is done usually, that is, allow step length to be negative), but usually the cycle is terminated in such a step and the method is reset to begin the next cycle with the steepest descent direction in step 1. It can be shown that these methods have superlinear convergence in terms of cycles. See [10.8, 10.13, 10.17, 10.26, 10.37].

10.8.8 Practical Termination Conditions for Unconstrained Minimization Algorithms

When the descent algorithm generates the sequence of points $\{x^r : r = 0, 1, \ldots\}$ in practical implementations for minimizing $\theta(x)$, the method can be terminated when some or all of the following conditions are met

$$|\theta(x^k) - \theta(x^{k-1})| < \varepsilon_1$$
$$\|x^k - x^{k-1}\| < \varepsilon_2$$
$$\|\nabla\theta(x^k)\| < \varepsilon_3$$

where the ε's are suitably chosen tolerances.

10.9 SURVEY OF SOME METHODS FOR LINEAR EQUALITY CONSTRAINED MINIMIZATION IN $\mathbf{R}^n$

Here we consider the NLP

$$\begin{array}{ll} \text{minimize} & \theta(x) \\ \text{subject to} & Ax = b \end{array} \tag{10.24}$$

where A is a matrix of order $m \times n$ and rank m, and $\theta(x)$ is a real valued continuously differentiable function. Given a feasible point x for this problem, the first order necessary conditions for it to be a local minimum are that there exist a Lagrange multiplier vector $\pi = (\pi_1, \ldots, \pi_m)$ satisfying

$$(\nabla\theta(x)) = \pi A. \tag{10.25}$$

Suppose (10.24) is feasible, and let x be any feasible solution for it. Then every feasible solution for (10.24) is of the form $\tilde{x} + z$ where z satisfies

$$Az = 0. \tag{10.26}$$

There exists a matrix Z of order $n \times (n-m)$ and rank $n-m$, such that every column vector of Z is a solution of (10.26) and conversely every solution of (10.26) is a linear combination of the column vectors of Z. To obtain a matrix like Z, find a basis B for (10.24). B is a square nonsingular submatrix of A of order m. Rearrange the variables and their columns in A so that A can be partitioned into basic and nonbasic parts as (B, D) where D is the $m \times (n-m)$ matrix of nonbasic columns. Then the matrix Z can be taken to be

$$Z = \begin{pmatrix} -B^{-1}D \\ I_{n-m} \end{pmatrix} \tag{10.27}$$

where I_{n-m} is the unit matrix of order $n-m$. It is not necessary to compute Z explicitly. All the computations in the algorithms discussed below can be carried out using a factorization for B^{-1}.

Since any solution for (10.24) is of the form $x = \tilde{x} + Z\xi$ where $\tilde{x}$ is a solution of (10.24) and $\xi \in \mathbf{R}^{n-m}$, (10.24) is equivalent to the problem of minimizing $f(\xi) = \theta(\tilde{x} + Z\xi)$ over $\xi \in \mathbf{R}^{n-m}$, that is the unconstrained minimum of $f(\xi)$ over $\xi \in \mathbf{R}^{n-m}$. It can be verified that $\nabla f(\xi) = (\nabla_x \theta(\tilde{x} + Z\xi))Z$. Also if $\theta(x)$ is twice continuously differentiable, $H(f(\xi)) = Z^T H_x(\theta(\tilde{x} + Z\xi))Z$. For x feasible to (10.24) the vector $(\nabla\theta(x))Z$ is known as the **projected gradient** or the **reduced gradient vector of** $\theta(x)$ at x, and the matrix $Z^T H(\theta(x))Z$ of order $(n-m) \times (n-m)$ is known as the **reduced** or **projected Hessian matrix of** $\theta(x)$ at x. The condition (10.25) implies

$$(\nabla\theta(x))Z = 0. \tag{10.28}$$

If $\theta(x)$ is twice continuously differentiable, a second order necessary condition for the feasible solution x of (10.24) to be a local minimum for it is that the matrix $Z^T H(\theta(x))Z$ is PSD.

The algorithms discussed in this section generate a sequence of feasible points $\{x^0, x^1, \ldots\}$ beginning with the initial feasible point x^0. If x^k is feasible, the search direction at x^k in step $k+1$ must satisfy $Ay^k = 0$, that is, $y^k = Z\xi^k$ for some $\xi^k \in \mathbf{R}^{n-m}$, such directions are called feasible search directions, because a move of any length in such a direction, starting from a feasible point, remains in the feasible region for (10.24). Step $k+1$ of the algorithm consists of the following tasks:

1. Compute a feasible search direction: First compute ξ^k and then compute the search direction $y^k = Z\xi^k$.
2. Determine step length: Compute the positive step length α_k.
3. Compute the new point $x^{k+1} = x^k + \alpha_k y^k$.
4. Check whether x^{k+1} satisfies the conditions for termination, if so, accept x^{k+1} as the solution of (10.24) and terminate. Otherwise go to the next step.

The feasible search direction y^k selected in 1. above is a descent direction at x^k if

$$((\nabla\theta(x^k))Z)\xi^k = (\nabla\theta(x^k))y^k < 0. \tag{10.29}$$

The method of steepest descent uses $(\xi^k)^T = -(\nabla\theta(x^k))Z$ to determine the feasible search direction at x^k, which is therefore $y^k = -ZZ^T(\nabla\theta(x^k))^T$, and uses step length procedures exactly as in the unconstrained case. However, this method has slow linear rate of convergence.

Newton's method is based on minimizing the second order Taylor approximation for $f(\xi) = \theta(x^k + Z\xi)$ around $\xi = 0$, that is $\theta(x^k) + (\nabla\theta(x^k))Z\xi + \frac{1}{2}\xi^T Z^T H(\theta(x^k))Z\xi$. So, Newton's method uses the search direction $y^k = Z\xi^k$, where ξ^k solves

$$(Z^T H(\theta(x^k))Z)\xi = -Z^T(\nabla\theta(x^k))^T \tag{10.30}$$

and uses fixed step lengths of $\alpha_k = 1$. Modified Newton methods replace the matrix $Z^T H(\theta(x^k))Z$ in (10.30) (when this matrix is not PD) by a PD approximation to it such as $Z^T H(\theta(x^k))Z + \nu I$ for some $\nu > 0$, and step lengths determined by line searches.

When the second derivatives are not available, the matrix $Z^T H(\theta(x^k))Z$ can be approximated by finite difference approximation. For this, let ε_i be an appropriate finite difference interval, and for $i = 1$ to $n-m$ let

$$W_{.i} = \frac{1}{\varepsilon_i}(\nabla\theta(x^k + \varepsilon_i Z_{.i}) - \nabla\theta(x^k))^T$$

and let W be the $n \times (n-m)$ matrix with column vectors $W_{.i}$, $i = 1$ to $n-m$. Then a symmetric approximation for $Z^T H(\theta(x^k))Z$ is $(1/2)(Z^T W + W^T Z)$.

Quasi-Newton methods can be developed for (10.24) by looking at the corresponding unconstrained minimization problem of minimizing $f(\xi) = \theta(x^k + Z\xi)$, but

carrying out all the operations in the x-space. In this case the search direction in step $k+1$ will be $y^k = Z\xi^k$, where $\xi^k = -D_k Z^T(\nabla\theta(x^k))^T$. The matrix D_k is of order $(n-m)\times(n-m)$. We choose $D_0 = I_{n-m}$, and in updating D_k from step to step, we use the updating formulas discussed in Section 10.8.6 with $\xi^k = Z^T(x^{k+1} - x^k)$, and $\eta^k = Z^T(\nabla\theta(x^{k+1}) - \nabla\theta(x^k))^T$ instead of (10.19).

Another approach for solving (10.24) is to use a conjugate gradient method on the corresponding reduced problem of minimizing $f(\xi) = \theta(x^k + Z\xi)$, but doing all the computations in the x-space. The search directions used are

$$y^1 = -Z(\nabla\theta(x^0)Z)^T$$
$$y^k = -Z(\nabla\theta(x^k)Z)^T + \beta_k y^{k-1}$$

where $\beta_k = \|(\nabla\theta(x^k))Z\|^2/\|(\nabla\theta(x^{k-1}))Z\|^2$ or $(\nabla\theta(x^k) - \nabla\theta(x^{k-1}))ZZ^T(\nabla\theta(x^k))^T/\|\nabla\theta(x^{k-1})Z\|^2$, or $-\|\nabla\theta(x^k)Z\|^2/(\nabla\theta(x^{k-1})Z)\xi^k$ (here ξ^k is the unique solution of $Z\xi^k = y^k$), as in (10.21), (10.22), (10.23), depending on the method used. Statements made in Section 10.8.7 about resetting the algorithm remain valid here also (here resetting is done after every $n-m$ steps or whenever the search direction generated is not a descent direction).

10.9.1 Computing the Lagrange Multiplier Vector

Let $\overline{x}$ be the terminal point obtained in the algorithm for solving (10.24). The corresponding Lagrange multiplier vector is the vector $\overline{\pi}$ which satisfies (10.25). Given $\overline{x}$, (10.25) is a system of n equations in the m unknowns $\pi_1, \ldots, \pi_m$, and since $n > m$, this is an overdetermined system of equations. We can determine $\overline{\pi}$ as the row vector in $\mathbf{R}^m$ which minimizes $\|(\nabla\theta(\overline{x}))^T - \pi A\|^2$ over $\pi \in \mathbf{R}^m$, for which the solution is given by

$$\overline{\pi} = (AA^T)^{-1}A\nabla\theta(\overline{x}). \tag{10.31}$$

If $\overline{x}$ is a local minimum for (10.24), the vector $\overline{\pi}$ given by (10.31) is an exact solution for (10.25). If $\overline{x}$ is an approximation to a local minimum (obtained when the algorithms discussed above are terminated using some practical termination criteria discussed in Section 10.8, there is no π satisfying (10.25) exactly, however, the $\overline{\pi}$ obtained from (10.31) is a corresponding approximation to the Lagrange multiplier vector for (10.24). For other approximating estimates to the Lagrange multiplier vector see references [10.13, 10.17].

10.10 SURVEY OF OPTIMIZATION SUBJECT TO GENERAL LINEAR CONSTRAINTS

10.10.1 The Use of Lagrange Multipliers to Identify Active Inequality Constraints

For the purpose of this discussion, consider the following NLP:

$$\begin{array}{ll} \text{minimize} & \theta(x) \\ \text{subject to} & Ax \geqq b \end{array} \tag{10.32}$$

where A is a matrix of order $m \times n$, say. If $\overline{x}$ is feasible, the ith constraint in (10.32) is said to be active or tight or binding at $\overline{x}$ if it holds as an equation at $\overline{x}$, that is, if $A_{i.}\overline{x} = b_i$; inactive if $A_{i.}\overline{x} > b_i$. For $\overline{x}$ feasible to (10.32), let $\mathbf{I}(\overline{x}) = \{i : i \text{ such that } A_{i.}\overline{x} = b_i\}$ = index set of active constraints in (10.32) at $\overline{x}$. Let $y \in \mathbf{R}^n$, $y \neq 0$. y is said to be a feasible direction at $\overline{x}$, if $\overline{x} + \lambda y$ remains feasible for (10.32) for all $0 \leqq \lambda \leqq \overline{\lambda}$, for some positive $\overline{\lambda}$. Clearly y is a feasible direction at $\overline{x}$ iff

$$A_{i.}y \geqq 0, \quad \text{for each } i \in \mathbf{I}(\overline{x}). \tag{10.33}$$

The direction y is said to be a **binding direction** or a **non-binding direction** at $\overline{x}$ with respect to the ith constraint for $i \in \mathbf{I}(\overline{x})$, depending on whether $A_{i.}y = 0$ or $A_{i.}y > 0$ respectively. A move in a binding direction continues to keep the constraint active, while any move of positive length in a non-binding direction makes the constraint inactive, that is, moves off the constraint.

Now consider the corresponding equality constrained NLP:

$$\begin{array}{ll} \text{minimize} & \theta(x) \\ \text{subject to} & Ax = b \end{array} \tag{10.34}$$

and further assume that the set of row vectors of A is linearly independent. Suppose $\overline{x}$ is a KKT point for (10.34) with the associated Lagrange multiplier vector $\overline{\pi} = (\overline{\pi}_1, \ldots, \overline{\pi}_m)$. So $\overline{x}$, $\overline{\pi}$ together satisfy the first order necessary optimality conditions

$$\nabla\theta(\overline{x}) = \overline{\pi}A. \tag{10.35}$$

Since the set of feasible solutions of (10.34) is a subset of the set of feasible solutions of (10.32), an optimum solution for (10.34) may not be optimal for (10.32) in general. The point $\overline{x}$ is of course feasible to (10.32) and clearly it is also a KKT point for (10.32) if $\overline{\pi} \geqq 0$.

Suppose there is a t such that $\overline{\pi}_t < 0$, we will now show that there exists a descent feasible direction at $\overline{x}$ for (10.32) which moves off the tth constraint. Since the set of row vectors of A is assumed to be linearly independent, by standard results in linear algebra, there exists a $y \in \mathbf{R}^n$ satisfying

$$\begin{aligned} A_i.y &= 1 \quad \text{for } i = t \\ &= 0 \quad \text{for } i \neq t. \end{aligned} \tag{10.36}$$

Let y be a solution for (10.36). From (10.35) and (10.36), we have $(\nabla\theta(\overline{x}))y = \overline{\pi}Ay = \overline{\pi}_t < 0$, and hence y is a descent feasible direction for (10.32) at $\overline{x}$.

Thus a necessary condition for a KKT point of (10.34) to be a KKT point for (10.32) is that all the Lagrange multipliers be nonnegative. Otherwise we can construct a descent feasible direction for (10.32) at such a point. These results are used in some of the algorithms discussed below, to solve NLP's involving linear inequality constraints using techniques for solving NLP's involving linear equality constraints only. They try to guess the set of active inequality constraints at the optimum, and apply the equality constraint techniques to the problem treating these active constraints as equations. Modifications are made in the active set using Lagrange mulitplier information gathered over each step.

10.10.2 The General Problem

Here we consider the NLP

$$\begin{array}{lll} \text{minimize} & \theta(x) & \\ \text{subject to} & A_i.x = b_i, \ i = 1 \text{ to } m & \\ & \geqq b_i, \ i = m+1 \text{ to } m+p & \end{array} \tag{10.37}$$

where $x \in \mathbf{R}^n$, and $\theta(x)$ is a real valued continuously differentiable function. Given a feasible point x, the first order necessary conditions for x to be a local minimum for this problem are that there exists a Lagrange multiplier vector $\pi = (\pi_1, \ldots, \pi_{m+p})$ satisfying

$$\begin{aligned} \nabla\theta(x) &= \sum_{i=1}^{m+p} \pi_i A_i. \\ \pi_i &\geqq 0, \ i = m+1 \text{ to } m+p \\ \pi_i(A_i.x - b_i) &= 0, \ i = m+1 \text{ to } m+p. \end{aligned} \tag{10.38}$$

Without any loss of generality we assume that $\{A_i. : i = 1 \text{ to } m\}$ is linearly independent. Let $\mathbf{K}$ denote the set of feasible solutions of (10.37). Given $\overline{x} \in \mathbf{K}$, all the equality constraints for $i = 1$ to m are active at $\overline{x}$ in (10.37). For $m+1 \leqq i \leqq m+p$, the ith constraint in (10.37) is active at $\overline{x}$ (also said to be an active inequality constraint at $\overline{x}$) if $A_i.\overline{x} = b_i$, inactive otherwise. Let $\mathbf{I}(\overline{x}) = \{i : A_i.\overline{x} = b_i\}$, the index set of active

constraints at $\overline{x}$. The point $y \in \mathbf{R}^n$, $y \neq 0$, is a feasible direction at $\overline{x}$ if $\overline{x} + \lambda y \in \mathbf{K}$ for $0 \leqq \lambda \leqq \overline{\lambda}$, for some positive $\overline{\lambda}$. Clearly y is a feasible direction at $\overline{x}$ iff

$$\begin{aligned} A_i.y &= 0, \ i = 1 \text{ to } m \\ &\geqq 0, \ i \in \mathbf{I}(\overline{x}) \cap \{m+1, \ldots, m+p\}. \end{aligned}$$

If y is a feasible direction at $\overline{x}$ and $A_i.y > 0$ for some $i \in \mathbf{I}(\overline{x}) \cap \{m+1, \ldots, m+p\}$, a move in the direction y from $\overline{x}$ is said to move off the ith constraint in (10.37).

We will now discuss some algorithms for solving (10.37).

10.10.3 The Frank-Wolfe Method

To solve (10.37), this method generates a descent sequence of feasible points $\{x^r : r = 0, 1, \ldots\}$ beginning with an initial feasible solution x^0, satisfying $\theta(x^{r+1}) < \theta(x^r)$ for all r.

For $k \geqq 0$, in step $k+1$, the initial point is x^k, the feasible point obtained at the end of the previous step if $k > 0$, or the feasible point with which the method is initiated, if $k = 0$. In this step the search direction y^k is of the form $z^k - x^k$ where z^k is a feasible point satisfying $(\nabla\theta(x^k))(z^k - x^k) < 0$, and so y^k is a descent direction at x^k. To find a point like z^k, we solve the LP in variables x

$$\begin{array}{ll} \text{minimize} & (\nabla\theta(x^k))x \\ \text{subject to} & x \in \mathbf{K}. \end{array} \tag{10.39}$$

If z^k is an optimum solution obtained when the LP is solved and $(\nabla\theta(x^k))^T z^k = (\nabla\theta(x^k))^T x^k$, then x^k is also optimal to the LP (10.39). By the duality theorem of linear programming, there exists a vector π^k such that x^k, π^k together satisfy the first order necessary optimality conditions (10.38) for (10.37), and so we terminate with x^k as the solution for (10.37). Otherwise, since $x^k \in \mathbf{K}$, we must have $(\nabla\theta(x^k))(z^k - x^k) < 0$, and so $y^k = z^k - x^k$ is a feasible descent direction at x^k. Now do a line search to find the minimum of $\theta(x^k + \alpha y^k)$ subject to $0 \leqq \alpha \leqq 1$. If α_k is the minimum for this line search problem, the next point in the sequence is $x^{k+1} = x^k + \alpha_k y^k$, continue.

We have the following results about the convergence properties of this method.

Theorem 10.2 *Suppose $\mathbf{K} \neq \emptyset$ and that the linear function in x, $(\nabla\theta(\tilde{x}))x$, is bounded below on $x \in \mathbf{K}$ for each $\tilde{x} \in \mathbf{K}$. Assume that $\mathbf{K}$ has at least one extreme point, and that for each k, the optimum solution z^k for the LP (10.39) obtained in the method is an extreme point of $\mathbf{K}$. If the method does not terminate after a finite number of steps, the sequence $\{x^r : r = 0, 1, \ldots\}$ generated by the above method has at least one limit point, and every limit point of this sequence is a KKT point for (10.37), if the line searches are carried out exactly in each step.*

Proof. Since $\nabla\theta(\tilde{x})x$ is bounded below for $x \in \mathbf{K}$ for each $\tilde{x} \in \mathbf{K}$, the LP (10.39) has an optimum solution always. The LP (10.39) may have alternate optima, and we are

assuming that z^k is an optimum solution for (10.39) which is an extreme point of $\mathbf{K}$ (this will be the case, for example, if $\mathbf{K}$ has at least one extreme point and (10.39) is solved by the simplex method). Since $\mathbf{K}$ is a convex polyhedron, it has a finite number of extreme points, and let $\mathbf{K}^{\Delta}$ be the convex hull of these extreme points. Because of the descent property $\theta(x^r)$ is monotonic decreasing as r increases, and by the manner in which the algorithm is carried out, it is clear that every point in the infinite sequence $\{x^r\}$ lies in the convex hull of $\mathbf{K}^{\Delta}$ and x^0, a compact set. So the sequence $\{x^r\}$ has at least one limit point. Let $\overline{x}$ be a limit point of the sequence $\{x^r\}$. Let $\mathbf{S}$ be an infinite set of positive integers such that $x^k \to \overline{x}$ as $k \to \infty$ with all $k \in \mathbf{S}$. For each $k \in \mathbf{S}$ we have an associated extreme point of $\mathbf{K}$, z^k, which is an optimum solution of (10.39). Since there are only a finite number of extreme points of $\mathbf{K}$, there must exist at least one extreme point of $\mathbf{K}$, say $\overline{z}$, which is equal to z^k for $k \in \mathbf{S}$ an infinite number of times. Let $\mathbf{S}_1 \subset \mathbf{S}$ such that for each $k \in \mathbf{S}_1$, $z^k = \overline{z}$. So $(\nabla\theta(x^k))^T(\overline{z} - x^k) < 0$ for each $k \in \mathbf{S}_1$. $x^k \to \overline{x}$ as $k \to \infty$ through $k \in \mathbf{S}_1$, so taking the limit in the above inequality as $k \to \infty$ through $k \in \mathbf{S}_1$, we get

$$(\nabla\theta(\overline{x}))(\overline{z} - \overline{x}) \leqq 0. \tag{10.40}$$

By our hypothesis, the line searches are carried out exactly in each step. Let $\mathbf{S}_1 = \{r_t : t = 1 \text{ to } \infty\}$, with the elements in $\mathbf{S}_1$ arranged in increasing order. So limit $x^{r_t} = \overline{x}$ as $t \to \infty$. In step $k = 1 + r_t$, the optimal step length is α_{1+r_t}, and so we must have, for $0 \leqq \alpha \leqq 1$,

$$\theta(x^{r_t} + \alpha(\overline{z} - x^{r_t})) \geqq \theta(x^{1+r_t}) \geqq \theta(x^{r_{t+1}}). \tag{10.41}$$

This follows because x^{1+r_t} is the point on the line segment $\{x^{r_t} + \alpha(\overline{z} - x^{r_t}) : 0 \leqq \alpha \leqq 1\}$ which minimizes $\theta(x)$ on this line segment. Also, since r_t is an increasing sequence, we have $r_{t+1} \geqq 1 + r_t$, and since $\{\theta(x^1), \theta(x^2), \ldots\}$ is a descent sequence we have $\theta(x^{1+r_t}) \geqq \theta(x^{r_{t+1}})$. In (10.41) let $t \to \infty$. This leads to

$$\theta(\overline{x} + \alpha(\overline{z} - \overline{x})) - \theta(\overline{x}) \geqq 0 \tag{10.42}$$

for all $0 \leqq \alpha \leqq 1$. When α is sufficiently small and positive, by the mean value theorem of calculus, (10.42) implies that $\alpha(\nabla\theta(\overline{x}))(\overline{z} - \overline{x}) \geqq 0$, that is, $(\nabla\theta(\overline{x}))(\overline{z} - \overline{x}) \geqq 0$. Combining this with (10.40) we have

$$\nabla\theta(\overline{x})(\overline{z} - \overline{x}) = 0. \tag{10.43}$$

Since $\overline{z}$ is an optimum solution of (10.39) whenever $k \in \mathbf{S}_1$, and since $x^k \to \overline{x}$ as $k \to \infty$ with all $k \in \mathbf{S}_1$, by (10.43) we conclude that $\overline{x}$ is a feasible solution for (10.37) satisfying the property that $x = \overline{x}$ is an optimum solution of the LP

$$\begin{array}{ll} \text{minimize} & (\nabla\theta(\overline{x}))x \\ \text{subject to} & A_i.x = b_i,\ i = 1 \text{ to } m \\ & \qquad \geqq b_i,\ i = m+1 \text{ to } m+p. \end{array} \tag{10.44}$$

Let $\overline{\pi} = (\overline{\pi}_1, \ldots, \overline{\pi}_{m+p})$ be an optimum dual solution associated with (10.44), then by the duality and complementary slackness theorems of linear programming, $\overline{x}$, $\overline{\pi}$ together satisfy (10.38), and hence $\overline{x}$ is a KKT point for (10.37).

□

If $\theta(x)$ is convex, and x^k is a point obtained during the Frank-Wolfe method, and satisfies $(\nabla\theta(x^k))(x^k - z^k) \leqq \varepsilon$, where z^k is an optimum solution of (10.39), then $\theta(x^k) \leqq \varepsilon +$ minimum value of $\theta(x)$ in (10.37). To see this, since $\theta(x)$ is convex, we have for $x \in \mathbf{K}$, $\theta(x) - \theta(x^k) \geqq (\nabla\theta(x^k))(x - x^k) \geqq (\nabla\theta(x^k))(z^k - x^k) \geqq -\varepsilon$, and so $\theta(x) \geqq \theta(x^k) - \varepsilon$ for all $x \in \mathbf{K}$. So if $\theta(x)$ is convex and x^k satisfies $(\nabla\theta(x^k))(x^k - z^k) < \varepsilon$, where ε is small, we can conclude that x^k is near optimum and terminate.

In each step of this method, an LP and a line search problem have to be solved. Even though the system of constraints in the LP to be solved in all the steps is the same, the objective function changes from step to step. The line search problem in each step has to be solved either optimally or at least to guarantee a sufficient decrease in the objective value. Since there is a considerable amount of work to be done in each step, the method tends to be slow. It is practical to use the method only on such problems for which the structure of the problem allows the solution of the LP in each step by an efficient special algorithm. One such application arises in the study of traffic flow along a city's street network using a traffic assignment model. We discuss this application briefly here.

The Traffic Assignment Problem

Let $G = (\mathcal{N}, \mathcal{A})$ be a city's street network. $\mathcal{N}$ is a set of points which are the various centers in the city or street intersections. $\mathcal{A}$ is a set of arcs or street segments, each arc joining a pair of points. The prupose of the study is to determine how the traffic will be distributed over alternate routes. Each driver makes his own choice of the route to take, but traffic flow on road network exhibits certain patterns. One broad principle for the analysis of traffic movement enunciates that traffic distributes itself over alternative routes so that the average journey time is a minimum.

The cost associated with an arc (i, j) in the network is a measure of the journey time from node i to node j along that arc. Journey time is influenced by traffic congestion, and tends to increase with traffic flow. Let f_{ij} denote the traffic flow on this arc (i. e., the number of cars entering this arc at node i per unit time) and let $c_{ij}(f_{ij})$ denote the journey time as a function of the flow f_{ij}. This function has the shape given in Figure 10.10, and so is a monotone increasing convex function.

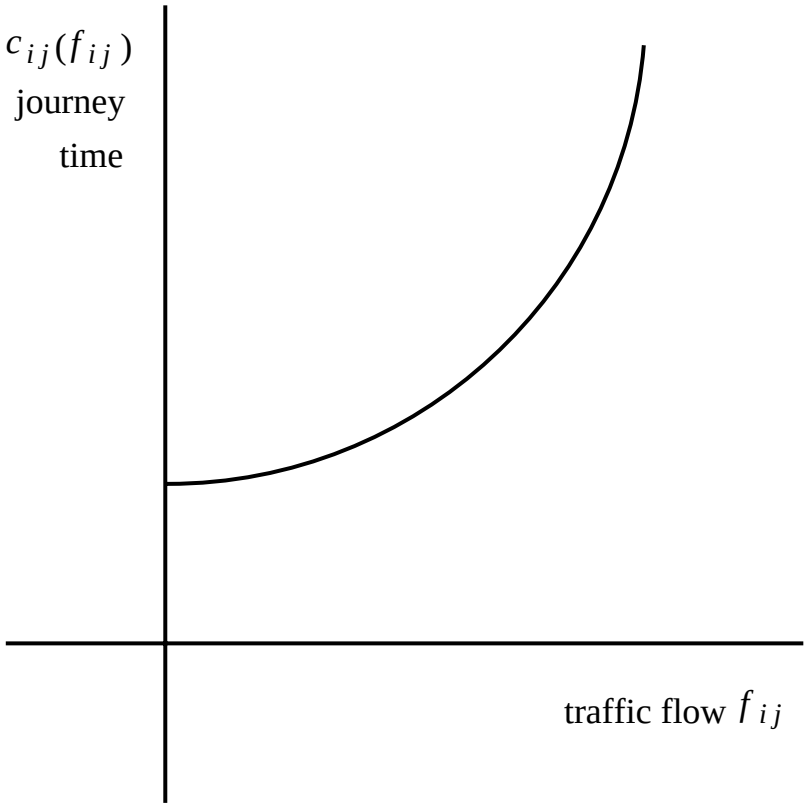


Figure 10.10

Traffic modelers construct these functions $c_{ij}(f_{ij})$ by actually collecting data. They also have data on the volumes of traffic (i. e., how many people travel and from where, to where) for different periods of the day. For example, during a particular peak period, suppose we know that V^u vehicles will be travelling from node s_u (origin) to node t_u (destination) in the network, $u = 1$ to g. Let f_{ij}^u be the number of these vehicles (with origin s_u and destination t_u) travelling along arc (i, j) in the network. For $u = 1$ to g let $f^u = (f_{ij}^u)$ be the vector of arc flows of the s_u to t_u vehicle flows. The problem is to determine these vectors f^u. The traffic assignment model states that the $(f^u : u = 1 \text{ to } g)$ form an optimum solution to the following nonlinear (flow dependent cost) multicommodity flow problem

$$\begin{aligned}
&\text{minimize} \quad \sum_{(i,j)\in\mathcal{A}} c_{ij}(f_{ij}) \\
&\text{subject to } f_{ij} = \text{flow on arc } (i,j) = \sum_{u=1}^{g} f_{ij}^u \\
&\qquad \sum(f_{ij}^u : j \text{ such that } (i,j) \in \mathcal{A}) - \sum(f_{ji}^u : j \text{ such that } (j,i) \in \mathcal{A}) \qquad (10.45) \\
&\qquad\quad = \quad 0, \text{ if } i \neq s_u \text{ or } t_u \\
&\qquad\quad = \quad V^u, \text{ if } i = s_u \\
&\qquad\quad = -V^u, \text{ if } i = t_u \\
&\qquad f_{ij}^u \geqq \quad 0, u = 1 \text{ to } g, (i,j) \in \mathcal{A}.
\end{aligned}$$

In this model it is quite reasonable to make the simplifying assumption that the flow variables are continuous variables rather than discrete integer variables. Also, since the cost function $c_{ij}(f_{ij})$ is constructed to reflect the journey time as a function of the flow f_{ij}, there is no need to include a constraint in the model corresponding to the capacity for flow of this arc. So, (10.45) is an uncapacitated, convex, multicommodity flow problem, and this can be solved efficiently using the Frank-Wolfe method. It begins with a feasible flow $((f^u)^0 : u = 1 \text{ to } g)$, which can be generated by standard network flow methods, and generates a sequence of feasible flow vectors $((f^u)^r : u = 1 \text{ to } g : r = 0, 1, \ldots)$ converging to the optimum solution of (10.45). In the $(k+1)$th step of this method, the initial flow vectors are $((f^u)^k, u = 1 \text{ to } g)$. Let $(f_{ij})^k = \sum_{u=1}^{g} (f_{ij}^u)^k$, the total flow on arc (i, j) in these flow vectors. Let $\overline{c}_{ij} = \left(\frac{dc_{ij}(f_{ij})}{df_{ij}} : \text{ evaluated at } f_{ij} = (f_{ij})^k\right)$. Then the LP to be solved in this step is

$$\begin{aligned}
\text{minimize } & \sum_{u=1}^{g} \sum_{(i,j)\in\mathcal{A}} \overline{c}_{ij} f_{ij}^u \\
\text{subject to } & \sum (f_{ij}^u : j \text{ such that } (i,j) \in \mathcal{A}) - \sum (f_{ji}^u : j \text{ such that } (j,i) \in \mathcal{A}) \\
& = \quad 0, \text{ if } i \neq s_u \text{ or } t_u \\
& = \quad V^u, \text{ if } i = s_u \\
& = -V^u, \text{ if } i = t_u \\
& f_{ij}^u \geqq \quad 0, u = 1 \text{ to } g, (i,j) \in \mathcal{A}.
\end{aligned} \tag{10.46}$$

Clearly, (10.46) can be broken up into g separate network flow problems one for each $u = 1$ to g. Also, the uth problem becomes the shortest chain problem from s_u to t_u in the network $G = (\mathcal{N}, \mathcal{A})$ with $(\overline{c}_{ij})$ as the vector of arc lengths, for which there are very efficient special algorithms.

Let $\mathcal{P}_u$ be the shortest chain from s_u to t_u in G with $(\overline{c}_{ij})$ as the vector of arc costs. Define the flow vector $z^u = (z_{ij})^u$ where

$$\begin{aligned}
z_{ij}^u &= V^u \text{ if } (i,j) \text{ is on } \mathcal{P}_u \\
&= 0 \text{ otherwise.}
\end{aligned}$$

Then the flow vectors $(z^u : u = 1 \text{ to } g)$ are an optimum solution of the LP (10.46), to be solved in this step.

Since the objective function in (10.45) is separable in the arcs, even the line search problem to be solved in this step, which is that of minimizing $\sum_{(i,j)\in\mathcal{A}} c_{ij}(f_{ij}^1 + \ldots + f_{ij}^g)$ over the line segment $\{f^u = \alpha(f^u)^k + (1-\alpha)z^u, u = 1 \text{ to } g, 0 \leqq \alpha \leqq 1\}$, can be simplified.

Thus the Frank-Wolfe method provides a reasonable approach for solving the traffic assignment problem (10.45). The main reason for this is the fact that the LP to be solved in each step of the method breaks down into g separate shortest chain problems, for which very efficient special algorithms are available.

10.10.4 Reduced Gradient Methods

The name reduced gradient method refers to a method which uses the equality constraints to eliminate some variables (the dependent or basic variables) from the problem, and treats the remaining problem in the space of the independent (or nonbasic) variables only, either explicitly or implicitly. The gradient of the objective function in the space of independent variables is the reduced gradient defined in Section 10.9, the search direction is usually the steepest descent vector in the space of the independent variables (the negative reduced gradient vector); or the Newton search direction in the space of the independent variables, determined using the reduced Hessian or an approximation for it.

We will consider the problem in the following form

$$\begin{array}{ll} \text{minimize} & \theta(x) \\ \text{subject to} & Ax = b \\ & l \leqq x \leqq u \end{array} \tag{10.47}$$

where A is a matrix of order $m \times n$ and rank m. As discussed in Chapter 1, the problem (10.37) can be put in this form. Here l, u are the lower and upper bound vectors for x in (10.47). Let B be a basis for A (i. e., a square nonsingular submatrix of A of order m) and partition A as (B, D), and let $x = (x_B, x_D)$ be the corresponding partition of the vector x. x_D is the vector of independent (nonbasic) variables and x_B is the vector of dependent (basic) variables. Let $\overline{x} = (\overline{x}_B, \overline{x}_D)$ be a feasible solution for (10.47). So $\overline{x}_B = B^{-1}(b - D\overline{x}_D)$. The problem can be transformed into one in the space of independent variables x_D only, by eliminating the dependent variables x_B. The reduced gradient at $\overline{x}$ is $\overline{c}_D = (\nabla_{x_D}\theta(\overline{x})) - (\nabla_{x_B}\theta(\overline{x}))B^{-1}D$. Define $\overline{y}_D = (\overline{y}_j)$ by

$$\begin{aligned} \overline{y}_j = -\overline{c}_j \text{ if } & x_j \text{is a nonbasic variable in } x_D \text{ and either } \overline{c}_j < 0 \text{ and} \\ & \overline{x}_j < u_j \text{ or } \overline{c}_j > 0 \text{ and } \overline{x}_j > l_j \\ = 0 \text{ if } & x_j \text{ is a nonbasic variable in } x_D, \\ & \text{and the above conditions not met.} \end{aligned}$$

If $\overline{y}_D = 0$, $\overline{x}$ satisfies the first order necessary optimality condition for being a local minimum for (10.47), and the method terminates. Otherwise verify that $\overline{c}_D\overline{y}_D < 0$, so $\overline{y}_D$ is a descent direction in the space of independent variables x_D. It is the steepest descent (negative reduced gradient) direction. Define $\overline{y}_B = -B^{-1}D\overline{y}_D$ and let $\overline{y} = (\overline{y}_B, \overline{y}_D)$. Then $\overline{y}$ is the search direction at $\overline{x}$. Since $A\overline{y} = 0$, the equality constraints in (10.47) continue to be satisfied when we move in this direction. Define

$$\begin{aligned} \lambda_1 &= \text{ minimum } \{(\overline{x}_j - l_j)/(-\overline{y}_j) : j \text{ such that } \overline{y}_j < 0\}, \\ \lambda_2 &= \text{ minimum } \{(u_j - \overline{x}_j)/(\overline{y}_j) : j \text{ such that } \overline{y}_j > 0\}, \\ \overline{\lambda} &= \text{ minimum } \{\lambda_1, \lambda_2\}. \end{aligned}$$

Do a line search for minimizing $\theta(\overline{x} + \lambda\overline{y})$ over $0 \leqq \lambda \leqq \overline{\lambda}$, and repeat the whole process with the optimum point in this line segment.

Let l_B, u_B denote the bound vectors for the dependent variables x_B. If $l_B < \overline{x}_B < u_B$, from the definition of the search direction $\overline{y}$, it can be verified that $\overline{\lambda} > 0$. If however, (10.47) is degenerate, given a feasible solution $\overline{x}$ for it, it may not be possible to find a basis B for (10.47) for which $l_B < \overline{x}_B < u_B$ holds. In this degenerate case, it may so happen that $\overline{\lambda} = 0$. In this case $\overline{y}$ is not a feasible direction at $\overline{x}$, and the line search problem does not make any sense, since any move of positive length in the direction $\overline{y}$ results in infeasibility. In this case the method can be continued by identifying the active constraints at $\overline{x}$, and moving from $\overline{x}$ in the direction of $\overline{y}^p$, the orthogonal projection of $\overline{y}$ in the subspace of active constraints at $\overline{x}$ (this will be a gradient projection step, see the next section, Section 10.10.5). This is equivalent to carrying out the line search problem exactly as above after replacing $\overline{y}$ by $\overline{y}^p$.

For convergence and rate of convergence results in this method see [10.2, 10.13, 10.15, 10.17, 10.26].

This method has been generalized very directly into the Generalized Reduced Gradient method (GRG) for solving NLPs involving nonlinear constraints. See [10.2, 10.13, 10.15, 10.17, 10.25, 10.26].

10.10.5 The Gradient Projection Method

When applied to solve the NLP (10.37), this method generates a descent sequence $\{x^r : r = 0, 1, \ldots\}$ beginning with a feasible point x^0, all the points in which are feasible. Step 1 begins with x^0, and in general for $k \geqq 1$ step $k+1$ begins with the point x^k at the end of step k.

For any feasible solution $\overline{x}$ of (10.37) define $\mathbf{I}(\overline{x}) = \{i : A_i.\overline{x} = b_i\}$. Clearly, $\{1, \ldots, m\} \subset \mathbf{I}(\overline{x})$ for all feasible solutions $\overline{x}$.

In step $k+1$, if there are no equality constraints in the problem and if $\mathbf{I}(x^k) = \emptyset$, choose the search direction at x^k to be $y^k = -(\nabla\theta(x^k))^T$. If $\mathbf{I}(x^k) \neq \emptyset$, the search direction in this step is determined by projecting the negative gradient of the objective function at x^k, onto the subspace parallel to the affine space of currently active constraints treated as equations. Let A_k denote the matrix whose rows are $A_i.$ for $i \in \mathbf{I}(x^k)$. So A_k is of order $|\mathbf{I}(x^k)| \times n$. Assume that the set of rows of A_k is linearly independent, otherwise delete some dependent row vectors of A_k from it until this property holds. The projection matrix corresponding to the active subspace is $P_k = I - A_k^T(A_kA_k^T)^{-1}A_k$. The projection of $-(\nabla\theta(x^k))^T$ onto the active subspace is $-P_k(\nabla\theta(x^k))^T$. It can be verified that this vector $-P_k(\nabla\theta(x^k))^T$ is a positive multiple of the vector which minimizes $(\nabla\theta(x^k))y$ subject to $A_ky = 0$ and $y^Ty \leqq 1$.

If $-P_k(\nabla\theta(x^k))^T = 0$, define $\beta^k = (A_kA_k^T)^{-1}A_k(\nabla\theta(x^k))^T$. Then $\nabla\theta(x^k) - (\beta^k)^TA_k = 0$. $(\beta^k)^T$ is a row vector of dimension $|\mathbf{I}(x^k)|$. Augment $(\beta^k)^T$ into a vector of dimension $m+p$, by inserting 0's for all $i \notin \mathbf{I}(x^k)$, and let the vector obtained be called π^k. Then $\nabla\theta(x^k) = \pi^kA$ where A is the $(m+p)\times n$ coefficient matrix in (10.37). So if $\pi_i^k \geqq 0$ for all $i = m+1$ to $m+p$, x^k, π^k together satisfy the first order necessary optimality conditions (10.38) and the method terminates with x^k as the KKT point

for (10.37). On the other hand if $\pi_i^k < 0$ for some i between $m+1$ to $m+p$, identify the i for which π_i^k is the most negative, say r, delete the rth constraint from the active set (that is, eliminate $A_{r.}$ from the matrix A_k) update the projection matrix, and the projection of $-(\nabla\theta(x^k))^T$ on the new active subspace, and repeat the whole process.

If $-P_k(\nabla\theta(x^k))^T \neq 0$, define $y^k = -P_k(\nabla\theta(x^k))^T$, y^k is the search direction at x^k. It can be verified that P_k is symmetric and $P_k^T P_k = P_k$, so P_k is PSD. Also $\nabla\theta(x^k)y^k = -\|y^k\|^2 < 0$. So y^k is a descent direction. Now find $\overline{\lambda}$ from

$$\overline{\lambda} = \text{ minimum } \left\{ \frac{A_{i.}x^k - b_i}{-A_{i.}y^k} : i \text{ such that } i \notin \mathbf{I}(x^k) \text{ and } A_{i.}y^k < 0 \right\}$$
$$= +\infty \text{ if } A_{i.}y^k \geqq 0 \text{ for all } i \notin \mathbf{I}(x^k).$$

Do a line search to minimize $\theta(x^k + \lambda y^k)$, $0 \leqq \lambda \leqq \overline{\lambda}$. If λ_k is the optimum step length in this line search problem, $x^{k+1} = x^k + \lambda_k y^k$ is the new point; go to the next step.

Methods for Updating the Projection Matrices

The periodic updating of the projection matrix is a considerable computational problem. However, the matrix A_k usually changes by one row, say $A_{r.}$, which is either dropped from the set of active constraint rows, or is added to it. Here we discuss how to efficiently update $(A_kA_k^T)^{-1}$ when a change like this takes place.

To Delete a Row From A_k

Let $A_{r.}$ be the sth row in A_k at the moment and suppose we want to delete it from A_k. After deletion suppose A_k becomes $\hat{A}$, of order $(q-1) \times n$.

Interchange the last row and the sth row in $(A_kA_k^T)^{-1}$. In the resulting matrix interchange the sth column and the last column. After these interchanges suppose this matrix $(A_kA_k^T)^{-1}$ is written down in partitioned form as

$$\begin{pmatrix} E & u \\ u^T & \delta \end{pmatrix}$$

where E is of order $(q-1) \times (q-1)$. Then it can be shown that $(\hat{A}\hat{A}^T)^{-1} = E - \frac{uu^T}{\delta}$.

To Add a Row to A_k

Suppose the row $A_{r.}$ has to be added to A_k. We will make $A_{r.}$ as the last row of the resulting matrix, which is $\tilde{A} = \begin{pmatrix} A_k \\ A_{r.} \end{pmatrix}$. Let $\overline{P}$ be the projection matrix corresponding to A_k, which is $I - A_k^T(A_kA_k^T)^{-1}A_k$. Compute $c = \|\overline{P}(A_{r.})^T\|^2 = A_{r.}\overline{P}(A_{r.})^T$, $w = (A_kA_k^T)^{-1}A_k(A_{r.})^T$, $u = -(w/c)$, $F = (A_kA_k^T)^{-1} + \frac{ww^T}{c}$. Then

$$(\tilde{A}\tilde{A}^T)^{-1} = \begin{pmatrix} F & u \\ u^T & 1/c \end{pmatrix}.$$

In the process of this updating, if c turns out to be zero, i. e., $\overline{P}A_r. = 0$, then the new active constraint row, $A_r.$, is linearly dependent on the previous active constraint rows, and the updating cannot be carried out. In this case the new active constraint row is ignored and the method can be continued with the same set of active constraint rows as before.

The updating procedure can also be used recursively to obtain the inverse $(A_k A_k^T)^{-1}$ in the first step of the algorithm, from the set of active constraints at that stage, by introducing them one at a time. An advantage of this recursion is that it selects the largest set of linearly independent active constraint rows among the set of all active constraint rows at this stage.

10.10.6 The Active Set Methods

We consider the NLP (10.37). These methods begin with a feasible solution x^0 and obtain a descent sequence $\{x^r : r = 0, 1, \ldots\}$, where each point in the sequence is feasible.

If $\overline{x}$ is an optimum solution for (10.37), and $\mathbf{I}(\overline{x}) = \{i : A_i.\overline{x} = b_i, i = 1 \text{ to } m+p\}$, then $\overline{x}$ is also an optimum solution of the equality constrained NLP

$$\begin{array}{ll} \text{minimize} & \theta(x) \\ \text{subject to} & A_i.x = b_i, i \in \mathbf{I}(\overline{x}). \end{array} \tag{10.48}$$

If we can guess the correct active set $\mathbf{I}(\overline{x})$, we could solve (10.48) by methods for solving equality constrained NLPs discussed in Section 10.9.

In these methods, a guess is built up over the steps, on the likely set of active constraint indices at the optimum. This set is known as the **working active set**. The working active set in step $k+1$ is denoted by $\mathbf{I}_k$. Clearly $\{1, \ldots, m\} \subset \mathbf{I}_k$ for all k. Changes are made in the set $\mathbf{I}_k$ using information gathered in each step. $\mathbf{I}_k$ always satisfies the property: $\{A_i. : i \in \mathbf{I}_k\}$ is linearly independent. The initial point in step 1 is x^0, in initial feasible solution with which the method is initiated. For $k \geqq 1$, the initial point in step $k+1$ is x^k, the feasible point obtained at the end of step k. Usually we have $\mathbf{I}_k \subset \mathbf{I}(x^k)$.

In step $k+1$, we carry a step for the equality constrained minimization problem

$$\begin{array}{ll} \text{minimize} & \theta(x) \\ \text{subject to} & A_i.x = b_i, i \in \mathbf{I}_k \end{array} \tag{10.49}$$

as discussed in Section 10.9. The search direction at x^k is the direction determined using the projected gradient, the projected Hessian or some quasi-Newton search direction at x^k for (10.49) as discussed in Section 10.9.

If x^k satisfies the termination criteria for (10.49), let A_k denote the matrix with rows $A_i.$, $i \in \mathbf{I}_k$. The corresponding Lagrange multiplier vector for (10.49) is $\beta^k = (A_k A_k^T)^{-1} A_k (\nabla\theta(x^k))^T$ from (10.31). If $\beta_i^k \geqq 0$ for all $i \in \mathbf{I}_k \cap \{m+1, \ldots, m+p\}$, as

discussed in Section 10.9.1, x^k is a KKT point for (10.37), terminate. If $\beta_i^k < 0$ for some $i \in \mathbf{I}_k \cap \{m+1, \ldots, m+p\}$, select the most negative among these, say β_r^k, then delete r from the working active set, and repeat the whole process.

If x^k does not satisfy the termination criteria for (10.49), let y^k be the search direction generated at x^k for solving (10.49). Find out $\overline{\lambda}$ from

$$\begin{aligned}\overline{\lambda} = &\text{ minimum } \left\{ \frac{A_{i.}x^k - b_i}{-A_{i.}y^k} : i \text{ such that } i \notin \mathbf{I}_k \text{ and } A_{i.}y^k < 0 \right\} \\ &= \infty \text{ if } A_{i.}y^k \geqq 0 \text{ for all } i \notin \mathbf{I}_k.\end{aligned}$$

Do a line search to minimize $\theta(x^k + \lambda y^k)$ over $0 \leqq \lambda \leqq \overline{\lambda}$. Let λ_k be the optimum step length for this line search problem. If $\lambda_k < \overline{\lambda}$, leave the working active set $\mathbf{I}_k$ unchanged, and with $x^{k+1} = x^k + \lambda_k y^k$ go to the next step. If $\lambda_k = \overline{\lambda}$, all the i which tie for the minimum in the definition of $\overline{\lambda}$ join the active set, select one of these and include it in $\mathbf{I}_k$. Then go to the next step.

To carry out a step of the algorithm discussed in Section 10.9 for the equality constrained minimization problem (10.49), we need the corresponding matrix Z, which we denote by Z_k here, as discussed in Section 10.9. Whenever we change the working active set $\mathbf{I}_k$ by dropping an element from it, or including a new element in it, it is necessary to make the corresponding changes in Z_k. Suppose Z_k is computed as in (10.27) using a basis B_k for A_k, and maintained by storing B_k either explicitly or in some factored form. Whenever $\mathbf{I}_k$ changes by one element, B_k changes by one row and one column, and B_k^{-1} can be updated by using the standard pivot methods of LP.

Several practical strategies have been developed to decide when to include a constraint in the working active set, and when to drop a constraint from it. Software packages for linearly constrained nonlinear programming based on such active set strategies seem to give the most satisfactory performance. Many of the commercially available packages usually include a combination of several of the strategies discussed above, in order to satisfactorily solve the widest class of problems.

All these methods become considerably simplified when applied to solve a quadratic programming problem, because of the special nature of the objective function.

10.11 Exercises

10.2 Fermat's Problem

Let $A_{.j} = (a_{1j}, \ldots, a_{mj})^T$, $j = 1$ to n be given distinct points in $\mathbf{R}^m$. Let w_j be a given positive weight associated with point $A_{.j}$. For any $x \in \mathbf{R}^m$ define $f(x) = \sum_{j=1}^n w_j \|x - A_{.j}\|$.

(i) If no three points among $\{A_{.1}, \ldots, A_{.n}\}$ are collinear, prove that $f(x)$ is positive and strictly convex on $\mathbf{R}^m$.

(ii) Assuming that no three points in the set $\{A_{.j} : j = 1 \text{ to } n\}$ are collinear prove that the problem of minimizing $f(x)$ over $\mathbf{R}^n$ has a unique solution, call it $\overline{x}$, and prove that $\overline{x}$ lies in the convex hull of $\{A_{.1}, \ldots, A_{.n}\}$.

(iii) Define

$$g(x) = \sum_{j=1}^{n} \Big(\frac{w_j(A_{.j} - x)}{\|x - A_{.j}\|} \Big), \quad \text{if } x \neq A_{.j}, \text{ for each } j = 1 \text{ to } n.$$

For such points, $g(x) = -\nabla f(x)$. This function $g(x)$ given above, is not defined if $x = A_{.j}$ for some j. By analogy, define for $j = 1$ to n,

$$h(A_{.j}) = \sum_{\substack{i=1 \\ i \neq j}}^{n} \Big(\frac{w_i(A_{.i} - A_{.j})}{\|A_{.i} - A_{.j}\|} \Big)$$

$$g(A_{.j}) = \text{ maximum } \{\|h(A_{.j})\| - w_j, 0\} \Big(\frac{h(A_{.j})}{\|h(A_{.j})\|} \Big).$$

Prove that a given point x is $\overline{x}$ (whether x is one of the points in the set $\{A_{.j} : j = 1 \text{ to } n\}$ or not) iff $g(x) = 0$, with $g(x)$ defined as above.

(iv) Assume that no three points in the set $\{A_{.j} : j = 1 \text{ to } n\}$ are collinear.
Define:

$$T(x) = \Big(\sum_{j=1}^{n} \frac{w_j A_{.j}}{\|x - A_{.j}\|} \Big) / \Big(\sum_{j=1}^{n} \frac{w_j}{\|x - A_{.j}\|} \Big), \quad \text{if } x \neq A_{.j} \text{ for each } j = 1 \text{ to } n$$

$$T(A_{.j}) = A_{.j}, \text{ for each } j = 1 \text{ to } n.$$

Prove that $T(\overline{x}) = \overline{x}$. Also prove that if x is such that $x \neq A_{.j}$ for each $j = 1$ to n and $T(x) = x$, then $x = \overline{x}$.
Prove that if $x \in \mathbf{R}^m$ satisfies $x \neq T(x)$, then $f(T(x)) < f(x)$.
Consider the interative method $x^0 =$ initial point in $\mathbf{R}^m$ choosen so that

$$x^0 \neq A_{.j} \text{ for each } j = 1 \text{ to } n$$
$$x^{r+1} = T(x^r), r = 0, 1, \ldots.$$

If $x^r \notin \{A_{.j} : j = 1 \text{ to } n\}$ for all r, prove that the sequence $\{x^r : r = 0, 1, \ldots\}$ converges to $\overline{x}$.

(v) Let $A_{.j}$ be the jth column vector of the following matrix for $j = 1$ to 6.

$$\begin{pmatrix} -2 & -1 & 1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{pmatrix}.$$

Let $w_j = 1$ for all $j = 1$ to 6. In this case prove that $\overline{x} = (0, 0)^T$.

Show that there is an x_1^0 (approximately 1.62) such that for $x^0 = (x_1^0, 0)^T$ we have $T(x^0) = A_{.3}$, which is not optimal. This shows that the iterative method discussed in (iv) may not always converge to $\overline{x}$ even if the initial point $x^0 \notin \{A_{.j} : j = 1 \text{ to } n\}$.

However, prove that there exists a countable set Γ of points in $\mathbf{R}^m$ such that if $x^0 \notin \Gamma$, then the sequence of points generated by the iterative method discussed in (iv) converges to $\overline{x}$.
(H. W. Kuhn [10.21])

10.3 Consider the NLP

$$\begin{array}{ll} \text{minimize} & \theta(x) \\ \text{subject to} & f(x) = 0 \end{array}$$

where $\theta(x)$ and $f(x)$ are both continuously differentiable real valued functions defined over $\mathbf{R}^n$. Using the ideas of the reduced gradient method and the results given by the implicit function theorem, develop an efficient algorithm for solving this problem.

10.4 Define the diameter of a convex hexagon (convex polytope with six extreme points in $\mathbf{R}^2$) $\mathbf{K}$ to be maximum $\{\|x - z : x, z \in \mathbf{K}\}$. Formulate the problem of finding a maximum area convex hexagon of diameter $\leqq 1$, as an NLP. Is this a convex programming problem? Find a solution to this problem using some of the algorithms discussed in this book.

10.5 $D = (d_{ij})$ is a square symmetric matrix of order n satisfying, $d_{ii} = 0$ for all i, the triangle inequality ($d_{ij} + d_{jk} \geqq d_{ik}$ for all i, j, k), and $d_{ij} > 0$ for all $i \neq j$. It is the matrix of Euclidean distances between pairs of points among a set of n points in $\mathbf{R}^2$. We are given the matrix D, but not the actual points from which D was calculated. It is required to find the coordinates (x_i, y_i), $i = 1$ to n, of n points in $\mathbf{R}^2$, for which the pairwise distance matrix is D. Formulate this as an NLP and discuss an efficient approach for solving it.

The rectilinear or L_1-distance between two points (x_1, y_1), (x_2, y_2) in $\mathbf{R}^2$ is defined to be $|x_1 - x_2| + |y_1 - y_2|$. Consider the version of the above problem of finding the coordinates of n points in $\mathbf{R}^2$, for which the matrix of pairwise rectilinear distances is a given matrix D. Formulate this problem. Is this easier or harder to solve than the version for the Euclidean distances? Why?
(S. M. Pollock)

10.6 Let $n > 1$, $x = (x_1, \ldots, x_n)^T$, $S_k = \sum_{j=1}^{n} x_j^{\mathbf{k}}$. Consider the NLP

$$\begin{array}{ll} \text{minimize} & S_3^{\mathbf{2}} - S_2 S_4 \\ \text{subject to} & 0 \leqq x_j \leqq 1, \ j = 1 \text{ to } n. \end{array}$$

Prove that the vector $x = (x_j)$ is a strict local minimum for this problem if m of the x_j are equal to 1, and p of the x_j are equal to $1/2$, where $m + p = n$ and $n > m > (1/9)n$. Also, prove that x is a global minimum for this problem if it is of the above form and either m or p is $\lfloor (1/2)n \rfloor$.
(P. Wolfe [10.41])

10.7 Automatic Voltage Regulator Control Panel (AVR) Design Problem. AVR's are used to stabilize voltage in electrical power systems. AVR contains many circuits, each circuit may consist of several components like resistors, transistors, capacitors, zener diodes etc. Each component is characterized by some variables (e. g. the resistence of a resistor measured in ohms, the gain value of a transistor measured in hFE, the capacitance of a capacitor measured in microfared (MF) etc.). The problem is to find an optimum design (i. e., find the optimal values of all the variables) which stabilizes the output voltage as far as possible, while the input voltage may fluctuate uncontrollably in some specified range. Here we provide a simplified example relating to the triggering circuit design in the AVR control panel for a diesel 2MW AC generator, to illustrate the basic principles involved in modelling and solving this class of problems (the general problem may have many more variables, and the functions involved are more complicated and may have many more terms, but the basic features remain identical). The functional form for the output voltage as a function of the input voltage and the design variables is available from electrical engineering theory. Given this function, and the range of fluctuation of the input voltage, the problem is to find optimal values for the design variables that stabilizes the output voltage as much as possible. In our example, the positive and negative voltages are denoted by v_1, v_2; each of these fluctuates between 14.25 to 15.75 and we have no way of controlling it. There are 5 design variables, x_1, x_2, x_3, x_4, x_5. The functional form for the output voltage v is the following:

$$\begin{aligned} v_3 &= v_1(1 - \mathrm{e}^{-(\mathbf{0.5/x_1x_5})}) \\ v_4 &= (x_4(x_3 + 100) + 100v_2)/(x_3 + 200) \\ v &= (v_3 - v_4)\mathrm{e}^{-(\mathbf{10/x_2x_5})}. \end{aligned}$$

The constraints on the variables are, $1 \leqq x_5 \leqq 10$, $3 \leqq x_4 \leqq 15$, $10 \leqq x_1 \leqq 200$, $100 \leqq x_2 \leqq 4000$, $1 \leqq x_3 \leqq 1000$. Formulate the problem as a nonlinear program and discuss an algorithm for solving it.
(Kirloskar Electricals Ltd., India)

10.8 The variable y represents the yield in a chemical process. There are n process variables $x_1, x_2, \ldots, x_n$ (such as temperature, flow rate, etc.) which influence the yield y. Data was collected to observe the yield y for various values of the process variable vector $x = (x_1, \ldots, x_n)$. This leads to k data points, $t = 1$ to k.

Process variable vector $x^t = (x_1^t, \ldots, x_n^t)$, corresponding yield y_t.

In the vectors x^t, $t = 1$ to k, each process variable takes several values spanning its possible range of variation, and each combination of process variables takes several values in the combined range of variation of the vector of these process variables. It is believed that y can be approximated reasonably well by a convex quadratic function of the form $Q(x) = cx + (\frac{1}{2})x^T Dx$. It is required to find the best convex quadratic fit $Q(x)$ for y, using the available data. Formulate this problem of finding the best convex quadratic approximation $Q(x)$ for y using the available data as a nonlinear programming problem, and discuss how this problem can be solved.

If l_j, u_j are known lower and upper bounds for the process variable x_j for $j = 1$ to n, and you are asked to design an experiment for collecting the necessary data in the above problem, outline how you will determine the process variable vectors $x^t = (x_1^t, \ldots, x_n^t)$ at which the yield has to be observed, in order to obtain the best fit.

10.9 Let $\theta(x)$ be a continuously differentiable real valued function defined on $\mathbf{R}^n$. It is required to find the unconstrained minimum of $\theta(x)$ over $\mathbf{R}^n$. Beginning with an initial point $x^0 \in \mathbf{R}^n$, the sequence of points $\{x^r : r = 0, 1, 2, \ldots\}$ was obtained by using Cauchy's Method of steepest descent with optimal step lengths in each step (the metric matrix for determining the steepest descent direction is always the unit matrix I). Prove that $(x^{r+2} - x^{r+1})^T(x^{r+1} - x^r) = 0$ for all r.

10.10 Let c be a given row vector in $\mathbf{R}^n$. Write down explicitly, an optimum solution for the following problem

$$\begin{array}{ll} \text{minimize} & cx \\ \text{subject to} & x^T x = 1 \\ & x \geqq 0. \end{array}$$

10.11 Let $\theta(x)$ be a continuously differentiable real valued convex function defined on a bounded convex set $\mathbf{K} \subset \mathbf{R}^n$, that attains its minimum over $\mathbf{K}$ at $x^* \in \mathbf{K}$. $\{x^r : r = 1, 2, \ldots\}$, $\{y^r : r = 1, 2, \ldots\}$ are sequences of points in $\mathbf{K}$ satisfying the following conditions

$$\nabla\theta(x^r)(y^r - x^r) \leqq \text{ Infimum } \{\varepsilon_r + \nabla\theta(x^r)(x - x^r) : x \in \mathbf{K}\}$$

$$\nabla\theta(x^r)(y^r - x^r) \to 0 \text{ as } r \to \infty$$

where $\varepsilon_r \geqq 0$ for all r and $\varepsilon_r \to 0$ as $r \to \infty$. Then, prove that $\theta(x^r) \to \theta(x^*)$ as $r \to \infty$.

10.12 We are given a set of n points in $\mathbf{R}^2$, say, $a^t = (a_1^t, a_2^t)$, $t = 1$ to n. It is required to fit a circle to these points. The objective function to be minimized is $\sum$ ((r^2-square of the Euclidean distance between a^t and the center)^2 : $t = 1$ to n), where r is the radius of the circle. Formulate this problem as an NLP and discuss an efficient method for solving it.
(R. Chandrasekaran)

10.13 We are given row vectors $c^1, \ldots, c^r$ in $\mathbf{R}^n$ and real numbers $d_1, \ldots, d_r$. Define

$$\theta(x) = \text{ Maximum } \{|c^t x - d_t| : t = 1 \text{ to } r\}.$$

It is required to find the unconstrained minimum of $\theta(x)$ over $x \in \mathbf{R}^n$. Discuss an efficient method for computing it.

10.14 Let $d_1, \ldots, d_n$ be given positive integers. The partition problem with this data, is to check whether there exists a subset $\mathbf{S} \subset \{1, \ldots, n\}$ such that

$$\sum_{i \in \mathbf{S}} d_i = \sum_{i \in \overline{\mathbf{S}}} d_i$$

where $\overline{\mathbf{S}} = \{1, \ldots, n\} \setminus \mathbf{S}$. This is a well known $\mathcal{NP}$-complete problem (see [8.12]). Formulate this problem as a special case of

$$\begin{array}{ll} \text{minimize} & \|x\|_p \\ \text{subject to} & x \in \mathbf{K} = \{x : Ax \geqq b\} \end{array} \tag{10.50}$$

where $\|x\|_p = (\sum_{i=1}^{n} |x_i|^{\mathbf{p}})^{\mathbf{1/p}}$, and A, b are integer matrices of orders $m \times n$ and $m \times 1$ respectively, p a positive integer $\geqq 1$, and $\mathbf{K}$ is known to be nonempty and bounded. $\|x\|_p$ is known as the p-norm of the vector x. Thereby establish that the problem of maximizing the p-norm on a convex polytope specified in terms of linear inequalities with integer data, is an $\mathcal{NP}$-hard problem.

Show that an upper bound on the optimum objective value in (10.50) can be obtained by solving a relaxed linear program.

The ∞-norm of the vector $x = (x_i) \in \mathbf{R}^n$, denoted by $\|x\|_\infty$ is defined to be maximum $\{|x_i| : i = 1 \text{ to } n\}$. Show that when $p = \infty$, (10.50) can be solved by solving at most $2n$ linear programs.

(O. L. Mangasarian and T.-H. Shiau, "A variable-complexity norm maximization problem", Technical Report 2780, Mathematics Research Center, University of Wisconsin, Madison, 1984)

10.15 Optimal Betting in a Race Track

The "market" at a race track in North America typically convenes for about 20 minutes, during which participants make bets on any number of 6 to 12 horses in the following race. To keep the discussion simple, we consider a race in which participants can bet on each horse either to **win** or **place**. All participants who have bet on a horse to **win**, realize a positive return on that bet only if that horse comes first, while a **place** bet realizes a positive return if that horse comes first or second. Consider a race with the following data declared at the time we are ready to bet.

n = number of horses running in the race.

W_i = total amount bet by public (all participants so far) on horse i to **win**.

W = $\sum_{i=1}^{n} W_i$ = **win** pool.

Q = track payback proportion (typically .83, it is the proportion of pool given away; the remaining proportion .17 is kept by the race track company).

P_j = total amount bet by public (all participants so far) on horse j to **place**.

P = $\sum_{j=1}^{n} P_j$ = **place** pool.

q_i = probability that horse i finishes first in the race.

q_{ij} = $\frac{q_i q_j}{1-q_i}$ = probability that horse i finishes first and horse j finishes second in the race.

$$\text{payoff per dollar bet on horse } i \text{ to } \textbf{win} = \begin{cases} \frac{WQ}{W_i}, & \text{iff horse } i \text{ comes in first place} \\ 0 & \text{otherwise.} \end{cases}$$

$$\text{payoff per dollar bet on horse } j \text{ to } \textbf{place} = \begin{cases} 1 + \frac{PQ - P_i - P_j}{2P_j}, & \text{if horses } i, j \text{ are first two winners in any order} \\ 0, & \text{if horse } j \text{ did not finish in the first two places in the race.} \end{cases}$$

Thus the payoff on horse j to **place** is independent of whether j finishes first or second, but dependent on which horse finishes with it in first two places.

We assume that $q_i = W_i/W$, that is that the crowd is good at picking a winner, or that the relative amount bet on a horse to **win** corresponds closely to its actual chances of winning.

The W_i, P_i are the public's bets in the race, are known. Consider the problem of determining the place bets to make optimally, given all the above data and the assumptions, subject to a budget of b \$. The Kelly criterion determines the optimal bets to maximize the expected logarithm of final wealth. The decision vector in this problem is $x = (x_1, \ldots, x_n)^T$, where x_i is the **place** bet on the ith horse, $i = 1$ to n. Define

$$f_{ij}(x) = \left(\frac{Q\left(P + \sum_{l=1}^{n} x_l\right) - x_i - x_j - P_i - P_j}{2} \right) \left(\frac{x_i}{x_i + P_i} + \frac{x_j}{x_j + P_j} \right).$$

Then the problem for determining the optimal x is

$$\begin{aligned} \text{minimize} \quad & \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} q_{ij} \log\left(f_{ij}(x) + b - \sum_{\substack{l=1 \\ l \neq i,j}}^{n} x_l \right) \\ \text{subject to} \quad & \sum_{l=1}^{n} x_l \leqq b \\ & x_l \geqq 0, \text{ for all } l = 1 \text{ to } n. \end{aligned}$$

Discuss an efficient approach for solving this problem. Solve the numerical problem using this approach, when the data is

$$n = 8, \; Q = 0.83, \; b = 500.$$

i	1	2	3	4	5	6	7	8
W_i	10,000	15,000	5,000	35,000	5,000	10,000	18,000	12,000
P_i	4,000	4,000	4,000	8,000	3,000	8,000	13,000	5,000

(See the delightful book, W. T. Ziemba and D. B. Hausch [10.42] for a complete treatment of this problem.)

10.16 Consider the following LP

$$\begin{array}{ll} \text{minimize} & cx \\ \text{subject to} & x \in \mathbf{K} = \{x : Ax \geqq b\} \end{array}$$

where A, b are given matrices of orders $m \times n$, $m \times 1$ respectively. Assume that $\mathbf{K} \neq \emptyset$. For $\lambda \geqq 0$, let $x(\lambda)$ denote the nearest point in $\mathbf{K}$ to λc in terms of the usual Euclidean distance.

If the above LP has an optimum solution, prove that there exists a $\overline{\lambda} > 0$ such that $x(\lambda)$ is constant for all $\lambda \geqq \overline{\lambda}$ and that $x(\overline{\lambda})$ is the least (Euclidean) norm optimum solution for the LP.

If the objective value is unbounded below on $\mathbf{K}$ in the above LP, prove that $\|x(\lambda)\| \to \infty$ as $\lambda \to \infty$.
(O. L. Mangasarian)

10.17 Consider the following NLP

$$\begin{array}{ll} \text{minimize} & \theta(x) \\ \text{subject to} & Ax \geqq b \end{array}$$

where $\theta(x)$ is a strictly convex function defined over $\mathbf{R}^n$ with a unique unconstrained minimum, $\overline{x}$, in $\mathbf{R}^n$; and A is a matrix of order $m \times n$. Suppose $\overline{x}$ satisfies

$$A_i.x - b_i \begin{cases} < 0, & i = 1 \text{ to } r \\ \geqq 0, & \text{for } i = r+1 \text{ to } m. \end{cases}$$

Let x^i denote the point which minimizes $\theta(x)$ subject to one constraint only "$A_i.x \geqq b_i$", for $i = 1$ to r. Suppose there is a unique $k \in \{1, \ldots, r\}$ such that x^k is feasible to the original NLP. Then prove that x^k is an optimum solution for the original NLP.

10.18 A Curve Fitting Application in High Voltage Coil Insulation Testing: The life of the insulation system on high voltage coils used in rotating electrical machines, depends on it's DLA (dielectric loss analyzer) value. The DLA value for a coil is expected to depend on it's $\Delta \tan \delta$ (increase in $\tan \delta$ or dissipation factor expressed as a percentage, with increase in test voltage) and ΔC (inrease in capacitance with increase in test voltage) values. The DLA value is hard to measure, but the $\Delta \tan \delta$ and ΔC values can be measured easily. Given below are the DLA, $\Delta \tan \delta$ and ΔC values for a sample of 95 test coils. Use this data to determine if the DLA value of a coil can be estimated reliably from it's $\Delta \tan \delta$ and ΔC values, and if so, determine the appropriate functional form. Using this analysis, design a scheme for checking the acceptability of coils (acceptable if DLA value is $\leqq 8.0$ units) using measurements of their $\Delta \tan \delta$ and ΔC values as far as possible.

ΔC and $\Delta \tan\delta$ with their corresponding DLA values for 95 test coils.

Sample coil No.	$\Delta \tan\delta$, 6.6 KV to 11 KV	Δ C, 6.6 KV to 11 KV	DLA at 11 KV	Sample coil No.	$\Delta \tan\delta$, 6.6 KV to 11 KV	Δ C, 6.6 KV to 11 KV	DLA at 11 KV
1	.0011	0.5	0.4	26	.0044	2.4	2.5
2	.0017	0.9	0.8	27	.0042	2.5	1.6
3	.0030	1.0	1.6	28	.0041	2.5	2.2
4	.0019	1.2	0.8	29	.0048	2.6	2.0
5	.0020	1.3	0.2	30	.0042	2.7	1.5
6	.0026	1.3	1.1	31	.0060	2.7	1.7
7	.0020	1.4	1.2	32	.0039	2.7	1.4
8	.0028	1.6	1.4	33	.0030	2.7	1.2
9	.0023	1.6	1.4	34	.0031	2.8	2.3
10	.0027	1.7	1.6	35	.0047	2.9	3.0
11	.0024	1.7	1.6	36	.0052	3.3	2.7
12	.0023	1.8	1.0	37	.0036	3.3	2.1
13	.0032	1.9	2.1	38	.0049	3.3	2.4
14	.0026	1.9	1.5	39	.0045	3.3	2.5
15	.0027	2.0	1.6	40	.0053	3.3	2.4
16	.0026	2.0	1.2	41	.0050	3.6	3.7
17	.0031	2.0	2.8	42	.0054	3.6	2.9
18	.0041	2.0	0.6	43	.0056	3.7	4.0
19	.0045	2.1	0.6	44	.0059	3.8	2.4
20	.0032	2.1	2.1	45	.0057	3.8	2.5
21	.0031	2.1	1.2	46	.0057	3.9	3.5
22	.0024	2.2	1.5	47	.0067	4.1	3.3
23	.0031	2.2	1.4	48	.0045	4.3	3.5
24	.0028	2.2	1.4	49	.0059	4.5	4.0
25	.0029	2.4	1.1	50	.0066	4.5	3.4

Sample coil No.	$\Delta \tan\delta$, 6.6 KV to 11 KV	Δ C, 6.6 KV to 11 KV	DLA at 11 KV	Sample coil No.	$\Delta \tan\delta$, 6.6 KV to 11 KV	Δ C, 6.6 KV to 11 KV	DLA at 11 KV
51	.0076	4.5	3.4	71	.0045	6.2	5.1
52	.0073	4.6	4.2	72	.0066	6.2	5.2
53	.0058	4.6	4.0	73	.0077	6.3	3.6
54	.0060	4.7	3.4	74	.0093	6.4	7.0
55	.0072	4.8	4.2	75	.0089	6.3	5.8
56	.0057	4.9	4.1	76	.0055	6.8	7.3
57	.0068	4.9	3.6	77	.0083	7.3	5.9
58	.0076	5.3	6.5	78	.0096	7.5	5.0
59	.0084	5.0	3.4	79	.0091	7.7	6.0
60	.0063	5.0	3.1	80	.0100	8.0	5.5
61	.0058	5.0	5.4	81	.0109	8.3	7.6
62	.0053	5.2	4.6	82	.0045	8.8	7.3
63	.0067	5.2	4.6	83	.0094	9.1	7.0
64	.0064	5.3	5.2	84	.0104	9.1	6.4
65	.0072	5.3	4.2	85	.0093	8.2	9.3
66	.0086	5.3	6.7	86	.0140	10.0	8.2
67	.0074	5.6	5.1	87	.0121	10.0	10.8
68	.0081	6.0	3.8	88	.0143	10.3	6.7
69	.0070	6.0	4.2	89	.0124	10.5	9.9
70	.0074	6.0	4.0	90	.0120	12.1	9.7
				91	.0131	11.1	10.4
				92	.0155	13.3	7.1
				93	.0127	14.6	9.0
				94	.0144	14.7	16.0
				95	.0139	15.4	13.2

10.12 References

10.1 L. Armijo, "Minimization of Functions Having Lipschitz Continuous First Partial Derivatives", *Pacific Journal of Mathematics*, 16 (1966) 1–3.

10.2 M. Avriel, *Nonlinear Programming, Analysis and Methods,* Prentice-Hall, Englewood Cliffs, New Jersey, 1976.

10.3 M. S. Bazaraa and C. M. Shetty, *Nonlinear Programming: Theory and Algorithms,* Wiley, New York, 1979.

10.4 J. Backen and G. P. McCormick, *Selected Applications of Nonlinear Programming,* Wiley, New York, 1968.

10.5 A. Ben-Israel, A. Ben-Tal and S. Zlobec, *Optimality in Nonlinear Programming: A Feasible Direction Approach,* Wiley-Interscience, New York, USA, 1981.

10.6 J. M. Borwein, "Necessary and sufficient conditions for quadratic minimality", *Numerical Function Analysis and Optimization*, 5 (1982) 127–140.

10.7 L. B. Contesse, "Une caratérisation complète des minima beaux en programmation Quadratique", *Numer. Math.,* 34 (1980) 315–332.

10.8 J. E. Dennis Jr., and J. J. Moŕe, "A Characterization of Superlinear Convergence and Its Application to Quasi-Newton Methods", *Mathematics of Computation*, 28 (1974) 549–560.

10.9 J. E. Dennis Jr., and R. B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations,* Prentice-Hall, Englewood Cliffs, NJ, USA, 1983.

10.10 G. P. Egorychev, "Reshenie Problemy van der Wardena, dha Permanentov", Inst. Fiziki im. L. V. Kirenskogo, USSR Acad. Sci. Siberian branch, Preprint IFSO – 13M, Krasnoiarsk, 1980.

10.11 G. P. Egorychev, "Novye formuly dha Permanenta", *Dokl. Acad. Sci. USSR,* 254 (1980) 784–787.

10.12 A. V. Fiacco and G. P. McCormick, *Nonlinear Programming: Sequential Unconstrained Minimization Techniques,* Wiley, New York, 1968.

10.13 R. Fletcher, *Practical Methods of Optimization, Vols. 1,2,* Wiley-Interscience, New York, 1981.

10.14 N. Frank and P. Wolfe, "An Algorithm for Quadratic Programming", *Naval Research Logistics Quarterly*, 3 (1956) 95–110.

10.15 P. E. Gill and W. Murray, *Numerical Methods for Cconstrained Optimization,* Academic Press, New York, 1974.

10.16 P. E. Gill, W. Murray, M. A. Sanders and M. H. Wright, "Model Building and Practical Implementations Aspects of Nonlinear Programming", Technical Report, Department of OR, Stanford University, Stanford, California, USA, 1984.

10.17 P. E. Gill, W. Murray and M. H. Wright, *Practical Optimization,* Academic Press, New York, 1981.

10.18 D. Goldfarb, "Extension of Davidson's Variable Metric Method to Maximization Under Linear Equality and Inequality Constraints", *SIAM Journal on Applied Mathematics*, 17 (1969) 739–764.

10.19 W. Hock and K. Schittkowski, "Test Examples for Nonlinear Programming Codes", *Lecture Notes in Economics and Mathematical Systems* Vol. 187, Springer-Verlag, New York (1981).

10.20 D. E. Knuth, "A Permanent Inequality", *The American Mathematical Monthly*, 88 (1981) 731–740.

10.21 H. W. Kuhn, "A Note on Fermat's Problem", *Mathematical Programming*, 4 (1973) 98–107.

10.22 C. L. Lawson and R. J. Hanson, *Solving Least Squares Problems,* Prentice-Hall, Englewood Cliffs, New Jersey, 1974.

10.23 C. Lemarechal and R. Mifflin, "Global and Superlinear Convergence of an Algorithm for One-Dimensional Minimization of Convex Functions", *Mathematical Programming*, 24 (1982) 241–256.

10.24 M. E. Lenard, "A Computational Study of Active Set Strategies in Nonlinear Programming With Linear Constraints", *Mathematical Programming*, 16 (1979) 81–97.

10.25 F. A. Lootsma, *Numerical Methods for Nonlinear Optimization,* Academic Press, New York, 1972.

10.26 G. P. McCormick, *Nonlinear Programming, Theory, Algorithms and Applications,* Wiley-Interscience, 1983.

10.27 G. P. McCormick, "A Modification of Armijo's Step-Size Rule for Negative Curvature", *Mathematical Programming*, 13 (1977) 111-115.

10.28 R. Mifflin, "A Superlinearly Convergent Algorithm for One-Dimensional Constrained Minimization Problems With Convex Functions", *Mathematics of Operations Research*, 8 (1983) 185–195.

10.29 R. Mifflin, "Stationarity and Superlinear Convergence of an Algorithm for Univariate Locally Lipschitz Constrained Minimization", *Mathematical Programming*, 28 (1984) 50–71.

10.30 J. J. More' and D. C. Sorensen, "On the Use of Directions of Negative Curvature in a Modified Newton Method", *Mathematical Programming*, 16 (1979) 1–20.

10.31 B. A. Murtagh and M. A. Saunders, "Large-Scale Linearly Constrained Optimization", *Mathematical Programming*, 14 (1978) 41–72.

10.32 K. G. Murty and S. N. Kabadi, "Some $\mathcal{NP}$-complete Problems in Quadratic and Nonlinear Programming", *Mathematical Programming*, 38 (1987), to appear.

10.33 J. M. Ortega and W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables,* Academic Press, New York, 1970.

10.34 P. Ribenboim, *13 Lectures on Fermat's Last Theorem,* Springer-Verlag, New York, 1980.

10.35 J. B. Rosen, "The gradient projection method for nonlinear programming, Part I: linear constraints", *SIAM J. Applied Math.,* 8 (1960) 181–217.

10.36 J. B. Rosen, "The gradient projection method for nonlinear programming, Part II: nonlinear constraints", *SIAM J. Applied Math.,* 9 (1961) 514–553.

10.37 L. E. Scales, *Introduction to Non-Linear Optimization,* Springer-Verlag, New York, 1985.

10.38 K. Schittkowski, *Nonlinear Programming Codes, Information, Tests, Performance,* Springer-Verlag, New York, 1980.

10.39 D. F. Shanno, "Conjugate Gradient Methods With Inexact Searches", *Mathematics of Operations Research*, 3 (1978) 244–254.

10.40 B. L. van der Waerden, "Problem, Jaresbericht", *Deutsch. Math. Verein*, 25 (1926) 117.

10.41 P. Wolfe, "Explicit Solution of an Optimization Problem", *Mathematical Programming*, 2 (1972) 258–260.

10.42 W. T. Ziemba and D. B. Hausch, *Beat the Race Track,* Harcourt Brace Jovanovich, New York, 1984.

Chapter 11

NEW LINEAR PROGRAMMING ALGORITHMS, AND SOME OPEN PROBLEMS IN LINEAR COMPLEMENTARITY

Some open research problems in linear complementarity have already been posed among the exercises in previous chapters. Here we discuss some more research problems briefly.

11.1 Classification of a Given Square Matrix M

Let M be a given square matrix of order n. In Section 1.3.1 we discussed algorithms to check whether M is PD os PSD, requiring a computational effort of at most n Gaussian pivot steps, or $\mathcal{O}(n^3)$ effort in terms of multiplications and additions. Such efficient algorithms are not known to check whether M belongs to other classes of matrices discussed in Chapters 2, 3.

As an example, consider the problem of checking whether M is a non-degenerate (i. e., principally non-degenerate to be specific) matrix. The question is: given M, to find whether there exists a subset of $\{1, \ldots, n\}$ such that the principal subdeterminant of M corresponding to that subset is zero. Since this question involves the existence of a subset of $\{1, \ldots, n\}$ satisfying a specified property which is easily checked (given a subset $\mathbf{J} \subset \{1, \ldots, n\}$, we can check whether $\mathbf{J}$ satisfies this property by computing the subdeterminant of M corresponding to $\mathbf{J}$, which takes at most $\mathcal{O}(r^3)$ effort, $r = |\mathbf{J}|$), this problem is in $\mathcal{NP}$, the class of decision problems which can be solved by a polynomially bounded non-deterministic algorithm (see M. Garey and D. Johnson's book [8.12] for precise definitions of these terms). We will now show that this problem is in fact $\mathcal{NP}$-complete. Given positive integers $d_0; d_1, \ldots, d_n$, the problem of checking

whether there exists a subset of $\{d_1, \ldots, d_n\}$ whose sum is equal to d_0, known as the **subset sum problem**, is the 0-1 problem of checking whether the following system has a solution

$$\begin{aligned} &\sum_{j=1}^{n} d_j x_j = d_0 \\ &x_j = 0 \text{ or } 1 \text{ for all } j \ . \end{aligned} \tag{11.1}$$

This problem is a well-known $\mathcal{NP}$-complete problem. Define M to be the matrix

$$M = \begin{pmatrix} d_0 & d_1 & d_2 & d_3 & \ldots & d_n \\ 1 & 1 & 0 & 0 & \ldots & 0 \\ 1 & 0 & 1 & 0 & \ldots & 0 \\ 1 & 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & 0 & \ldots & 1 \end{pmatrix} = \begin{pmatrix} d_0 & d \\ e & I_n \end{pmatrix}$$

where $d = (d_1, \ldots, d_n)$, e is the column vector of all 1's in $\mathbf{R}^n$, and I_n is the unit matrix of order n. A principal submatrix of M corresponding to a non-empty subset of $\{1, \ldots, n+1\}$ not containing 1 is a unit matrix of appropriate order, and hence has determinant 1. The principal subdeterminant of M corresponding to a subset of $\{1, \ldots, n+1\}$ of the form $\{1, i_1, \ldots, i_r\}$ can be verified to be $d_0 - (d_{i_1} + \ldots + d_{i_r})$. Thus the matrix M given above has a zero principal subdeterminant iff the system (11.1) has a solution. Since the $\mathcal{NP}$-complete problem (11.1) is a special case of the problem of checking whether a given square matrix has zero principal subdeterminant, this later problem is also an $\mathcal{NP}$-complete problem. This result is from [11.1] of R. Chandrasekaran, S. N. Kabadi and K. G. Murty.

The computational complexity of checking whether a given square matrix M is a P-matrix, P_0-matrix, Q-matrix, or Q_0-matrix is not known. For all these problems, finite algorithms are known. P- and P_0-properties can be checked by computing all the principal subdeterminants (requiring the evaluation of $2^{\mathbf{n}}$ determinantes when M is of order n). Finite algorithms for checking the Q- and Q_0-properties are provided in Exercises 3.87, 3.89 (when applied on a matrix of order n, these methods require the solution of at most $n^{2^{\mathbf{n}}}$ systems of linear inequalities, hence these methods though finite, are utterly impractical even for $n = 4$). No polynomially bounded algorithms for any of these problems are known so far, and it is also not known whether any of these problems is $\mathcal{NP}$-complete.

11.2 Worst Case Computational Complexity of Algorithms

In Chapter 6 we established that several of the pivotal algorithms for LCP are exponential growth algorithms in the worst case. However, the worst case computational complexity of the algorithm for solving the LCP (q, M) when M is PD symmetric matrix (or the corresponding nearest point problem) based on orthogonal projections discussed in Chapter 7 is still an open question.

11.2.1 Computational Complexity of the LCP Associated with a P-Matrix

In Chapter 8 we discussed polynomially bounded algorithms for the LCP (q, M) when M is either a Z-matrix, or a principally triangular P-matrix, or a PSD-matrix. The polynomially bounded ellipsoid methods work only when M is PSD, since they depend on the fact that the set $\{z : z^T(Mz + q) \leqq 0\}$ is convex, which may not hold when M is not PSD. None of the methods discussed in Chapter 8 are guaranteed to process the LCP (q, M) when M is a P-matrix which is not PSD. In this case the set $\{z : z^T(Mz + q) \leqq 0\}$ may not be convex. When M is a P-matrix, by the results in Chapter 3, the LCP (q, M) has the nice property of having a unique solution, but as yet no polynomially bounded algorithm is known for computing it. Establishing whether the LCP (q, M), where M is a P-matrix, can be solved by a polynomially bounded algorithm, remains an important mathematical problem in LCP theory.

11.2.2 A Principal Pivoting Descent Algorithm For the LCP Associated with a P-Matrix

In the LCP there is of course no objective function. In this algorithm from K. G. Murty [3.52] an extraneous distance function is computed and this distance decreases strictly in each step. The distance provides a measure of progress in the algorithm, it becomes zero iff we obtain a complementary feasible basis. The algorithm is a principal pivoting algorithm employing only single principal pivot steps, it can be used to solve the LCP (q, M) when M is a P-matrix. The algorithm can be initiated with any complementary basis. We now describe the algorithm.

Let A be the current complementary basis and y the corresponding complementary basic vector.

Find the nearest point in $\text{Pos}(A)$ to q in terms of the usual Euclidean distance (this can be found in polynomial time by the ellipsoid algorithm discussed in Section 8.4, or by the practically efficient algorithm discussed in Chapter 7). Let $\bar{x}$ be this nearest point and $d = ||\bar{x} - q||$, the Euclidean distance between $\bar{x}$ and q.

We will have $d = 0$ and $\bar{x} = q$ iff $q \in \text{Pos}(A)$. In this case y is a complementary feasible basic vector, and the solution of the LCP (q, M) is $(y = A^{-1}q, t = 0)$, where $t = (t_j)$ and t_j is the complement of y_j for all j.

If $d > 0$, let $\mathbf{B}(q, d) = \{x : ||x - q|| \leqq d\}$. $\mathbf{B}(q, d)$ is the closed ball with q as center and d as radius. Let $\mathbf{T}(q, \bar{x}) = \{x : (q - \bar{x})^T(x - \bar{x}) = 0\}$, it is the tangent hyperplane to $\mathbf{B}(q, d)$ at its boundary point $\bar{x}$. Since $\bar{x}$ is the nearest point in $\text{Pos}(A)$ to q, by the results in Chapter 7, $\bar{x}^T(q - \bar{x}) = 0$, $\mathbf{T}(q, \bar{x}) = \{x : x^T(q - \bar{x}) = 0\}$, it is a hyperplane containing the origin, 0. Since $\bar{x} \in \text{Pos}(A)$, we have $\bar{x} = \sum_{j=1}^{n} \alpha_j A_{.j}$ where $\alpha_j \geqq 0$ for all j. Let $\mathbf{J} = \{j : \alpha_j > 0\}$, $\bar{\mathbf{J}} = \{1, \ldots, n\} \setminus \mathbf{J}$. In this case since $q \notin \text{Pos}(A)$, by the results in Chapter 7, $\bar{x}$ must be a boundary point of $\text{Pos}(A)$, so $\bar{\mathbf{J}} \neq \emptyset$. For each

j let $D_{.j}$ be the complement of $A_{.j}$. By the results in Chapter 7, $\bar{x}$ is the orthogonal projection of q in the linear hull of $\{A_{.j} : j \in \mathbf{J}\}$, so the tangent hyperplane $\mathbf{T}(q, \bar{x})$ contains the linear hull of $\{A_{.j} : j \in \mathbf{J}\}$. By Theorem 3.20 of Section 3.3, $\mathbf{T}(q, \bar{x})$ must separate strictly, at least one of the pair of column vectors $\{A_{.j}, D_{.j}\}$ for some $j \in \bar{\mathbf{J}}$. Let $\mathbf{\Delta} = \{j : j \in \bar{\mathbf{J}}, A_{.j}$ and its complement are strictly separated by $\mathbf{T}(q, \bar{x})\}$. So $\mathbf{\Delta} \neq \emptyset$, select a $p \in \mathbf{\Delta}$ arbitrarily. Then in the notation of Chapter 7, $D_{.p}$ is on the near side of $\mathbf{T}(q, \bar{x})$, and $\text{Pos}\{\bar{x}, D_{.p}\}$ contains points which are strictly closer to q than $\bar{x}$. Thus if we make a single principal pivot step in position p in the complementary basis A, we get a new complementary basis whose pos cone contains points strictly nearer than $\bar{x}$ to q.

With $(y_1, \ldots, y_{p-1}, t_p, y_{p+1}, \ldots, y_n)$ as the new complementary basic vector, we repeat the whole process.

After each principal pivot step, the distance d strictly decreases, so a complementary basic vector can never reappear in the algorithm. Since there are only 2^n complementary basic vectors, the method must terminate after a finite number of principal pivot steps with the complementary solution for the problem.

Since the problem of finding the nearest point in a complementary cone which has a non-empty interior, to q, is equivalent to another LCP associated with a PD symmetric matrix, the method can be viewed as one for solving the LCP (q, M) associated with a P-matrix M by solving a finite number of LCP's associated with PD symmetric matrices.

The worst case computational complexity of this algorithm is still an open question.

One can get different variants of the algorithm by choosing p from $\mathbf{\Delta}$ according to different rules. One can consider the least index rule in which the p chosen from $\mathbf{\Delta}$ is always the least; or a cyclical rule like the least recently considered rule popular in implementations of the simplex algorithm. We can also consider a block principal pivoting method in which the new complementary basic vector at the end of the step is obtained by replacing each y_p in the present complementary basic vector, by its complement for each $p \in \mathbf{\Delta}$, in a block principal pivot step. The worst case computational complexity of each of these variants is currently under investigation.

Exercise

11.1 The rectilinear or L_1-distance between two points $x = (x_j)$, $y = (y_j)$ in $\mathbf{R}^n$ is defined to be $\sum_{j=1}^{n}(|x_j - y_j|)$. Consider the LCP (q, M) with M being a P-matrix. Let $y = (y_j)$ be a complementary basic vector for this problem associated with the complementary basis A. The nearest point in the complementary cone $\text{Pos}(A)$ to q in terms of the L_1-distance can be obtained by solving the LP

$$\begin{aligned} \text{minimize} \quad & \sum_{j=1}^{n}(u_i + v_i) \\ \text{subject to} \quad & Ay + u - v = q \\ & y, u, v \geqq 0 \ . \end{aligned}$$

If $(\bar{y}, \bar{u}, \bar{v})$ is an optimum solution to this LP, $\bar{x} = A\bar{y}$ is a nearest point in Pos(A) to q in terms of the L_1-distance.

If M is a P-matrix and $q \notin$ Pos(A), does there always exist a p such that the cone Pos $\{A_{.1}, \ldots, A_{.p-1}, D_{.p}, A_{.p+1}, \ldots, A_{.n}\}$, where $D_{.p}$ is the complement of $A_{.p}$, contains points which are strictly closer to q in terms of the L_1-distance, than $\bar{x}$? If so, discuss an efficient method for identifying such a p.

Develop a method for solving the LCP (q, M) when M is a P-matrix, that moves from one complementary cone to another, decreasing the L_1-distance to q in each step. Study the worst case computational complexity of this method.

11.3 Alternate Solutions of the LCP (q,M)

There are very nice conditions to check the uniqueness of a given solution for a linear programming problem, and to characterize and enumerate alternative optimum solutions when they exist. See [2.26].

For LCP, such characterizations or methods do not exist yet. A sufficient condition for the uniqueness of the solution for the LCP (q, M) is that M be a P-matrix. When M is not a P-matrix, alternate solutions may exist for the LCP (q, M), but in this case the algorithms discussed for the LCP find only one solution for the problem if they are able to process it, and then terminate.

Consider the LCP (q, M). Let $y = (y_j)$ be a complementary vector of variables for it, that is, for each j, $y_j \in \{w_j, z_j\}$. Let A be the complementary matrix corresponding to y. Let $t = (t_j)$ where t_j is the complement of y_j for each j. The complementary vector y leads to a solution of the LCP (q, M) iff the system

$$\begin{aligned} Ay &= q \\ y &\geqq 0 \end{aligned}$$

has a feasible solution. If $\bar{y}$ is a feasible solution of this system, $(y = \bar{y}, t = 0)$ is a solution of the LCP (q, M). If A is nonsingular, the above system has a feasible solution iff $A^{-1}q \geqq 0$, and in this case if it does have a solution, it is unique. If A is singular, the above system may have many feasible solutions. Whether it has a feasible solution or not can be determined by using Phase I of the simplex method for linear programming. If the above system is feasible, all alternate feasible solutions of it can be enumerated and the set of alternate feasible solutions compactly represented using standard results in linear programming [2.26], each such feasible solution leads to a solution of the LCP (q, M), as discussed above.

By solving the system of the type discussed above, for each of the complementary vectors of variables y, we can check whether the LCP (q, M) has a solution, and in fact

obtain all its solutions. This is a total enumeration method, requiring the solution of 2^n separate systems of linear equations in non-negative variables.

Since 2^n grows rapidly, the above total enumeration method for checking whether alternate solutions exist for a given LCP, or to obtain all solutions of it, is impractical unless n is very small. It would be nice if some efficient partial enumeration methods can be developed for doing the same job. These partial enumeration methods should identify subsets of complementary vectors of variables which do not lead to a solution of the LCP, and prune them, thereby saving some of the effort needed to carry out the enumeration. These methods would be similar to branch and bounds for 0-1 integer programming problems (see [1.28]) which are also partial enumeration methods.

We will now describe briefly one partial enumeration method for generating all the solutions of the LCP (q, M) discussed in K. G. Murty [11.3]. To keep the discussion simple, we make the assumption that q is nondegenerate. In this case, every complementary solution is a complementary BFS and it is adequate to enumerate among complementary basic vectors for all complementary solutions of the LCP (q, M).

The set of all variables in the LCP (q, M) is $\{w_1, \ldots, w_n; z_1, \ldots, z_n\}$. Given any subset $\mathbf{\Delta}$ of these variables, we will represent $\mathbf{\Delta}$ by a 0-1 incidence vector $a = (a_p) \in \mathbf{R}^{2n}$, a row vector, where

$$\text{for } j = 1 \text{ to } n, \quad a_j = \begin{cases} 1, & \text{if } w_j \in \mathbf{\Delta} \\ 0, & \text{if } w_j \notin \mathbf{\Delta} \end{cases}$$

$$a_{n+j} = \begin{cases} 1, & \text{if } z_j \in \mathbf{\Delta} \\ 0, & \text{if } z_j \notin \mathbf{\Delta} \end{cases}.$$

As an example, for $n = 4$, the incidence vector of the subcomplementary set $\{z_1, w_2, z_3\}$ is $(0, 1, 0, 0; 1, 0, 1, 0)$. So a complementary feasible basic vector for the LCP (q, M) corresponds to an incidence vector $x = (x_p) \in \mathbf{R}^{2n}$ satisfying $\sum_{p=1}^{2n} x_p = n$ and $x_j + x_{n+j} \geqq 1$, for each $j = 1$ to n, and the vector is a feasible basic vector, $x_p = 0$ or 1 for all $p = 1$ to $2n$. The second constraint that the vector be a feasible basic vector is not available explicitly in the form of a system of linear constraints, at the beginning; but we develop linear constraints in the x_p-variables corresponding to it during the course of the algorithm. In each step, more constraints of this type in the x_p-variables are generated and augmented to the system.

A set covering problem is a 0-1 integer programming problem of the following form.

$$\begin{aligned} \text{minimize} \quad & \sum_{p=1}^{2n} x_p \\ \text{subject to} \quad & Ex \geqq e_r \\ & x_p = 0 \text{ or } 1 \text{ for all } p \end{aligned}$$

where E is a 0-1 matrix of order $r \times 2n$ and e_r in the column vector of all 1's in $\mathbf{R}^r$. In each step, we solve a set covering problem of this form, and generate additional constraints for the set covering problem in the next step.

The set covering problem itself is an $\mathcal{NP}$-hard combinatorial optimization problem, but practically efficient branch and bound algorithms are available for it. The branch

and bound algorithm discussed in [1.28] for the set covering problem using the lower bounding strategy based on Lagrangian Relaxation may be particularly suitable, since we have to solve the problem repeatedly, with the only change between the problem in one step and the next being a few additional constraints.

A solution stack is maintained. Any solution to the LCP (q, M) found out during the algorithm is stored in the solution stack. At termination of the algorithm, this stack contains all the solutions of the LCP (q, M).

The initial set covering problem is

$$\begin{aligned} \text{minimize} \quad & \sum_{p=1}^{2n} x_p \\ \text{subject to} \quad & x_j + x_{n+j} \geqq 1, \text{ for each } j = 1 \text{ to } n \\ & x_p = 0 \text{ or } 1, \text{ for } p = 1 \text{ to } 2n \end{aligned}$$

The initial complementary basic vector is w. We will now describe a general step in the algorithm.

General Step

Let $y = (y_j)$ be the current complementary vector of variables with $y_j \in \{w_j, z_j\}$ for each $j = 1$ to n, and let A be the corresponding complementary matrix. Let $t = (t_j)$ where t_j is the complement of y_j for each $j = 1$ to n.

If A is singular, every complementary basic vector must include one of the variables from $\{t_1, \ldots, t_n\}$. Let $a \in \mathbf{R}^{2n}$ be the incidence vector of $\{t_1, \ldots, t_n\}$. Add the additional constraint "$ax \geqq 1$" to the set covering problem.

If A is nonsingular, y is a complementary basic vector, obtain the canonical tableau of the LCP (q, M) with respect to it. Suppose it is

y	t	
I	$-D$	$\bar{q}$

If $\bar{q} \geqq 0$, y is a complementary feasible basic vector, and $(y = \bar{q}, t = 0)$ is the corresponding complementary solution, include it in the stack. Every complementary basic vector different from y must include one of the variables from $\{t_1, \ldots, t_n\}$. Let $a \in \mathbf{R}^{2n}$ be the incident vector of $\{t_1, \ldots, t_n\}$. Include the additional constraint "$ax \geqq 1$" in the set covering problem.

If $\bar{q} \not\geqq 0$, y is not a feasible basic vector. For each i such that $\bar{q}_i < 0$, let $\mathbf{S}_i = \{t_j : j$ such that $-d_{ij} < 0\}$, where d_{ij} is the $(i, j)^{th}$ entry in the matrix D in the canonical tableau. Clearly, any feasible basic vector must include one of the variables from $\mathbf{S}_i$. Let a_i be the incidence vector of $\mathbf{S}_i$, include the additional constraint "$a_i x \geqq 1$" for each i satisfying $\bar{q}_i < 0$, in the set covering problem.

Solve the set covering problem together with the additional constraints added in this step.

If the optimum objective value in the set covering problem is $\geqq n+1$, terminate. The solution stack at this stage contains all the complementary solutions of the LCP (q, M).

If the optimum objective value in the set covering problem is n, let $\bar{x}$ be an optimum solution for it. Let $\bar{y}$ be the complementary vector of variables corresponding to the incidence vector $\bar{x}$. Make $\bar{y}$ the new complementary vector of variables. Go to the next step with it and the current set covering problem.

This algorithm has not been computationally tested and it is not known how it may work in practice.

Developing practically efficient partial enumeration methods for the general LCP remains a problem worth investigating.

11.4 New Approaches for Linear Programming

The well known primal simplex algorithm for linear programming starts at an extreme point of the set of feasible solutions, moves along an edge direction to an adjacent extreme point, and repeats the whole process until an optimal extreme point or an unbounded edge along which the objective value is unbounded below (for minimization problems) is reached. Thus all the direction used in the primal simplex algorithm are edge directions. Recently K. G. Murty and Y. Fathi [11.4] discussed versions of the simplex algorithm based on profitable directions of movement through the interior or relative interior of the set of feasible solutions or faces of it of dimension greater than 1. They showed that with simple modifications these methods can be proved to be finite, and can be implemented using basis inverses just as the usual versions of the simplex algorithm. Computational testes indicate that these modifications leads to improvements in the running time for solving linear programs.

N. Karmarkar [11.2] has developed an entirely new polynomially bounded algorithm for solving linear programs based on profitable search directions through the interior of the set of feasible solutions. This method closes in on an optimum by creating a sequence of spheres inside the feasible region for the LP. It is claimed that preliminary computational testing has shown this method to be much faster than the simplex algorithm for large scale linear programs. A statement of this algorithm with an intuitive justification is given in the Notation section in front of this book. Here we provide a detailed treatment of the algorithm and its polynomial boundedness.

Throughout this section the symbol e denotes the column vector of all 1s of appropriate dimension, and e^T denotes its transpose.

11.4.1 Karmarkar's Algorithm for Linear Programming

The Barrier Function Approach to Handle Inequality Constraints in Nonlinear Programming

Consider the following optimization problem (P).

$$\begin{array}{ll} \text{minimize} & \theta(x) \\ \text{subject to} & Ax = b \\ & g_i(x) \geqq 0, i = 1 \text{ to } m. \end{array} \tag{P}$$

A feasible solution x for this problem is said to be **strictly feasible** if $g_i(x) > 0$ for all $i = 1$ to m. The barrier function approach for solving this problem needs an initial strictly feasible point x^0. It generates a sequence of points $\{x^r : r = 0, 1, \ldots\}$, each x^r being a strictly feasible solution of the problem.

Barrier methods work by establishing a barrier on the boundary of the feasible region that prevents the search procedure from leaving the strictly feasible part of the feasible region. A barrier function for this problem is a continuous function $B(x)$ defined on $\mathbf{\Gamma} = \{x : g_i(x) > 0, \text{ for all } i = 1 \text{ to } m\}$ that tends to $+\infty$ as the point x approaches the boundary of $\mathbf{\Gamma}$. One commonly used barrier function is the logarithmic barrier function (suggested by K. R. Frisch in 1955)

$$B(x) = -\sum_{i=1}^{m} \log(g_i(x)).$$

Here log represents the natural logarithm. The barrier function method for (P) looks at the problem

$$\begin{array}{ll} \text{minimize} & F(x) = \theta(x) - \alpha \sum_{i=1}^{m} \log(g_i(x)) \\ \text{subject to} & Ax = b \end{array} \tag{B}$$

where α is a positive parameter known as the barrier parameter. Giving α some positive value and fixing it, the barrier method tries to solve (B), by some feasible direction descent method beginning with the initial strictly feasible point x^0. Consider the line search problem of minimizing $F(x)$ along the half-line $\{x + \lambda y : \lambda \geqq 0\}$, where x is a strictly feasible point. If $\lambda_1 > 0$ is such that $g_i(x + \lambda_1 y) = 0$ for some i between 1 to m, then the step length choosen in this line search problem will be $< \lambda_1$, since $-\log g_i(x + \lambda y) \to +\infty$ as $\lambda \to \lambda_1$ from below. Thus any line searches carried out for solving (B) beginning with a strictly feasible point will always lead to another strictly feasible point.

The barrier function method for solving (P) proceeds as follows. It selects a monotonic decreasing sequence of positive values $\{\alpha_r : r = 1, 2, \ldots\}$ converging to 0. Fixing $\alpha = \alpha_1$, it solves (B) by a feasible direction descent method, beginning with the initial strictly feasible point x^0. Suppose this terminates with the strictly feasible point x^1. Now α is changed to α_2, and the new (B) solved again beginning with the

initial strictly feasible solution x^1. The process is repeated in the same way, generating the sequence of strictly feasible points $\{x^r : r = 0, 1, \ldots\}$. Under mild conditions it can be shown that this sequence converges to a solution of (P). Karmarkar's algorithm for linear programming, closely resembles this nonlinear interior point barrier method. In his algorithm, Karmarkar uses a potential function which closely resembles the logarithmic barrier function.

We will now provide a theoretical description of Karmarkar's algorithm and proofs of its polynomial boundedness. A brief discussion on issues in implementing Karmarkar's algorithm will then follow. We divide this section into various numbered subsections, for ease of cross referencing.

1 Transforming Any LP Into Another With an Optimum Objective Value of Zero

We show that any LP can be transformed into another one with a known (minimum) objective value of zero.

Consider the LP

$$\begin{array}{lr} \text{minimize} & h\chi \\ \text{subject to} & E\chi \geqq p \\ & \chi \geqq 0. \end{array} \tag{11.2}$$

Let π denote the row vector of dual variables. It is well known (see [2.26]) that solving (11.2) is equivalent to solving the following system of linear inequalities.

$$\begin{array}{rl} h\chi - \pi p & \leqq 0 \\ E\chi & \geqq p \\ \pi E & \leqq h \\ \chi, \quad \pi & \geqq 0 \end{array} \tag{11.3}$$

There is no objective function in (11.3). If $(\bar{\chi}, \bar{\pi})$ is a feasible solution for (11.3), $\bar{\chi}$ is an optimum solution for the LP (11.2) and $\bar{\pi}$ is an optimum dual solution. If (11.3) is infeasible, either (11.2) is itself infeasible, or (11.2) may be feasible but its dual may be infeasible (in the later case, the objective value is unbounded below on the set of feasible solutions of (11.2)).

The system (11.3) can be expressed as a system of equations in nonnegative variables by introducing the appropriate slack variables. To solve the resulting system, construct the usual Phase I problem by introducing the appropriate artificial variables (see Chapter 2 in [2.26]). Let u denote the vector consisting of the variables χ_j, π_i, and the artificial variables. Let the Phase I problem corresponding to (11.3) be

$$\begin{array}{lr} \text{minimize} & gu \\ \text{subject to} & Fu = d \\ & u \geqq 0 \ . \end{array} \tag{11.4}$$

The optimum objective value in (11.4) is $\geqq 0$ (since it is a Phase I problem corresponding to (11.3)) and (11.3) is feasible iff it is zero. Let v denote the row vector of dual variables for (11.4). Consider the LP

$$\begin{array}{ll} \text{minimize} & gu - vd \\ \text{subject to} & Fu = d \\ & vF \leqq g \\ & u \geqq 0 \end{array} \tag{11.5}$$

The LP (11.5) consists of the constraints in (11.4) and its dual. From the duality theory of linear programming, the optimum objective value in (11.5) is zero (since (11.4) has a finite optimum solution). The LP (11.5) can be put in standard form for LPs by the usual transformations of introducing slack variables etc., see Chapter 2 in [2.26]. If $(\bar{u}, \bar{v})$ is optimal to (11.5), then $\bar{u}$ is optimal to (11.4). If $g\bar{u} = 0$, then the χ-portion of $\bar{u}$ is an optimum solution for (11.2). If $g\bar{u} > 0$, (11.3) is infeasible and hence (11.2) is either infeasible or has no finite optimum solution.

2 Transforming an LP Into Another With a Known Strictly Positive Feasible Solution

An LP in standard form with an optimum objective value of zero, can be transformed into another with the same property, but with a known strictly positive feasible solution. Consider the LP

$$\begin{array}{ll} \text{minimize} & gy \\ \text{subject to} & Gy = d \\ & y \geqq 0 \end{array} \tag{11.6}$$

where G is a matrix of order $m \times n$, and suppose all the data is integer and the optimum objective value in (11.6) is zero. Let $y^0 > 0$ by any integer vector in $\mathbf{R}^n$. Consider the new LP

$$\begin{array}{ll} \text{minimize} & gy + g_{n+1}y_{n+1} \\ \text{subject to} & Gy + y_{n+1}(d - Gy^0) = d \\ & y \geqq 0, \quad y_{n+1} \geqq 0 \end{array} \tag{11.7}$$

clearly $(y = y^0, y_{n+1} = 1) > 0$ is a feasible solution of (11.7). Since the optimum objective value in (11.6) is zero, the same property holds in (11.7) if g_{n+1} is sufficiently large (mathematically, it is sufficient to take $g_{n+1} > 2^s$, where s is the size of $\begin{pmatrix} G & G_{.n+1} \\ g & 0 \end{pmatrix}$, $G_{.n+1} = d - Gy^0$).

3 Transforming the Feasible Set into the Intersection of a Subspace with a Simplex

Given an LP in standard form with integer or rational data, with the optimum objective value of zero, and a strictly positive feasible solution, we can transform it into another,

for which the set of feasible solutions is $\mathbf{H} \cap \mathbf{S}$, where $\mathbf{H}$ is a subspace and $\mathbf{S}$ is the standard simplex. Consider the LP

$$\begin{array}{ll} \text{minimize} & \bar{g}y \\ \text{subject to} & \overline{G}y = \bar{d} \\ & y \geqq 0 \end{array} \tag{11.8}$$

where $\overline{G}$ is of order $m \times n$, and all the data is assumed to be integer. Let L be the size of this LP (i. e., L is the total number of digits in all the data in the LP (11.8) in binary encoding, see Sections 8.3 to 8.6 and Chapters 14, 15 in [2.26]).

Since (11.8) has an optimum solution, it has an optimum solution satisfying the additional constraint

$$\sum_{j=1}^{n} y_j \leqq M$$

where M is an upper bound depending on the size L. By the results in Chapter 8 (see also Chapter 15 in [2.26]) taking $M = 2^{\mathbf{L}}$ will do. Hence (11.8) is equivalent to

$$\begin{array}{ll} \text{minimize} & \bar{g}y \\ \text{subject to} & \overline{G}y = \bar{d} \\ & e^T y \leqq M \\ & y \geqq 0 \end{array}$$

where $e^T = (1, 1, \ldots, 1) \in \mathbf{R}^n$. By introducing the slack variable y_{n+1}, this LP is the same as

$$\begin{array}{ll} \text{minimize} & \bar{g}y \\ \text{subject to} & \overline{G}y - \frac{1}{M}\bar{d}\left(\sum\limits_{j=1}^{n+1} y_j\right) = 0 \\ & \sum\limits_{j=1}^{n+1} y_j = 1 \\ & y_j \geqq 0, \quad j = 1 \text{ to } n+1 \end{array} \tag{11.9}$$

The system $\overline{G}y - \frac{1}{M}\bar{d}\left(\sum_{j=1}^{n+1} y_j\right) = 0$ is a homogeneous system of equations, and hence its set of feasible solutions is a subspace $\mathbf{H}$ in $\mathbf{R}^{n+1}$. The system $\sum_{j=1}^{n+1} y_j = 1$, $y_j \geqq 0$ for $j = 1$ to $n+1$ defines the standard simplex $\mathbf{S}$ in $\mathbf{R}^{n+1}$. So the set of feasible solutions of (11.9) is $\mathbf{H} \cap \mathbf{S}$, as desired.

4 Minimization of a Linear Function Over a Spherical Ball or an Ellipsoid

Consider the problem

$$\begin{array}{ll} \text{minimize} & cx \\ \text{subject to} & (x - x^0)^T(x - x^0) \leqq \rho^{\mathbf{2}} \end{array}$$

If $c = 0$, every point in the sphere is optimal to this problem. If $c \neq 0$, the optimal solution of this problem is $x^0 - c^T\rho$, it is the point obtained by taking a step of length ρ (radius of the sphere) from the center x^0 in the direction of $-c^T$.

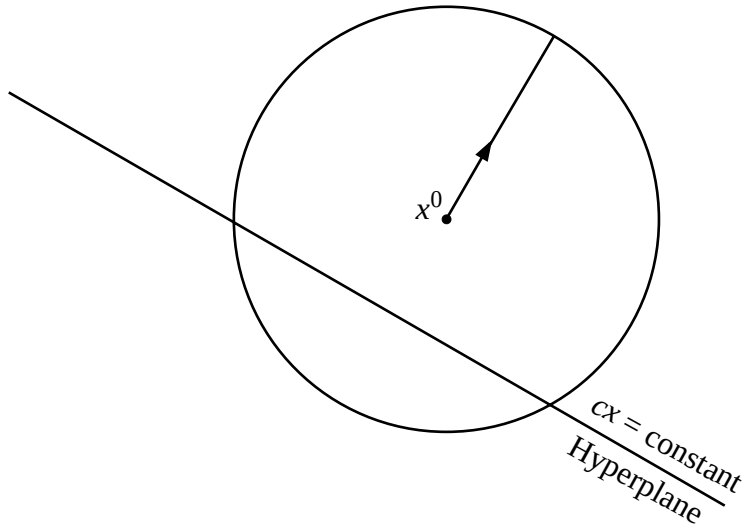


Figure 11.1 To minimize cx on the sphere, walk from the center x^0 in the direction $-c^T$, a step of length ρ =radius. The direction $-c^T$ is the steepest descent direction for the linear function cx.

Now consider the problem

$$\begin{array}{ll} \text{minimize} & cx \\ \text{subject to} & Ax = b \\ \text{and} & x \in \mathbf{B} = \{x : ||x - x^0|| \leqq \rho\} \ . \end{array}$$

Let $\mathbf{H} = \{x : Ax = b\}$. $\mathbf{H}$ is an affine space. If $\mathbf{H} \cap \mathbf{B} \neq \emptyset$, it is a lower dimensional sphere inside the affine space $\mathbf{H}$. Again if $c = 0$, every point in $\mathbf{H} \cap \mathbf{B}$ is optimal to this problem. If $c \neq 0$, let $\bar{c}$ be the orthogonal projection of c onto $\mathbf{H}$. $\bar{c} = 0$ if c is a linear combination of the rows of A, in this case the objective function is a constant on $\mathbf{H} \cap \mathbf{B}$, and every point in it is optimal. If $\bar{c} \neq 0$, the optimal solution of this problem is the point obtained by taking a step of length equal to the radius of the lower dimensional sphere $\mathbf{H} \cap \mathbf{B}$ from its center in the direction of $-\bar{c}^T$.

Consider the following problem

$$\begin{array}{ll} \text{minimize} & cx \\ \text{subject to} & Ax = b \\ \text{and} & x \in \mathbf{E} = \{x : (x - x^0)^T \Gamma (x - x^0) \leqq 1\} \end{array}$$

where Γ is a symmetric PD matrix of order n. So $\mathbf{E}$ is an ellipsoid. Let $\mathbf{H} = \{x : Ax = b\}$. Let F be the Cholesky factor of Γ (i. e., it is the lower triangular matrix

satisfying $FF^T = \Gamma$). To solve this problem, apply the linear transformation that transforms the ellipsoid into a sphere $\mathbf{B}$, this is

$$y = F^T(x - x^0) \quad \text{or}$$
$$x = x^0 + (F^T)^{-1}y \ .$$

This transform the affine space $\mathbf{H}$ into another affine space $\widehat{\mathbf{H}} = \{y : A(F^T)^{-1}y = (b - Ax^0)\}$, the ellipsoid $\mathbf{E}$ into the unit sphere $\mathbf{B} = \{y : ||y|| \leqq 1\}$, and the objective function cx into $c(F^T)^{-1}y + cx^0$. So the transformed problem is :

$$\begin{array}{ll} \text{minimize} & c(F^T)^{-1}y \\ \text{subject to} & y \in \widehat{\mathbf{H}} \cap \mathbf{B} \end{array}$$

which can be solved as discussed above. From the optimum solution y of this problem, we compute the optimum solution x, of the original problem using the equation $x = x^0 + (F^T)^{-1}y$.

5 Converting a Near Optimum Feasible Solution into an Optimum Feasible Solution

Consider the LP

$$\begin{array}{ll} \text{minimize} & z(x) = cx \\ \text{subject to} & Ax = b \\ & x \geqq 0 \end{array} \tag{11.10}$$

Let $\bar{x}$ be a feasible solution for it. A well known result in LP says that if $\bar{x}$ is not a BFS for this problem, then a BFS $\hat{x}$ for it satisfying $c\hat{x} \leqq c\bar{x}$ can be obtained, or it can be established that cx is unbounded below in this problem. See [2.26]. We describe the procedure for doing it here.

Let $\mathbf{J} = \{j : \bar{x}_j > 0\}$. If $\{A_{.j} : j \in \mathbf{J}\}$ is linearly independent, $\bar{x}$ is itself a BFS. If $\bar{x}$ is not BFS, $\{A_{.j} : j \in \mathbf{J}\}$ is linearly dependent. Let a linear dependence relation among these vectors be

$$\sum_{j \in \mathbf{J}} \alpha_j A_{\cdot j} = 0$$

where $(\alpha_j : j \in \mathbf{J}) \neq 0$. Such a vector $(\alpha_j : j \in \mathbf{J})$ can be computed by pivotal methods for checking linear independence of the set $\{A_{.j} : j \in \mathbf{J}\}$, see [2.26].

Since $\bar{x}$ is feasible, and from the definition of $\mathbf{J}$, we also have

$$\sum_{j \in \mathbf{J}} \bar{x}_j A_{.j} = b$$

$$\therefore \quad \sum_{j \in \mathbf{J}} (\bar{x}_j + \theta\alpha_j) A_{.j} = b$$

for all real values of θ. Define the vector $x(\theta)$ by

$$x_j(\theta) = \begin{cases} \bar{x}_j + \theta\alpha_j & \text{for } j \in \mathbf{J} \\ 0 & \text{for } j \notin \mathbf{J} \ . \end{cases}$$

Now define

$$\theta_1 = \begin{cases} -\infty, & \text{if } \alpha_j \leqq 0 \text{ for all } j \in \mathbf{J}; \\ \max\{-\frac{\bar{x}_j}{\alpha_j} : j \in \mathbf{J} \text{ and such that } \alpha_j > 0\}, & \text{otherwise} \end{cases}$$

$$\theta_2 = \begin{cases} +\infty, & \text{if } \alpha_j \geqq 0 \text{ for all } j \in \mathbf{J}; \\ \min\{-\frac{\bar{x}_j}{\alpha_j} : j \in \mathbf{J} \text{ and such that } \alpha_j < 0\}, & \text{otherwise} \ . \end{cases}$$

Clearly $\theta_1 < 0$, $\theta_2 > 0$, and $x(\theta) \geqq 0$ and hence feasible to the LP for all $\theta_1 \leqq \theta \leqq \theta_2$. Since $(\alpha_j : j \in \mathbf{J}) \neq 0$, at least one among θ_1 or θ_2 is finite. If $\sum_{j\in\mathbf{J}} c_j\alpha_j = 0$, let $\gamma = \theta_1$ or θ_2 whichever is finite, break ties arbitrarily.

If $\sum_{j\in\mathbf{J}} c_j\alpha_j > 0$, and $\theta_1 = -\infty$ then $\{x(\theta) : \theta \leqq \theta_2\}$ is a feasible half-line along which cx diverges to $-\infty$. Likewise if $\sum_{j\in\mathbf{J}} c_j\alpha_j < 0$, and $\theta_2 = +\infty$, then $\{x(\theta) : \theta \geqq \theta_1\}$ is a feasible half-line along which cx diverges to $-\infty$. If neither of these unboundedness conditions are satisfied, select $\gamma = \theta_1$ if $\sum_{j\in\mathbf{J}} c_j\alpha_j > 0$, or $\gamma = \theta_2$ if $\sum_{j\in\mathbf{J}} c_j\alpha_j < 0$.

Then $x(\gamma)$ is a feasible solution satisfying $cx(\gamma) \leqq c\bar{x}$, and the number of positive components in $x(\gamma)$ is at least one less than that in $\bar{x}$.

Repeat the same process now with the feasible solution $x(\gamma)$. After at most $|\mathbf{J}|$ of these steps, we will either obtain a BFS $\hat{x}$ satisfying $c\hat{x} \leqq c\bar{x}$, or establish that cx is unbounded below in this LP.

Example 11.1

Consider the following LP

x_1	x_2	x_3	x_4	x_5	x_6	x_7	b
1	0	0	1	0	1	−1	3
0	1	0	0	−1	2	−1	5
0	0	1	−1	1	1	−2	7
−10	4	6	2	4	8	10	$= z(x)$ minimize

$x_j \geqq 0$ for all j

Let $x^0 = (\frac{5}{2}, 6, \frac{13}{2}, \frac{1}{2}, 1, 0, 0)^T$ be the feasible solution with an objective value $z(x^0) = 43$. Denote the coeficient of x_j in $z(x)$ by c_j, and the column vector of x_j in the constraint matrix by $A_{.j}$. $\mathbf{J}$ = the set of subscripts of positive variables in x^0 is $\{1, 2, 3, 4, 5\}$. The set of columns $\{A_{.j} : j = 1 \text{ to } 5\}$ is linearly dependent, and a linear dependence relation among them is

$$-A_{.1} + A_{.3} + A_{.4} = 0$$

So the vector α^0 leading to this linear dependence relation is $(-1,0,1,1,0,0,0)^T$ and $z(\alpha^0) = 18 > 0$. The feasible solution $x^0(\theta)$ constructed in the procedure is

$$x^0(\theta) = \left(\frac{5}{2} - \theta, 6, \frac{13}{2} + \theta, \frac{1}{2} + \theta, 1, 0, 0\right)^T$$

and so $\theta_1 = -\frac{1}{2}$, $\theta_2 = \frac{5}{2}$. Since $z(\alpha^0) > 0$, we choose $\gamma = \theta_1 = -\frac{1}{2}$. The next feasible solution in $x^0(\theta_1) = x^1$ is

$$x^1 = (3,6,6,0,1,0,0)^T$$

It can be verified that $z(x^1) = 34$ and that x^1 has only 4 positive components. Continuing the procedure with x^1, the set of columns to examine is $\{A_{.1}, A_{.2}, A_{.3}, A_{.5}\}$ which again is linearly dependent, with the linear dependence relation

$$A_{.2} - A_{.3} + A_{.5} = 0 \ .$$

The vector α^1 corresponding to this linear dependence relation is $(0,1,-1,0,1,0,0)^T$ and $z(\alpha^1) = 2 > 0$. The feasible solution $x^1(\theta)$ constructed in the procedure is

$$x^1(\theta) = (3, 6+\theta, 6-\theta, 0, 1+\theta, 0, 0)^T$$

and so $\theta_1 = -1$, $\theta_2 = 6$, and since $z(\alpha^1) > 0$, we choose $\alpha = \theta_1 = -1$. The next feasible solution is $x^1(\theta_1) = x^2$, $x^2 = (3,5,6,0,0,0,0)^T$, $z(x^2) = 32$. Now x^2 is a BFS and it satisfies $z(x^2) < z(x^0)$.

Consider the LP (11.10) again. Suppose the data is integer, and L is the size of this LP. Let z^* be the unknown optimum objective value in this LP. If $\bar{x}$ is a feasible solution for this LP whose objective value is sufficiently close to the optimum objective value, e.g. if $c\bar{x}$ is within $2^{-\mathbf{L}}$ of z^*, then the BFS obtained by applying the above procedure beginning with $\bar{x}$, will be an optimum solution for the LP, by the results proved in the ellipsoid algorithm, see Chapter 8 and [2.26] and Figure 11.2. This follows because when L is the size of the LP, any BFS x satisfying, objective value at x, $z(x) \leqq z^* + 2^{-\mathbf{L}}$, has to be an optimum BFS, by the results proved under the ellipsoid algorithm.

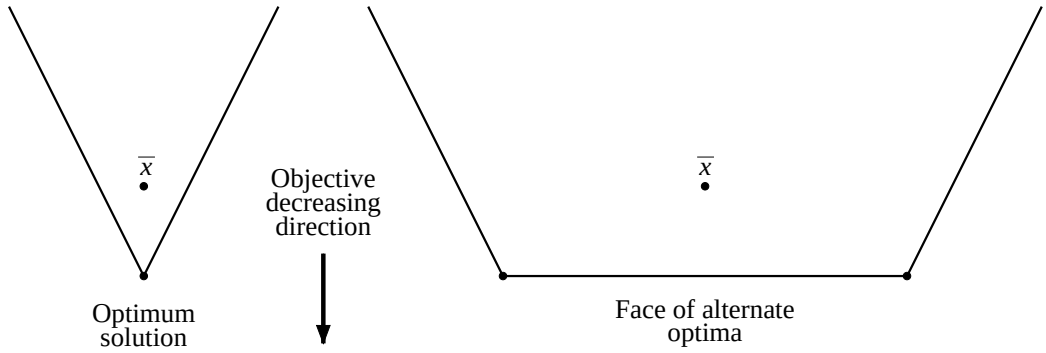


Figure 11.2 If $\bar{x}$ is near optimal, a BFS obtained by above procedure will be optimal, whether problem has unique optimum solution or has alternate optima.

Thus if a near optimal feasible solution with objective value sufficiently close to the optimum can be found, the procedure discussed in this subsection can be used to convert it into an exact optimum solution for the LP. This result is used in Karmarkar's algorithm. Karmarkar's algorithm computes a near optimal solution for an LP and then converts it into an exact optimum solution of the problem using the procedure discussed here.

6 Karmarkar's Algorithm

Consider the LP in the form

$$\begin{array}{ll} \text{minimize} & cx \\ \text{subject to} & x \in \Omega \cap \mathbf{S} \end{array} \tag{11.11}$$

where

$$\Omega = \{x : Ax = 0\}$$

$$\mathbf{S} = \left\{x : x \geqq 0, \quad \sum_{j=1}^{n} x_j = 1\right\}$$

A is of order $m \times n$. Without any loss of generality we assume that the rank of A is m. We make the Following assumptions.

(i) $x^0 = \frac{1}{n}e$, where e is the column vector of all 1's in $\mathbf{R}^n$ is feasible to this LP.

(ii) The optimum objective value in (11.11) is zero.

Karmarkar's algorithm generates a finite sequence of feasible points $x^0, x^1, \ldots$, all of them > 0, such that cx^r is strictly decreasing. L denotes the size of (11.11).

These assumptions also imply that the rank of $\begin{pmatrix} A \\ e^T \end{pmatrix}$ is $m+1$. If $cx^0 = 0$, by the assumptions, x^0 is optimal to (11.11), we terminate. So we assume that $cx^0 > 0$. The method terminates when a feasible solution x^r satisfying $cx^r \leqq 2^{-\mathcal{O}(L)}$ is obtained, and then converts this approximate optimal solution x^r into an exact optimal solution as in Subsection 5.

If c is a linear combination of the rows of A, $cx = 0$ at all feasible solutions x, and so our assumptions imply that c is not a linear combination of the rows of A.

Now we shall describe the general step of the algorithm.

Step $r+1$: Assume we are given $x^r > 0$, $x^r \in \Omega \cap \mathbf{S}$. Let $x^r = a = (a_1, \ldots, a_n)^T$. Let $D = \text{diag}\{a_1, \ldots, a_n\} = (d_{ij})$ with $d_{ii} = a_i$, $i = 1$ to n, and $d_{ij} = 0$ for $i \neq j$. So D is a positive diagonal matrix of order $n \times n$.

We now construct a projective transformation $T : \mathbf{S} \to \mathbf{S}$, which depends on the vector a. For $x \in \mathbf{S}$,

$$T(x) = \frac{D^{-1}x}{e^T D^{-1}x} \ .$$

It can be verified that $T(x) \in \mathbf{S}$ for all $x \in \mathbf{S}$. Also, if $x \in \mathbf{S}$ satisfies $x > 0$, so is $T(x)$. So, the transformation $T(x)$ maps every point in the relative interior of $\mathbf{S}$ (i. e., a point in $\mathbf{S}$ which is > 0) into another point in the relative interior of $\mathbf{S}$. It can be verified that

$$T(a) = a^0 = \frac{1}{n}e \ .$$

If $T(x) = x'$, the inverse transformation yielding $T^{-1}(x') = x$ is

$$T^{-1}(x') = \frac{Dx'}{e^T Dx'} \ .$$

Associate the objective function cx with the potential function $f(x)$ defined over the intersection of Ω with the relative interior of $\mathbf{S}$, given by

$$f(x) = \sum_{j=1}^{n} \log\left(\frac{cx}{x_j}\right)$$

where log denotes the natural logarithm. Since all the points obtained in the algorithm will be strictly positive, they are in the relative interior of $\mathbf{S}$, and $f(x)$ is well defined at them. For x from the relative interior of $\mathbf{S}$ (i. e., $x \in \mathbf{S}$ and $x > 0$) with $T(x) = x'$, define the transformed potential function $f'(x')$ so that it satisfies $f(x) = f'(T(x)) = f'(x')$. Then it can be verified that

$$f'(y) = \sum_{j=1}^{n} \log\left(\frac{\hat{c}y}{y_j}\right) - \sum_{j=1}^{n} \log(a_j)$$

where $\hat{c} = cD$.

Let Ω' denote the tranformation of the subspace Ω under T. Thus

$$\Omega' = \{x' : ADx' = 0\} \ .$$

Now define

$$\Omega'' = \left\{ \begin{array}{r} y : \ ADy = 0 \\ e^T y = 1 \end{array} \right\}$$

$$B = \left(\begin{array}{c} AD \\ e^T \end{array} \right) \ .$$

As discussed earlier, B is of full row rank. Since $a \in \Omega$, we have $ADe = 0$, so $a^0 \in \Omega''$.

Let $\bar{\rho}$, $\underline{\rho}$ be respectively the radii of the largest sphere with center a^0 contained in the simplex $\mathbf{S}$, smallest shpere with center a^0 containing $\mathbf{S}$. See Figure 11.3.

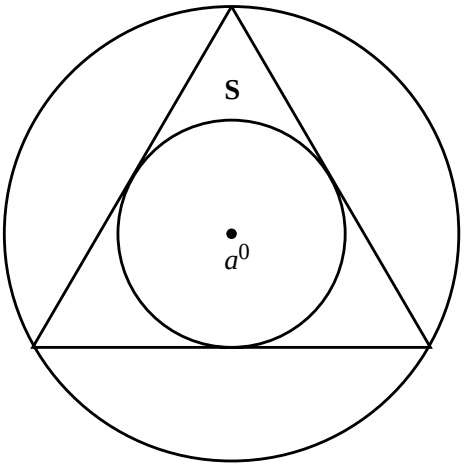


Figure 11.3 Inscribed sphere and circumscribing sphere

Then

$$\underline{\rho} = \frac{1}{\sqrt{n(n-1)}}, \qquad \bar{\rho} = \sqrt{\frac{n-1}{n}} = (n-1)\underline{\rho} \ .$$

For $0 < \alpha < 1$, define

$$\mathbf{B}(a^0, \alpha\underline{\rho}) = \{x : \ ||x - a^0|| \leqq \alpha\underline{\rho}\} \ .$$

Since $0 < \alpha < 1$, the sphere $\{x : \sum_{j=1}^{n} x_j = 1\} \cap \mathbf{B}(a^0, \alpha\underline{\rho}) \subset \mathbf{S}$. The projective transformation $T(x)$, transforms the set of feasible solutions of (11.11) into $\Omega'' \cap \mathbf{S}$. However, $T(x)$ does not transform cx into a linear function. But the potential function $f(x)$, which depends on ratios of linear functions is transformed into another function of the same form, $f'(x')$. We will show later on that a reduction in $f(x)$ leads to a reduction in cx. The problem of minimizing $f(x)$ gets transformed into that of minimizing $f'(x')$. We show later on that minimizing $f'(x')$ can be achieved approximately by optimizing a linear approximation, $\hat{c}x'$.

Instead of optimizing over $\Omega'' \cap \mathbf{S}$ in the transformed problem, we optimize over the simpler subset $\Omega'' \cap \mathbf{B}(a^0, \alpha\underline{\rho})$. The reasons for this are explained below.

Our original problem is transformed into that of optimizing $f'(x')$ over $\Omega' \cap \mathbf{S}$. Since

$$\left\{x : \sum_{j=1}^{n} x_j = 1\right\} \cap \mathbf{B}(a^0, \underline{\rho}) \subset \mathbf{S} \subset \mathbf{B}(a^0, \bar{\rho}) \cap \left\{x : \sum_{j=1}^{n} x_j = 1\right\}$$

$$\begin{array}{ccccc} \text{min value of } f'(x') & & \text{min value of } f'(x') & & \text{min value of } f'(x') \\ \text{over } \Omega'' \cap \mathbf{B}(a^0, \underline{\rho}) & \geqq & \text{over } \Omega'' \cap \mathbf{S} & \geqq & \text{over } \Omega'' \cap \mathbf{B}(a^0, \bar{\rho}) \end{array}$$

Since $\Omega'' \cap \mathbf{B}(a^0, \rho)$ for any ρ is a sphere, optimizing over it is much easier than optimizing over $\Omega'' \cap \mathbf{S}$. To optimize $f'(x')$ over $\Omega'' \cap \mathbf{B}(a^0, \rho)$, we approximate $f'(x')$ by a linear function, $\hat{c}x'$, and the minimization of this linear function over $\Omega'' \cap \mathbf{B}(a^0, \rho)$ can be carried out very easily by the simple techniques discussed in Subsection 4.

If $\gamma_0, \gamma_1, \gamma_2$ denote the minimum value of this linear function $\hat{c}x'$ over $\Omega'' \cap \mathbf{B}(a^0, \underline{\rho})$, $\Omega' \cap \mathbf{S}$, $\Omega'' \cap \mathbf{B}(a^0, \bar{\rho})$ respectively, we have $\gamma_0 \geqq \gamma_1 \geqq \gamma_2$, and so

$$\hat{c}a^0 - \gamma_0 \leqq \hat{c}a^0 - \gamma_1 \leqq \hat{c}a^0 - \gamma_2 = \left(\frac{\bar{\rho}}{\underline{\rho}}\right)(\hat{c}a^0 - \gamma^0)$$

the last equation follows from the results in Subsection 4. So

$$\frac{\hat{c}a^0 - \gamma_0}{\hat{c}a^0 - \gamma_1} \geqq \frac{1}{n-1}$$
$$\frac{\gamma_0 - \gamma_1}{\hat{c}a^0 - \gamma_1} \leqq 1 - \frac{1}{n-1}$$

So by going from the point a^0 to the point that minimizes $\hat{c}x'$ over $x' \in \Omega'' \cap \mathbf{B}(a^0, \underline{\rho})$, we come closer to the minimum value of the objective function by a factor of $\left(1 - \frac{1}{n-1}\right)$.

In practice, we optimize over a smaller subset $\Omega'' \cap \mathbf{B}(a^0, \alpha\underline{\rho})$ for $0 < \alpha < 1$ for the following reasons.

a) it allows for optimization of $f'(x')$ to be approximated closely by optimization of a linear function.
b) Under finite precision or other approximate arithmetic, it provides us a margin to absorb errors without going outside the simplex.

See Figure 11.4. The choice of $\alpha = \frac{1}{4}$ works (this leads to the factor δ discussed later on in Theorem 11.4 to be $> \frac{1}{32}$). In practical implementation, one may want to choose a value of α much closer to 1 for rapid convergence.

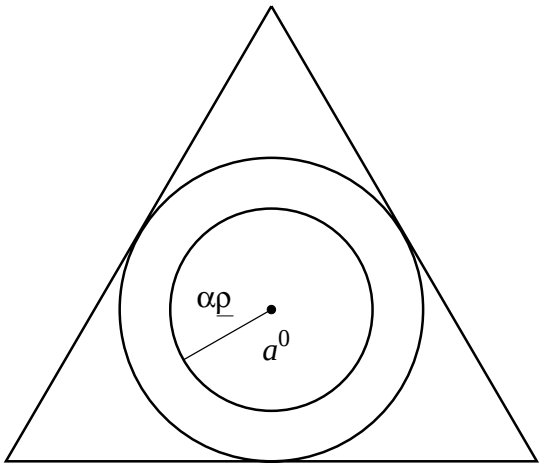


Figure 11.4 The simplex $\mathbf{S}$, and the inscribed sphere $\mathbf{B}(a^0, \alpha\underline{\rho}) \cap \{x : \sum_{j=1}^{n} x_j = 1\}$ inside it, for $0 < \alpha < 1$.

Since $\mathbf{B}(a^0, \alpha\underline{\rho})$ is a sphere with center a^0, and Ω'' is an affine space containing the point a^0, the intersection $\Omega'' \cap \mathbf{B}(a^0, \alpha\underline{\rho})$ is a lower dimensional sphere. As discussed in Subsection 4 above, minimizing a linear function over the lower dimensional sphere $\Omega'' \cap \mathbf{B}(a^0, \alpha\underline{\rho})$ requires taking a step from the center a^0, in the direction of the negative gradient, with step length equal to the radius of the sphere, in the affine space Ω''. We provide the details of this algorithm.

Subroutine to minimize $\hat{c}x'$ Over $x' \in \Omega'' \cap \mathbf{B}(a^0, \alpha\underline{\rho})$

First project $\hat{c}$ orthogonally onto the subspace $\{y : By = 0\}$. This yields

$$\hat{c}_p = \hat{c}\left[I - B^T(BB^T)^{-1}B\right]$$

If $\hat{c}_p = 0$, the objective function will have the same value at all feasible solutions, contradicting our assumptions. So $\hat{c}_p \neq 0$. Let

$$\tilde{c}_p = \frac{\hat{c}_p}{||\hat{c}_p||}$$
$$g' = a^0 - \alpha\underline{\rho}\tilde{c}_p$$

Then g' is the point which minimizes $\hat{c}x'$ over $x' \in \Omega'' \cap \mathbf{B}(a^0, \alpha\underline{\rho})$. We will prove this in Theorem 11.1 given below.

Now define

$$x^{r+1} = T^{-1}(g') = \frac{Dg'}{e^T Dg'} \ .$$

If $cx^{r+1} = 0$, x^{r+1} is optimal to (11.11), terminate. If $cx^{r+1} > 0$ but sufficiently small (i. e., $cx^{r+1} \leqq 2^{-\mathcal{O}(L)}$) terminate with the conclusion that x^{r+1} is near optimal to (11.11) and convert it into an exact optimal solution as in Subsection 5. If these conditions are not satisfied, go to the next step.

Proof of the Algorithm and its Polynomial Boundedness

Theorem 11.1 *The vector g' minimizes $\hat{c}x'$ over $x' \in \Omega'' \cap \mathbf{B}(a^0, \alpha\underline{\rho})$.*

Proof. Let $z \in \Omega'' \cap \mathbf{B}(a^0, \alpha\underline{\rho})$. Since Ω'' is an affine space and both g', $z \in \Omega''$, we have $B(g' - z) = 0$. So, $B^T(BB^T)^{-1}B(g' - z) = 0$. Therefore $(\hat{c} - \hat{c}_p)(g' - z) = 0$. Thus $\hat{c}(g' - z) = \hat{c}_p(g' - z) = ||\hat{c}_p||\tilde{c}_p(a^0 - \alpha\underline{\rho}\tilde{c}_p^T - z) = ||\hat{c}_p||\big(\tilde{c}_p(a^0 - z) - \alpha\underline{\rho}\big)$ (since $\tilde{c}_p\tilde{c}_p^T = ||\tilde{c}_p|| = 1$). But, $\tilde{c}_p(a^0 - z) \leqq ||\tilde{c}_p||\ ||a^0 - z||$ (by Cauchy-Schwartz inequality) $= ||a^0 - z||$ (since $||\tilde{c}_p|| = 1$) $\leqq \alpha\underline{\rho}$, since $z \in \mathbf{B}(a^0, \alpha\underline{\rho})$. Therefore $\tilde{c}_p(a^0 - z) - \alpha\underline{\rho} \leqq 0$, and therefore by the above $\hat{c}(g' - z) \leqq 0$. Hence, $\hat{c}g' \leqq \hat{c}z$ for all $z \in \Omega'' \cap \mathbf{B}(a^0, \alpha\underline{\rho})$, that is, g' minimizes $\hat{c}x'$ over $x' \in \Omega'' \cap \mathbf{B}(a^0, \alpha\underline{\rho})$.

□

Theorem 11.2 *There exists a point $\bar{x} \in \Omega'' \cap \mathbf{B}(a^0, \alpha\underline{\rho})$ such that*

$$\begin{array}{lll} \textit{either} & \textit{(i)} & \hat{c}\bar{x} = 0 \\ \textit{or} & \textit{(ii)} & f'(\bar{x}) \leqq f'(a_0) - \delta \end{array}$$

where δ is a positive constant depending on α.

Proof. Let x^* minimize cx over $\Omega \cap \mathbf{S}$. By hypothesis $cx^* = 0$. Define $\xi = \frac{D^{-1}x^*}{e^T D^{-1}x^*}$.

Case 1 : $\xi \in \mathbf{B}(a^0, \alpha\underline{\rho})$. In this case let $\bar{x} = \xi$. Then $\bar{x} \in \Omega'' \cap \mathbf{B}(a^0, \alpha\underline{\rho})$ and $\hat{c}\bar{x} = 0$, so (i) is satisfied.

Case 2 : $\xi \notin \mathbf{B}(a^0, \alpha\underline{\rho})$. In this case, let $\bar{x}$ be the point at which the line segment joining a^0 with ξ intersects the boundary of the sphere $\mathbf{B}(a^0, \alpha\underline{\rho})$. Then $\bar{x} = (1-\lambda)a^0 + \lambda\xi$ for some $0 < \lambda < 1$. Since a^0 and ξ are in Ω'', so is $\bar{x}$. So $\bar{x} \in \Omega'' \cap \mathbf{B}(a^0, \alpha\underline{\rho})$, and $\hat{c}\bar{x} = (1-\lambda)\hat{c}a^0 + \lambda\hat{c}\xi = (1-\lambda)\hat{c}a^0$ (since $\hat{c}\xi = cD\xi = 0$ because $cx^* = 0$). So

$$\frac{\hat{c}a^0}{\hat{c}\bar{x}} = \frac{1}{1-\lambda} \tag{11.12}$$

Now

$$\begin{aligned}
f'(a^0) - f'(\bar{x}) &= \sum_{j=1}^{n} \log\left(\frac{\hat{c}a^0}{a^0_j}\right) - \sum_{j=1}^{n} \log\left(\frac{\hat{c}\bar{x}}{\bar{x}_j}\right) \\
&= \sum_{j=1}^{n} \log\left(\left(\frac{\hat{c}a^0}{\hat{c}\bar{x}}\right)\left(\frac{\bar{x}_j}{a^0_j}\right)\right) \\
&= \sum_{j=1}^{n} \log\left(\frac{\bar{x}_j}{(1-\lambda)a^0_j}\right) \qquad \text{by (11.12)} \\
&= \sum_{j=1}^{n} \log\left(\frac{(1-\lambda)a^0_j + \lambda\xi_j}{(1-\lambda)a^0_j}\right) \\
&= \sum_{j=1}^{n} \log\left(1 + \left(\frac{\lambda}{1-\lambda}\right)\left(\frac{\xi_j}{a^0_j}\right)\right)
\end{aligned}$$

It can easily be verified that if $\gamma_i \geqq 0$ for all i, then the product $\prod_i(1+\gamma_i) \geqq 1 + \sum_i \gamma_i$. Taking logs on both sides we have $\sum_i \log(1+\gamma_i) \geqq \log(1 + \sum_i \gamma_i)$. Applying this to the above, we have

$$\begin{aligned}
f'(a^0) - f'(\bar{x}) &\geqq \log\left(1 + \left(\tfrac{\lambda}{1-\lambda}\right)\frac{\sum_{j=1}^{n}\xi_j}{(1/n)}\right), \text{ since } a^0_j = 1/n \text{ for all } j \\
&\geqq \log\left(1 + \tfrac{n\lambda}{1-\lambda}\right), \qquad \text{since } \textstyle\sum_{j=1}^{n}\xi_j = 1
\end{aligned}$$

Now, $\bar{x} = (1-\lambda)a^0 + \lambda\xi$. So $\bar{x} - a^0 = \lambda(\xi - a^0)$. Since $\bar{x}$ is on the boundary of the sphere $\mathbf{B}(a^0, \alpha\underline{\rho})$, we have $||\bar{x} - a^0|| = \alpha\underline{\rho}$, so from the above $\alpha\underline{\rho} = ||\bar{x} - a^0|| = \lambda||\xi - a^0|| \leqq \lambda\bar{\rho}$. So $\lambda \geqq \frac{(\alpha\underline{\rho})}{\bar{\rho}} = \frac{\alpha}{n-1}$. So

$$1 + \frac{n\lambda}{1-\lambda} \geqq 1 + \frac{n\left(\frac{\alpha}{n-1}\right)}{1 - \frac{\alpha}{n-1}} = 1 + \frac{n\alpha}{n-1-\alpha} \geqq 1 + \alpha$$

Therefore, from the above

$$f'(a^0) - f'(\bar{x}) \geqq \log(1+\alpha)$$

Thus taking $\delta = \log(1+\alpha)$ establishes (ii).

□

Lemma 11.1 *Let γ be a real number. If $|\gamma| \leqq \beta < 1$ then $|\log(1+\gamma) - \gamma| \leqq \frac{\gamma^2}{2(1-\beta)^2}$.*

Proof. Let $\psi(\gamma) = \log(1+\gamma)$. Then

$$\frac{d}{d\gamma}\psi(\gamma) = \frac{1}{1+\gamma}, \text{ and } \frac{d^2}{d\gamma^2}\psi(\gamma) = \frac{-1}{(1+\gamma)^2} \ .$$

By the mean value theorem of calculus applied to the function $\log(1+\gamma)$, we have

$$\log(1+\gamma) = \log(1) + \gamma\Big(\frac{d}{d\gamma}\psi(\gamma)\Big)_{\gamma=0} + \frac{\gamma^2}{2}\Big(\frac{d^2}{d\gamma^2}\psi(\bar{\gamma})\Big)$$

for some $\bar{\gamma}$ satisfying $|\bar{\gamma}| \leqq |\gamma|$. So

$$\log(1+\gamma) = \gamma - \frac{\gamma^2}{2}\Big(\frac{1}{(1+\bar{\gamma})^2}\Big)$$

$$|\log(1+\gamma) - \gamma| = \frac{\gamma^2}{2}\Big(\frac{1}{(1+\bar{\gamma})^2}\Big) \leqq \frac{\gamma^2}{2(1-\beta)^2}$$

□

Lemma 11.2 *Let $\beta = \alpha\sqrt{\frac{n}{n-1}}$. Then*

$$\left|\sum_{j=1}^{n} \log\Big(\frac{x_j}{a_j^o}\Big)\right| \leqq \frac{\beta^2}{2(1-\beta)^2} \quad \textit{for all } x \in \mathbf{B}(a^0, \alpha\underline{\rho}) \cap \mathbf{S} \ .$$

Proof. Let $x \in \mathbf{B}(a^0, \alpha\underline{\rho}) \cap \mathbf{S}$. Then $||x - a^0||^2 \leqq \alpha^2\underline{\rho}^2$. So (since $a_j^0 = \frac{1}{n}$ for all j)

$$\sum_{j=1}^{n}\Big(\frac{x_j - a_j^0}{a_j^0}\Big)^2 \leqq \frac{\alpha^2\underline{\rho}^2}{(1/n)^2} = \alpha^2\underline{\rho}^2 n^2 = \frac{\alpha^2 n^2}{n(n-1)} = \beta^2$$

So, $\left\|\frac{x_j - a_j^0}{a_j^0}\right\| \leqq \beta$ for all j. Therefore, by Lemma 11.1

$$\left|\log\Big(1 + \frac{x_j - a_j^0}{a_j^0}\Big) - \Big(\frac{x_j - a_j^0}{a_j^0}\Big)\right| \leqq \Big(\frac{x_j - a_j^0}{a_j^0}\Big)^2\Big(\frac{1}{2(1-\beta)^2}\Big)$$

$$\therefore \quad \left|\sum_{j=1}^{n}\log\Big(\frac{x_j}{a_j^0}\Big) - \sum_{j=1}^{n}\Big(\frac{x_j - a_j^0}{a_j^0}\Big)\right|$$

$$\leqq \Big(\frac{1}{2(1-\beta)^2}\Big)\left(\sum_{j=1}^{n}\Big(\frac{x_j - a_j^0}{a_j^0}\Big)^2\right)$$

$$\leqq \frac{\beta^2}{2(1-\beta)^2}$$

This implies that $\left|\sum_{j=1}^{n}\log\big(\frac{x_j}{a_j^0}\big)\right| \leqq \frac{\beta^2}{2(1-\beta)^2}$, since $\sum_{j=1}^{n}\Big(\frac{x_j - a_j^0}{a_j^0}\Big) = \frac{1}{n}\left(\sum_{j=1}^{n}(x_j - a_j^0)\right) = 0$ (as x and $a^0 \in \mathbf{S}$).

□

Theorem 11.3 *The point g' which minimimizes $\hat{c}x'$ over $x' \in \Omega'' \cap \mathbf{B}(a^0, \alpha\underline{\rho})$ satisfies*

$$\begin{array}{rrll} \textit{either} & \textit{(i)} & \hat{c}g' & = 0 \\ \textit{or} & \textit{(ii)} & f'(g') & \leqq f'(a^0) - \delta \end{array}$$

where δ is a constant depending on α. If $\alpha = \frac{1}{4}$, $\delta \geqq \frac{1}{32}$.

Proof. Define

$$\tilde{f}(x) = n \log\Big(\frac{\hat{c}x}{\hat{c}a^0}\Big) \ .$$

Let h be the point where $f'(x')$ achieves its minimum value over $x' \in \Omega'' \cap \mathbf{B}(a^0, \alpha\underline{\rho})$. Then,

$$\begin{aligned} f'(a^0) - f'(g') =& f'(a^0) - f'(h) + f'(h) - f'(g') \\ =& [f'(a^0) - f'(h)] + \big[f'(h) - \big(f'(a^0) + \tilde{f}(h)\big)\big] \\ & - \big[f'(g') - \big(f'(a^0) + \tilde{f}(g')\big)\big] + \big[\tilde{f}(h) - \tilde{f}(g')\big] \end{aligned} \tag{11.13}$$

Now if the minimum value of $\hat{c}x'$ over $x' \in \Omega'' \cap \mathbf{B}(a^0, \alpha\underline{\rho})$ is zero, condition (i) of the theorem holds trivially. Let us assume that this is not the case. Then by Theorem 11.2

$$f'(a^0) - f'(h) \geqq \log(1+\alpha) \ . \tag{11.14}$$

For $x' \in \mathbf{B}(a^0, \alpha\underline{\rho}) \cap \Omega''$, we have

$$\begin{aligned} f'(x') - \big(f'(a^0) + \tilde{f}(x')\big) &= \sum_{j=1}^{n} \log\Big(\frac{\hat{c}x'}{x'_j}\Big) - \sum_{j=1}^{n} \log\Big(\frac{\hat{c}a^0}{a^0_j}\Big) - n \log\Big(\frac{\hat{c}x'}{\hat{c}a^0}\Big) \\ &= -\sum_{j=1}^{n} \log\Big(\frac{x'_j}{a^0_j}\Big) \end{aligned}$$

So

$$\begin{aligned} \big|f'(x') - \big(f'(a^0) + \tilde{f}(x')\big)\big| &= \Big|\sum_{j=1}^{n} \log\Big(\frac{x'_j}{a^0_j}\Big)\Big| \\ &\leqq \frac{\beta^{\mathbf{2}}}{2(1-\beta)^{\mathbf{2}}} \quad \text{by Lemma 11.2} \end{aligned} \tag{11.15}$$

But $\tilde{f}(x')$ depends on $\hat{c}x'$ in a monotonically increasing manner. So $\tilde{f}(x')$ and $\hat{c}x'$ attain their minimum value over $x' \in \Omega'' \cap \mathbf{B}(a^0, \alpha\underline{\rho})$ at the same point, that is g'. So

$$\tilde{f}(h) \geqq \tilde{f}(g') \ . \tag{11.16}$$

Now from (11.15) we have, for $x' \in \Omega'' \cap \mathbf{B}(a^0, \alpha\underline{\rho})$,

$$f'(x') - \big(f'(a^0) + \tilde{f}(x')\big) \geqq -\frac{\beta^{\mathbf{2}}}{2(1-\beta)^{\mathbf{2}}} \ . \tag{11.17}$$

Also $\Omega'' \subset \Omega'$. So both h and $g' \in \Omega'$. From (11.13), (11.14), (11.15), (11.16), (11.17), we have

$$f'(a^0) - f'(g') \geqq \log(1+\alpha) - \frac{\beta^{\mathbf{2}}}{(1-\beta)^{\mathbf{2}}} \ .$$

We know that $\log(1+\alpha) \geqq \alpha - \frac{\alpha^{\mathbf{2}}}{2}$, for $0 < \alpha < 1$. Also

$$\frac{\beta^{\mathbf{2}}}{(1-\beta)^{\mathbf{2}}} = \frac{\alpha^{\mathbf{2}} n}{(n-1)\left(1 - \alpha\left(\frac{n}{n-1}\right)^{\frac{1}{\mathbf{2}}}\right)^{\mathbf{2}}}$$

So from the above, we have

$$f'(a^0) - f'(g') \geqq \delta(n) = \alpha - \frac{\alpha^{\mathbf{2}}}{2} - \frac{\alpha^{\mathbf{2}} n}{(n-1)\left(1 - \alpha\left(\frac{n}{n-1}\right)^{\frac{1}{\mathbf{2}}}\right)^{\mathbf{2}}}$$

As $n \to \infty$, $\delta(n) \to \alpha - \frac{\alpha^{\mathbf{2}}}{2} - \frac{\alpha^{\mathbf{2}}}{(1-\alpha)^{\mathbf{2}}}$. If $n \geqq 4$, $\alpha = \frac{1}{4}$, we have $\delta(n) \geqq \frac{1}{32}$.

□

Theorem 11.4 *Either $cx^{r+1} = 0$, or $f(x^{r+1}) \leqq f(x^r) - \delta$, where δ is a constant depending only on α, as in Theorem 11.3.*

Proof. We have proved in Theorem 11.3 that either $\hat{c}g' = 0$, or $f'(g') \leqq f'(a^0) - \delta$. Now

$$\begin{aligned} x^r &= T^{-1}(a^0) \\ x^{r+1} &= T^{-1}(g') \\ f'(T(x)) &= f(x) \text{ for all } x \in \mathbf{S} \ . \end{aligned}$$

So, by applying T^{-1}, we have from the above, that either $cx^{r+1} = 0$, or $f(x^{r+1}) \leqq f(x^r) - \delta$.

□

Theorem 11.5 *In $\mathcal{O}\big(n(l + \log n)\big)$ steps, the algorithm finds a feasible point $x > 0$ such that*

$$\begin{aligned} &\textit{either} \quad cx \ = 0 \\ &\textit{or} \quad \tfrac{cx}{ca^0} \leqq 2^{-l} \end{aligned}$$

Proof. Suppose $cx^r = 0$ did not occur in the first N steps. Then, by Theorem 11.4

$$\begin{aligned} f(x^r) &\leqq f(x^{r-1}) - \delta, \ \text{ for } r = 1 \text{ to } N \\ \therefore \ f(x^r) &\leqq f(x^0) - r\delta \end{aligned}$$

$$\begin{aligned} \therefore \quad \sum_{j=1}^{n} \log\left(\tfrac{cx^r}{x_j^r}\right) &\leqq \sum_{j=1}^{n} \log\left(\tfrac{ca^0}{a_j^0}\right) - r\delta \\ \text{i. e.,} \quad n \log\left(\tfrac{cx^r}{ca^0}\right) &\leqq \sum_{j=1}^{n} \log(x_j^r) - \sum_{j=1}^{n} \log(a_j^0) - r\delta \\ &\leqq n \log(n) - r\delta, \text{ since } x_j^r \leqq 1 \text{ and } a_j^0 = \tfrac{1}{n} \text{ for all } j \\ \therefore \quad \log\left(\tfrac{cx^r}{ca^0}\right) &\leqq \log n - \tfrac{r\delta}{n} \ . \end{aligned}$$

So if $r = \left\lceil \frac{n}{\delta}(l + \log n) \right\rceil$, we have

$$\log\left(\frac{cx^r}{ca^0}\right) \leqq -l$$

$$\text{i. e., } \left(\frac{cx^r}{ca^0}\right) \leqq 2^{-l}$$

The computation in each step involves $\mathcal{O}(n^3)$ arithmetic operations on the data in the worst case. By Theorem 11.5 and the termination conditions used in the algorithm it has to run for at most $\mathcal{O}(nL)$ steps, to come within $2^{-\mathcal{O}(L)}$ of the optimum, at which point we round the solution to get an exact optimum solution as discussed in Subsection 5. So, the algorithm needs at most $\mathcal{O}(n^4L)$ arithmetical operations on the data in the worst case, it is clearly polynomially bounded.

The final operation of converting the near optimal solution obtained at the termination of the algorithm into an exact optimal solution as discussed in Subsection 5 could be computationally expensive (it may need up to $\mathcal{O}(n)$ pivot steps). In most practical applications the data usually consists of unknown error terms and it makes sense to take the near optimal solution as it is, without the expensive final conversion. In practical LP applications, because of unknown errors in the data, a near optimal and aproximately feasible solution to the model is the usual goal, and Karmarkar's algorithm is well suited to achieve this goal.

7 Efficient Implementation of the Karmarkar Algorithm

The major piece of computation in each step of the algorithm is the computation of the projection $\hat{c}_p = \hat{c}[I - B^T(BB^T)^{-1}B]$. For this we have to find the inverse, $(BB^T)^{-1}$. Since $B = \begin{pmatrix} AD \\ e \end{pmatrix}$, we have

$$BB^T = \begin{pmatrix} AD^2A^T & ADe \\ (ADe)^T & e^TDe \end{pmatrix} = \begin{pmatrix} AD^2A^T & 0 \\ 0 & 1 \end{pmatrix}$$

since the point a used in defining the diagonal matrix D is in $\mathbf{S}$, and $a^0 = \frac{e}{n} \in \Omega''$. $(BB^T)^{-1}$ can be found efficiently if $(AD^2A^T)^{-1}$ can be. The only thing that changes in AD^2A^T from step to step is the diagonal matrix D. Let $D_r = \text{diag}(d_{11}^r, \ldots, d_{nn}^r)$ denote the diagonal matrix D in step r. We do not compute $(AD^2A^T)^{-1}$ in each step from scratch. Instead we update it to the extent necessary as we move from one step to the next.

If D_r and D_{r+1} differ in only one entry, the inverse of $AD_{r+1}^2A^T$ can be computed in $\mathcal{O}(n^2)$ arithmetic operations from $AD_r^2A^T$. For this, consider a nonsingular square matrix M of order n, $u = (u_1, \ldots, u_n)^T$, $v = (v_1, \ldots, v_n)^T$. Then the Sherman-Morrison formula states that

$$(M + uv^T)^{-1} = M^{-1} - \frac{(M^{-1}u)(M^{-1}v)^T}{1 + u^TM^{-1}v}$$

uv^T is a rank-one modification of M, and the formula shows that computation of $(M+uv^T)^{-1}$ can be done with $\mathcal{O}(n^2)$ arithmetical operations given M^{-1}. If D_r and D_{r+1} differ in only the i^{th} diagonal entry, then

$$AD_{r+1}A^T = AD_r^2A^T + \left(\left(d_{ii}^{r+1}\right)^2 - \left(d_{ii}^r\right)^2\right) A_{.i}(A_{.i})^T \ .$$

So, in this case $AD_{r+1}^2A^T$ is obtained from a rank-one modification of $AD_r^2A^T$, and the above formula can be used to get $(AD_{r+1}^2A^T)^{-1}$ from $(AD_r^2A^T)^{-1}$ with $\mathcal{O}(n^2)$ arithmetical operations. If D_r and D_{r+1} differ in t diagonal entries, we can perform t successive rank-one updates as above and obtain $(AD_{r+1}^2A^T)^{-1}$ from $(AD_r^2A^T)^{-1}$ with $\mathcal{O}(n^2t)$ arithmetical operations.

We now show that with a simple modification of the algorithm, we get a version in which $(AD_r^2A^T)^{-1}$ can be used in place of $(AD_{r+1}^2A^T)^{-1}$ as long as D_r and D_{r+1} are close in some sense.

We define the diagonal matrix $\overline{D} = \text{diag}(\overline{d}_{11}, \ldots, \overline{d}_{nn})$ as an approximation to $D_{r+1} = \text{diag}\ (d_{11}^{r+1}, \ldots, d_{nn}^{r+1})$ if

$$\frac{1}{2} \leq \left(\frac{\bar{d}_{ii}}{d_{ii}^{r+1}}\right)^2 \leqq 2 \text{ for all } i \ .$$

We will now analyse the effect of replacing D_{r+1} by such a $\overline{D}$. Consider the following modification of the optimization problem over the inscribed sphere in the transformed space.

$$\begin{array}{lll} \text{minimize} & \hat{c}x' & \\ \text{subject to} & x' \in \Omega'' & \\ \text{and} & h(x') = (x'-a^0)^TQ(x'-a^o) \leqq \bar{\alpha}\underline{\rho} & \end{array} \tag{11.18}$$

where Q is some positive diagonal matrix. Taking $Q = I$ and $\bar{\alpha} = \alpha$ corresponds to the original problem used in Subsection 6.

Letting the row vector π, and scalar μ to be the Lagrange multipliers for (11.18), the KKT conditions for (11.18) imply

$$\hat{c} - \pi B + 2\mu(x'-a^0)^TQ = 0$$

$$\begin{aligned} \therefore \quad \hat{c}Q^{-1}B^T &= \pi BQ^{-1}B^T - 2\mu(x'-a^0)^TB^T \\ &= \pi BQ^{-1}B^T \end{aligned}$$

since both $x', a^0 \in \Omega''$ implies that $B(x'-a^0) = 0$. Using this we conclude that the optimum solution of (11.18), x', satisfies $(x'-a^0)^T = \gamma\hat{c}\big(I - Q^{-1}B^T(BQ^{-1}B^T)^{-1}B\big)Q^{-1}$ where γ is a positive scalar to be determined so that x' satisfies $(x'-a^0)^TQ(x'-a^0) = \bar{\alpha}\underline{\rho}$. Computation of this requires $(BQ^{-1}B^T)^{-1}$. Substituting $B = \begin{pmatrix} AD_{r+1} \\ e^T \end{pmatrix}$ we get

$$BQ^{-1}B^T = \begin{pmatrix} AD_{r+1}Q^{-1}D_{r+1}A^T & AD_{r+1}Q^{-1}e \\ AD_{r+1}Q^{-1}e^T & e^TQ^{-1}e \end{pmatrix} \ .$$

If the inverse of $AD_{r+1}Q^{-1}D_{r+1}A^T$ is known, $(BQ^{-1}B^T)^{-1}$ can be computed with $\mathcal{O}(n^{\mathbf{2}})$ arithmetical operations using the formula

$$\begin{pmatrix} M & p \\ p^T & q \end{pmatrix} = \frac{1}{q - p^T M^{-1} p} \begin{pmatrix} (q - p^T M^{-1} p) M^{-1} + (M^{-1}p)(M^{-1}p)^T & \vdots & -M^{-1}p \\ -(M^{-1}p)^T & \vdots & 1 \end{pmatrix}$$

Suppose $\overline{D} = D_{r+1}E$ where E is a diagonal error matrix such that $E = (e_{ij})$ with $\frac{1}{2} \leqq e_{ii}^{\mathbf{2}} \leqq 2$ for all i, and we know $(A\overline{D}A^T)^{-1}$. Then setting $Q = E^{-\mathbf{2}}$, we have $AD_{r+1}Q^{-1}D_{r+1}A^T = A\overline{D}A^T$. So using the known $(A\overline{D}A^T)^{-1}$, we can compute the optimum solution of the modified problem (11.18) using the above formulae.

Now we relate the solution of (11.18) to the main optimization problem. Since $Q_{ii} = e_{ii}^{-\mathbf{2}} \in \left[\frac{1}{2}, 2\right]$, we have

$$\frac{1}{2}(x' - a^0)^T(x' - a^0) \leqq (x' - a^0)^T Q (x' - a^0) \leqq 2(x' - a^0)^T(x' - a^0)$$

$$\begin{aligned} \mathbf{B}\Big(a^0, (\frac{\bar{\alpha}}{2})\underline{\rho}\Big) &\subseteq \{x' : (x' - a^0)^T Q (x' - a^0) \leqq \bar{\alpha}\underline{\rho}\} \\ &\subseteq \mathbf{B}(a^0, 2\bar{\alpha}\underline{\rho}) \ . \end{aligned}$$

Take $\bar{\alpha} = \frac{\alpha}{2}$ where α is the quantity used in Subsection 6 (there, we used typically $\alpha = \frac{1}{4}$). So

$$\begin{aligned} \mathbf{B}\Big(a^0, (\frac{\alpha}{4})\underline{\rho}\Big) \cap \Omega'' &\subseteq \{x' : x' \in \Omega'' \text{ and } (x' - a^0)^T Q (x' - a^0) \leqq \bar{\alpha}\underline{\rho}\} \\ &\subseteq \mathbf{B}(a^0, \alpha\underline{\rho}) \cap \Omega'' \ . \end{aligned}$$

From the first inclusion we have

$$\begin{array}{l} \text{minimum value of } f'(x') \\ \text{subject to } x' \in \Omega'' \\ \text{and } (x' - a^0)^T Q (x' - a^0) \leqq \bar{\alpha}\underline{\rho} \end{array} \qquad \leqq \qquad \begin{array}{l} \text{minimum value of } f'(x') \\ \text{subject to } x' \in \Omega'' \cap \mathbf{B}\Big(a^0, (\frac{\alpha}{4})\underline{\rho}\Big) \end{array}$$

and by Theorem 11.2 we have

$$\begin{array}{l} \text{minimum value of } f'(x') \\ \text{subject to } x' \in \Omega'' \cap \mathbf{B}\Big(a^0, (\frac{\alpha}{4})\underline{\rho}\Big) \end{array} \qquad \leqq \qquad f'(a^0) - \log\big(1 + \frac{\alpha}{4}\big) \ .$$

So, for $\bar{g}'$, the optimum solution corresponding to the modified problem (11.18), we can claim

$$f'(\bar{g}') \leqq f'(a^0) - \log\big(1 + \frac{\alpha}{4}\big)$$

and if we define $\bar{x}^{r+1} = T^{-1}(\bar{g}')$, we can as in Theorem 11.4, claim

$$f(\bar{x}^{r+1}) \leqq f(x^r) - \bar{\delta}$$

where $\bar{\delta}$ is redefined as

$$\bar{\delta} = \log\big(1 + \frac{\alpha}{4}\big) - \frac{\beta^{\mathbf{2}}}{2(1-\beta)^{\mathbf{2}}} \ .$$

This affects the number of steps by only a constant factor and the algorithm still works.

So, this is what we do, to implement the modified algorithm in an efficient manner. We maintain $(AD^{\mathbf{2}}A^T)^{-1}$. We do not change all diagonal elements of D in each step. Let $y = (y_1, \ldots, y_n)^T$ be the new solution at the end of a step. It is time to update $(AD^{\mathbf{2}}A^T)^{-1}$. Before, we defined the new D to be diag $(y_1, \ldots, y_n)$. Instead, we modify D in two stages.

Compute $\sigma = \frac{1}{n}\sum_{j=1}^{n} \frac{y_j}{d_{jj}}$ where d_{jj} are the diagonal entries in the current D. First multiply D by σ, this needs dividing $(AD^{\mathbf{2}}A^T)^{-1}$ by $\sigma^{\mathbf{2}}$ to update it accordingly. This completes stage 1.

Then, for each $j = 1$ to n, if in the matrix D at the end of stage 1, $\left(\frac{d_{jj}}{y_j}\right)^{\mathbf{2}} \notin \left[\frac{1}{2}, 2\right]$, reset $d_{jj} = y_j$ and update $(AD^{\mathbf{2}}A^T)^{-1}$ corresponding to this change by a rank-one modification as discussed above.

In essence, we carry out fewer updating operations by optimizing (after the projective transformation) over an inscribed ellipsoid (dashed in Figure 11.5) and not the inscribed sphere. (Of course we do not optimize over this sphere or ellipsoid exactly, but scale it by α or $\bar{\alpha}$ before the optimization.) We make enough updating operations to make sure that the current D matrix and current solution y always satisfy $\left(\frac{d_{jj}}{y_j}\right)^{\mathbf{2}} \in \left[\frac{1}{2}, 2\right]$, this insures that the ellipsoid is close to the inscribed sphere

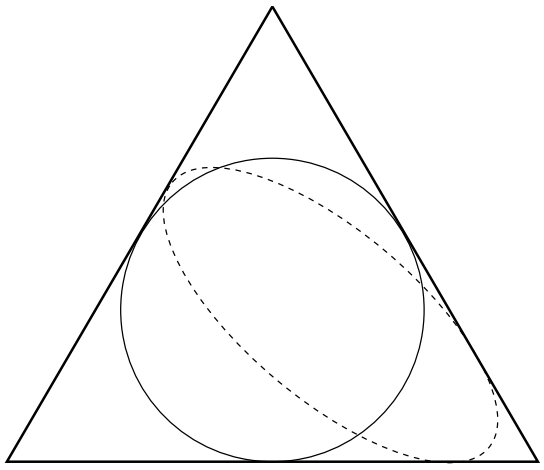


Figure 11.5

We still need only $\mathcal{O}(nL)$ steps to shrink the objective value by the required factor of $2^{-\mathcal{O}(L)}$. With this modification, N. Karmarkar has shown in [11.2] that we need to do only $\mathcal{O}(n^{\frac{\mathbf{3}}{\mathbf{2}}}L)$ updating operations. Since each updating operation requires $\mathcal{O}(n^{\mathbf{2}})$ arithmetic operations on the data, the overall algorithm needs $\mathcal{O}(n^{\mathbf{3.5}}L)$ arithmetic operations on the data in the worst case, with this modification.

8 The Sliding Objective Function Method

From Subsection 6, it is clear that Karmarkar's algorithm solves LPs for which the optimum objective value is known to be zero. As shown in Subsection 1, any LP can be transformed into one with this property, but this transformation increases the number of constraints and blows up the order of the problem, and hence may be undesirable in practical applications. In this subsection, we discuss a sliding objective value approach that can be used to solve the original problem by itself using Karmarkar's algorithm, when the optimum objective value is unknown.

For a given LP, the first problem is to determine whether it is feasible or not. Let the system of constraints be

$$\begin{aligned} Ax &= b \\ x &\geqq 0 \end{aligned}$$

where A is of order $m \times n$. As shown in Subsection 2, to check whether this system is feasible, we solve the following LP with the artificial variable x_{n+1}. Let $x^0 > 0$ be any vector.

$$\begin{array}{ll} \text{minimize} & x_{n+1} \\ \text{subject to} & Ax - x_{n+1}(Ax^0 - b) = b \\ & x \geqq 0, \quad x_{n+1} \geqq 0 \end{array} \tag{11.19}$$

$(x^0, 1) > 0$ is a feasible solution to this problem. The original problem is feasible iff the optimum objective value in this problem is zero. Even though the exact optimum objective value in this problem is unknown, we know that it lies between 0 and 1. Using it, this problem could be solved by Karmarkar's algorithm with the sliding objective value approach discussed below.

Now consider the general LP

$$\begin{array}{ll} \text{minimize} & cx \\ \text{subject to} & Ax = b \\ & x \geqq 0 \end{array} \tag{11.20}$$

This problem can be solved in two stages. First we check whether it is feasible, as discussed above. If a feasible solution $\bar{x}$ is obtained, $c\bar{x}$ is an upper bound on the optimum objective value in (11.20). We could then check whether the dual problem is feasible. If the dual is infeasible, from the duality theory of linear programming we know that cx is unbounded below in (11.20) (since (11.20) has already been verified to be feasible). If the dual is feasible, the dual objective value at the dual feasible solution obtained is lower bound on the optimum objective value in (11.20).

Now, consider the LP in the form discussed in (11.11)

$$\begin{array}{ll} \text{minimize} & dx \\ \text{subject to} & x \in \Omega \cap \mathbf{S} \\ \text{where} & \Omega = \{x : Ax = 0\} \\ & \mathbf{S} = \{x : x \geqq 0, \sum_{j=1}^{n} x_j = 1\} \end{array} \tag{11.21}$$

where A is a matrix of order $m \times n$ and rank m. We assusme that an optimum solution exists and that the optimum objective value is known to be between the given lower and upper bound l_0, u_0 (if the original problem is transformed directly into this form using the techniques discussed in Subsections 2, 3, we could take $l_0 = -2^{\mathbf{L}}$ and $u_0 = 2^{\mathbf{L}}$, where L is the size of the problem, under the assumption that an optimum solution exists). The difference between the current lower and upper bounds on the objective value is called the range. The sliding objective value approach is divided into several phases. At the end of each phase the range reduces to at least $\frac{2}{3}$ of its length at the beginning of the phase and takes no more than $n(k+\log(n))$ steps where k is a constant satisfying

$$\left(1 - \frac{\delta}{n}\right)^{\mathbf{kn}} \leq \frac{1}{2} .$$

Let z^* denote the unknown optimum objective value in (11.21). We run the algorithm pretending that a selected value, $\bar{z}$ is the minimum objective value (the value of $\bar{z}$ is updated at the beginning of each phase), that is, we try to minimize $dx - \bar{z} = (d - \bar{z}e^T)x$. This leads to the problem

$$\begin{array}{ll} \text{minimize} & cx \\ \text{subject to} & x \in \Omega \cap \mathbf{S} \end{array}$$

with $c = d - \bar{z}e^T$. We need to modify the computation of the vector g' in each step of the algorithm as follows. Compute g' as in the subroutine discussed in Subsection 6. Check if $\hat{c}g' < 0$. If so, choose the point g'' on the line segment joining a^0 and g' which satisfies $\hat{c}g'' = 0$, and make the point g'' the output of the subroutine instead of g'.

If $z^* \leqq \bar{z}$, let x^m be the point where $\hat{c}x$ achieves its minimum over $\Omega'' \cap \mathbf{B}(a^0, \alpha\underline{\rho})$. If $\hat{c}x^m < 0$, then define x^* to be the point on the line segment joining a^0 and x^m satisfying $\hat{c}x^* = 0$. Then all the proofs go through, and each step of the algorithm leads to a reduction of δ in the potential function or finds a point where the original objective function is $\bar{z}$.

Now a phase in the sliding objective value approach consists of the following. Let l, u be the current lower and upper bounds for the objective value dx at the beginning of the phase. Let

$$\bar{l} = l + \frac{1}{3}(u - l)$$
$$\bar{u} = l + \frac{2}{3}(u - l)$$

Figure 11.6

Run the algorithm as described above with $\bar{z}$ = pretended minimum objective value = $\bar{l}$.

If we obtain a feasible solution x which satisfies $dx < \bar{u}$, then terminate the phase, make dx the new upper bound u, and go to the next phase with the new bounds for the objective value.

Suppose after $n(k + \log(n))$ steps we have not reached a solution x with $dx < \bar{u}$. If $z^* \leqq \bar{l}$, we must have achieved a reduction δ in the associated potential function in each step, forcing the objective value dx to be $< \bar{u}$. So, if after $n(k + \log(n))$ steps we have not reached a solution x with $dx < \bar{u}$, we must have $z^* \geqq \bar{l}$. So make $\bar{l}$ the new lower bound l, and go to the next phase with the new bounds for the objective value.

Thus the length of the range gets multiplied by a factor $\frac{2}{3}$ or less during each phase. So after $\mathcal{O}(L)$ phases (i. e., after $\mathcal{O}(nL \log n)$ steps) we narrow the range to within $2^{-\mathcal{O}(L)}$ of the optimum objective value, and then obtain the exact optimum solution from the solution at that stage.

9 Implementation Difficulties

Consider the LP in standard form, find $y \in \mathbf{R}^n$ to

$$\begin{array}{ll} \text{minimize} & gy \\ \text{subject to} & Gy = d \\ & y \geqq 0 \ . \end{array} \tag{11.22}$$

The primal simplex algorithm for solving this problem processes the problem as it is in (11.22). It performs a sequence of operations on the data G, d, g until the problem is solved.

To solve (11.22) by Karmarkar's algorithm in the form discussed in Subsection 6, we have to first convert the problem into the form (11.11). As pointed out in Subsection 3, we add the additional constraint

$$\sum_{j=1}^{n} y_j + y_{n+1} = M \ .$$

Mathematically, taking M to be $2^{\mathbf{L}}$ where L is the size of the LP (11.22), would suffice; but in practical implementations M could be any practically reasonable upper bound for $\sum_{j=1}^{n} y_j$ in the problem. Using this additional constraint, (11.22) is transformed into the form

$$\begin{array}{ll} \text{minimize} & gy \\ \text{subject to} & Gy - \left(\frac{1}{M}\right) d \left(\sum\limits_{j=1}^{n+1} y_j \right) = 0 \\ & \sum\limits_{j=1}^{n+1} y_j = 1 \\ & y \geqq 0, j = 1 \text{ to } n+1 \end{array} \tag{11.23}$$

which is in Karmarkar's form.

LP models arising in practical applications lead to problems of the form (11.22) in which the coefficient matrix G is very sparse, that is, most of the entries in it are zero. Commercial implementations of the primal simplex algorithm exploit this sparsity and are able to take tremendous advantage of it. When the problem is transformed into the form (11.23) as discussed above, the resulting coefficient matrix A is usually totally dense, that is, almost all the entries in it are nonzero. This makes it very difficult to produce a practically viable implementation of Karmarkar's algorithm, at least for the algorithm in the form that is stated above. One may be able to overcome this problem by not computing A explicitly, but storing it as $G - \left(\frac{1}{M}\right)de^T$.

Now, consider the LP in the following form

$$\begin{array}{ll} \text{minimize} & cx \\ \text{subject to} & Ax = 0 \\ & e^T x = 1 \\ & x \geqq 0 \end{array} \tag{11.24}$$

The primal simplex algorithm would solve (11.24) by performing operations on the constraint matrix A directly. Karmarkar's algorithm operates on AA^T or $AD^{\mathbf{2}}A^T$ where D is a positive diagonal matrix. The computation of this matrix product is an additional burden in Karmarkar's algorithm. In fact an implementation of Karmarkar's algorithm which maintains $(AD^{\mathbf{2}}A^T)^{-1}$ in any form and updates it exactly from step to step in the algorithm, is not likely to be competitive with efficient implementations of the primal simplex algorithm.

Let D_r denote the diagonal matrix in step $r+1$ of Karmarkar's algorithm applied to (11.24). The computations (as discussed in Subsections 6, 7) in this step of the algorithm can be carried out by doing the following.

First solve the following system of equations for the row vector of variables $u = (u_1, \ldots, u_m)$

$$u(AD_r^{\mathbf{2}}A^T) = cD_r^{\mathbf{2}}A^T \ . \tag{11.25}$$

Let u^r denote the exact solution of this system. Then compute the $1 \times n$ row vector $\hat{c}_p^r$ from

$$\hat{c}_p^r = cD_r - u^r AD_r - cD_r ee^T \ .$$

This $(\hat{c}_p^r)^T$ is the direction for moving from a^0 to the boundary of the sphere $\mathbf{B}(a^0, \alpha\underline{\rho})$ in this step. It provides the steepest descent direction for minimizing the linear function $\hat{c}x'$ over $\Omega'' \cap \mathbf{B}(a^0, \alpha\underline{\rho})$ in this step. See Figure 11.7. In reality, we dot not need $\hat{c}_p^r$ exactly. Any approximate vector $\bar{c}_p^r$ that makes a strict acute angle will be adequate (the closer this angle is to 0 the better), it produces a decrease in objective function which may suffice in practice.

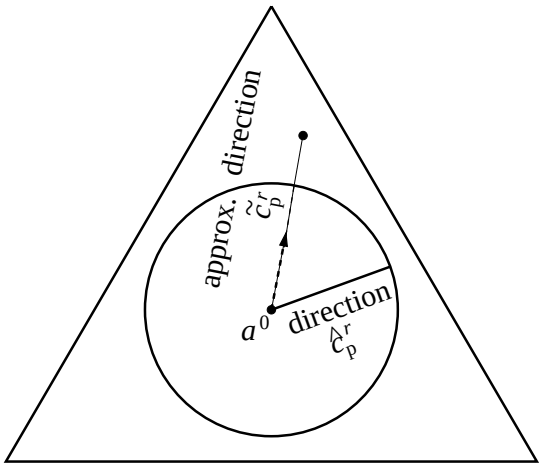


Figure 11.7 Steepest descent direction $\hat{c}_p^r$ for linear objective funtion $\hat{c}x$ in step r. Approximate descent direction $\tilde{c}_p^r$.

The key point is to get an approximate solution $\bar{u}^r$ for (11.25) efficiently, so that if

$$\tilde{c}_p^r = cD_r - \bar{u}^r AD_r - (cD_r ee^T)$$

that would satisfy $\tilde{c}_p^r \hat{c}_p^r > 0$ (acute angle condition). We also need $AD_r \tilde{c}_p^r = 0$ and $e^T \tilde{c}_p^r = 0$, so that moving from a^0 in the direction $\tilde{c}_p^r$ keeps the point within Ω''. Also given the approximate $\bar{u}^r$, how to update it into $\bar{u}^{r+1}$ that works for the $(r+1)$ in step the same way, when D_r changes to D_{r+1}. Some iterative methods for solving linear equations that produce approximate solutions efficiently may provide the key to this computation, and these are being investigated.

Also, once the direction of movement $\tilde{c}_p^r$ is obtained, in practical implementations one may want to move all the way closer to the boundary of the simplex, rather than to the boundary of the insphere $\mathbf{B}(a^0, \alpha\underline{\rho})$ as indicated in Figure 11.7. Since the simplex is determined by linear constraints, this can be done efficiently through a minimum ratio computation to determine how far you can move in this direction while retaining feasibility, and you can stop just a little bit short of it.

These and various other ideas are being explored for producing a practically useful implementation of Karmarkar's algorithm.

10 Solving Quadratic and Convex Programs by Karmarkar's Approach

It should be possible to extend Karmarkar's algorithm to solve convex quadratic programs and LCPs associated with PSD matrices, and possibly even smooth nonlinear convex programming problems. These extensions, and the best implementations of them, are now active research topics.

11.4.2 Tardos' New Strongly Polynomial Minimum Cost Circulation Algorithm

Consider a directed single commodity flow capacitated network with n nodes and m arcs. In [11.7] E. Tardos developed an algorithm for finding a minimum cost circulation in this network, with worst case computational complexity of $\mathcal{O}(m^{\frac{3}{2}}n^5)$ or $\mathcal{O}(m^2n^3 \log m)$ with some improvements. She has applied the idea of this algorithm and developed an algorithm to solve the general linear programming problem

$$\begin{array}{ll} \text{minimize} & cx \\ \text{subject to} & Ax = b \\ & x \geqq 0 \end{array}$$

in time polynomial in the size of A.

It remains to be investigated whether this approach can be extended to solve LCPs (q, M) when M is PSD, in time polynomial in the size of M.

11.4.3 The Ellipsoid Method for Linear Programming

A version of the ellipsoid method for solving linear programming problems is presented in Chapter 15 of [2.26]. The approach outlined there, uses a scheme suggested by P. Gacs and L. Lovász in a terminal step in order to obtain an optimum solution of the LP. Here we show how that terminal step can be replaced by a much more efficient scheme similar to the one discussed in Subsection 5 of Section 11.4.1. This has been suggested by R. Chandrasekaran and K. Truemper.

Consider an LP with rational data. By the techniques discussed in Section 1.2 and by scaling, this LP can be transformed into the problem

$$\min\ cu, \quad Fu \geqq g, \quad u \geqq 0 \tag{11.26}$$

where F, g, c are integer matrices. Let v denote the column vector of dual variables. By the duality theorem of linear programming (also see Subsection 1 of Section 11.4.1) solving this LP is equivalent to solving the system of linear inequalities (11.27).

$$\begin{array}{rr} -Fu & \leqq -g \\ F^Tv & \leqq c^T \\ cu\ -g^Tv & \leqq 0 \\ -u & \leqq 0 \\ -v & \leqq 0 \end{array} \tag{11.27}$$

Let $x = \begin{pmatrix} u \\ v \end{pmatrix}$. The system (11.27) is a system of linear inequalities in which all the coeficients are integer. Let D, b denote the coefficient matrix and right hand

side constants vector in (11.27), including the sign restrictions on the variables. Then (11.27) can be written as

$$Dx \leqq b \ . \tag{11.28}$$

Let D be of order $m \times n$, and L be the size of (11.28) (that is, L is the total number of binary digits in all the data in (11.28), see Section 8.3). Let $L_1 = 3\big((m+1)(n+1)+1\big)L$. As in Section 15.3 of [2.26], consider the perturbed system

$$2^{\mathbf{L_1}}(D_i.x) < 2^{\mathbf{L_1}}b_i + 1, \quad i = 1 \text{ to } m \tag{11.29}$$

(11.29) is now an open system of linear inequalities with integer data, and hence it can be solved by the ellipsoid method discussed in Section 15.2 of [2.26] in polynomial time. The method begins with an arbitrary point $x^0 \in \mathbf{R}^n$, and the matrix $A_0 = 2^{\mathbf{2(1+L_1)}}I$, where I is the unit matrix of order n, and generates the sequence (x^r, A_r), $r = 1, 2, \ldots$ using the iterative scheme (8.7) discussed in Section 8.4. For some r, if x^r satisfies (11.29), define $\tilde{x}$ to be that feasible x^r and go to the terminal step discussed below. If x^r violates (11.29), find a constraint in (11.29) violated by x^r, suppose it is the p^{th} constraint in (11.29). Then define $a = 2^{\mathbf{L_1}}D_p.$ and $d = 1 + 2^{\mathbf{L_1}}b_p$, and compute γ_{r+1} as in (8.7) using this a, d, x^r and A_r. If $\gamma_{r+1} \leqq -1$, (11.29) is infeasible, terminate the ellipsoid algorithm. If $\gamma_{r+1} > -1$, compute x^{r+1}, A_{r+1} as in (8.7), and continue.

If the ellipsoid algorithm continues for $r = 0$ to $6(n+1)^{\mathbf{2}}(m+1)(n+1)(L+L_1)$ steps and all the points x^r obtained in the algorithm are infeasible to (11.29), terminate with the conclusion that (11.29) has no feasible solution. The proofs of this and the other infeasibility conclusion stated earlier, are given in Chapter 15 of [2.26]. Under this infeasibility termination, (11.28), that is, (11.27), has no feasible solution, this implies that either the LP (11.26) is infeasible, or it is feasible and the objective function is unbounded below on its set of feasible solutions, we terminate.

Otherwise, let $\tilde{x}$ be the feasible solution for (11.29) obtained by the ellipsoid algorithm discussed above. If $\tilde{x} = \begin{pmatrix} \tilde{u} \\ \tilde{v} \end{pmatrix}$ is feasible to (11.28), then $\tilde{u}$ is an optimal solution of (11.26) and $\tilde{v}$ is an optimal dual solution, terminate. If not, consider the following system.

$$Dx^+ - Dx^- + It = b + e2^{-\mathbf{L_1}} \tag{11.30}$$

where $x^+ = (x_1^+, \ldots, x_n^+)$, $x^- = (x_1^-, \ldots, x_n^-)$, $t = (t_i) \in \mathbf{R}^m$ and e is the column vector in $\mathbf{R}^m$ of all 1's. Define for $j = 1$ to n

$$\tilde{x}_j^+ = \begin{cases} 0 & \text{if } \tilde{x}_j \leqq 0 \\ \tilde{x}_j & \text{if } \tilde{x}_j \geqq 0 \end{cases}$$

$$\tilde{x}_j^- = \begin{cases} |\tilde{x}_j| & \text{if } \tilde{x}_j \leqq 0 \\ 0 & \text{if } \tilde{x}_j \geqq 0 \end{cases}$$

$$\tilde{t}_i = -D_i.\tilde{x} + b_i + 2^{-\mathbf{L_1}}$$

$$\tilde{x}^+ = (\tilde{x}_j^+), \quad \tilde{x}^- = (\tilde{x}_j^-), \quad \tilde{t} = (\tilde{t}_i)$$

Then $(\tilde{x}^+, \tilde{x}^-, \tilde{t})$ is feasible to (11.30). Using the method discussed in Subsection 5 of Section 11.4.1, or Section 3.5.4 of [2.26] (here there is no objective function involved, so

we just apply this method without worrying about the objective value), obtain a BFS $(\hat{x}^+, \hat{x}^-, \hat{t}) \in \mathbf{R}^{2n+m}$ to (11.30). Denote the vector $(x^+, x^-, t) \in \mathbf{R}^{2n+m}$ by y and let $\hat{y} = (\hat{x}^+, \hat{x}^-, \hat{t})$. Since $\hat{y}$ is a BFS of (11.30), there exists a basis B, a square submatrix of $(D \vdots -D \vdots I)$ of order m, so that $\hat{y} = (\hat{y}_B, \hat{y}_E)$ is given by

$$\begin{aligned} \hat{y}_E &= 0 \\ \hat{y}_B &= B^{-1}(b + e2^{-\mathbf{L_1}}) \ . \end{aligned} \tag{11.31}$$

Here E is the submatrix of $(D \vdots -D \vdots I)$ consisting of all the columns other than those in B, and y_B, y_E are the basic, nonbasic vectors of variables y_j corresponding to the basic, nonbasic partition $(B \vdots E)$ of $(D \vdots -D \vdots I)$. Now define the vector $y^* = (y_B^*, y_E^*)$ by

$$\begin{aligned} y_E^* &= 0 \\ y_B^* &= B^{-1}b \end{aligned} \tag{11.32}$$

and let $y^* = (x^{*+}, x^{*-}, t^*)$ in terms of the original variables. Let

$$x^* = x^{*+} - x^{*-} \ .$$

The vector y^* is the basic solution of the system

$$Dx^+ - Dx^- + It = b \tag{11.33}$$

corresponding to the basis B. By Theorem 15.1 of [2.26], |determinant of B| $< 2^{\mathbf{L}}$, and hence using an argument similar to that in Theorem 15.2 of [2.26] we have, for $i = 1$ to m

$$\text{either} \quad t_i^* = 0 \quad \text{or } |t_i^*| > 2^{-\mathbf{L}} \ . \tag{11.34}$$

Let $\mathbf{J} = \{i : 1 \leqq i \leqq m$, and i such that t_i is a basic variable corresponding to the basis $B\}$. So, from the definition of $\hat{y}$, and from (11.31), (11.32), we have

$$\hat{t}_i = \begin{cases} 0, & \text{for all } i \notin \mathbf{J} \\ t_i^* + (B^{-1}e2^{-\mathbf{L_1}})_i, & \text{for } i \in \mathbf{J} \ . \end{cases} \tag{11.35}$$

From well known results in the theory of determinants, B^{-1} is the adjoint of B multiplied by a scalar, which is the inverse of the determinant of B. The determinant of the basis B is a nonzero integer and hence has absolute value $\geqq 1$. Each entry in the adjoint of B is the determinant of a square submatrix of B, by Theorem 15.1 of [2.26] its absolute value is $\leqq \frac{2^{\mathbf{L}}}{n}$. So $|(B^{-1}e2^{-\mathbf{L_1}})_i| \leqq \frac{m2^{\mathbf{L}}}{n2^{\mathbf{L_1}}} \leqq \frac{2^{\mathbf{L}}}{2^{\mathbf{L_1}}} \leqq 2^{-\mathbf{nL}}$. But $\hat{y}$ is a BFS of (11.30), so $\hat{t}_i \geqq 0$ for all i. Using this and (11.34) in (11.35) we conclude that t_i^* must be $\geqq 0$ for all $i \in \mathbf{J}$. We already know that $t_i^* = 0$ for all $i \notin \mathbf{J}$. So $t^* \geqq 0$. This clearly implies that x^* is feasible to (11.28). Therefore if $x^* = \begin{pmatrix} u^* \\ v^* \end{pmatrix}$, u^* is an optimum solution of (11.26) and v^* is an optimum dual solution. From u^*, a

basic feasible optimum solution of (11.26) can be obtained by the method described in Subsection 5 of Section 11.4.1.

The ellipsoid method is the first mathematical device used to prove that linear programming is in the class $\mathcal{P}$ of problems solvable in polynomial time. The modified terminal step given above is not adequate to make the ellipsoid method practically useful. However, the ellipsoid method remains a very important mathematical tool in the study of computational complexity of optimization problems.

11.4.4 The Gravitational Method for Linear Programming

Here we briefly describe an interior point variant of the gradient projection method for linear programming proposed by K. G. Murty [11.5, 11.6]. We consider the LP in the following form

$$\begin{array}{ll} \text{minimize} & z(x) = cx \\ \text{subject to} & Ax \geqq b \end{array} \tag{11.36}$$

where A is a matrix of order $m \times n$. Sign restrictions on the variables and any other lower or upper bound conditions on the variables, if any, are all included in the above system of constraints. Clearly every LP can be put in this form by well known simple transformations discussed in most LP textbooks (for example, see [2.26]).

Note: In practical applications, it usually turns out that the LP model for a practical problem is in standard form

$$\begin{array}{c} \min\ p\chi \\ \text{subject to } B\chi = d \\ \chi \geqq 0. \end{array} \tag{11.37}$$

The dual of this model is directly in form (11.36) and the gravitational method can be applied to solve the dual of (11.37) directly. As it will be shown later on, when the gravitational method is applied on the dual of (11.37), at termination, it will produce an optimum solution for (11.37), if one exists.

Assumptions

Let $\mathbf{K}$ denote the set of feasible solutions of (11.36). We assume that $\mathbf{K} \neq \emptyset$, and that $\mathbf{K}$ has a nonempty interior in $\mathbf{R}^n$, and that an initial interior feasible solution x^0 (this is a point x^0 satisfying $Ax^0 > b$) of (11.36) is available.

If these assumptions are not satisfied, introduce an artificial variable x_{n+1} and modify the problem as follows

$$\begin{array}{ll} \text{minimize} & cx + vx_{n+1} \\ \text{subject to} & Ax + ex_{n+1} \geqq b,\ x_{n+1} \geqq 0 \end{array} \tag{11.38}$$

where $e = (1, \ldots, 1)^T \in \mathbf{R}^m$ and v is a large positive number. For any $\hat{x} \in \mathbf{R}^n$, let $\hat{x}_{n+1} > \max\{|\min\{0, A_i.\hat{x} - b_i\}| : i = 1 \text{ to } m\}$, then $(\hat{x}, \hat{x}_{n+1})$ satisfies the constraints in (11.38) as strict inequalities. Thus the modified problem (11.38) satisfies all the assumptions made in the above paragraph.

We also assume that $c \neq 0$, as otherwise x^0 is optimal to (11.36), and we can terminate.

The Gravitational Method

The Euclidean distance of x^0 from the hyperplane $\{x : A_i.x = b_i\}$ is $(A_i.x^0 - b_i)/\|A_i.\|$.

The gravitational approach for solving (11.36) is the following. Assume that the boundary of $\mathbf{K}$ is an impermeable layer separating the inside of $\mathbf{K}$ from the outside. Introduce a powerful gravitational force inside $\mathbf{K}$ pulling everything down in the direction $-c^T$. Choose $0 < \varepsilon < \min\{(A_i.x^0 - b_i)/\|A_i.\| : i = 1 \text{ to } m\}$. Release a small spherical n-dimensional drop of mercury of diameter 2ε with its center at the initial interior feasible solution $x^0 \in \mathbf{K}$. The drop will fall under the influence of gravity. During its fall, the drop may touch the boundary, but the center of the drop will always be in the interior of $\mathbf{K}$ at a distance $\geqq \varepsilon$ from the nearest point to it on the boundary. Whenever the drop touches a face of $\mathbf{K}$, it will change direction and will continue to move, if possible, in the gravitational direction that keeps it within $\mathbf{K}$. If the objective function is unbounded below in (11.36), after changing direction a finite number of times, the drop will continue to fall forever along a half-line in $\mathbf{K}$ along which the objective function diverges to $-\infty$. If $z(x)$ is bounded below on $\mathbf{K}$, after changing direction a finite number of times, the drop will come to a halt. The algorithm tracks the path of the center of the drop as it falls in free fall under the influence of gravity. Let $\mathbf{P}$ denote this path of the center of this drop in its fall.

The Gravitational Direction at an Interior Point $\overline{x} \in \mathbf{K}$

Suppose a drop of radius ε, with its center at $\overline{x}$ is inside $\mathbf{K}$. So

$$(A_i.\overline{x} - b_i)/\|A_i.\| \geqq \varepsilon, \quad i = 1 \text{ to } m. \tag{11.39}$$

At every point $\overline{x}$ on the locus $\mathbf{P}$ of the center of the drop in the gravitational method, (11.39) will always be satisfied. Given a point $\overline{x}$ on $\mathbf{P}$, define

$$\mathbf{J}(\overline{x}) = \{i : (A_i.\overline{x} - b_i)/\|A_i.\| = \varepsilon\}. \tag{11.40}$$

The hyperplane $\{x : A_i.x = b_i\}$ is touching the drop of radius ε when its center is at the interior point $\overline{x} \in \mathbf{K}$ only if $i \in \mathbf{J}(\overline{x})$. Now, define

$$y^0 = -c^T/\|c\|. \tag{11.41}$$

If $\mathbf{J}(\overline{x}) = \emptyset$ (i. e., if $(A_i.\overline{x} - b_i)/\|A_i.\| > \varepsilon$ for all $i = 1$ to m), when the drop is in a position with its center at $\overline{x}$, it will move in the gravitational direction y^0. The distance that it will move in this direction is

$$\theta = \text{ minimum } \left\{ \frac{(A_i.\overline{x} - b_i) - \varepsilon\|A_i.\|}{-A_i.y^0} : 1 \leqq i \leqq m \text{ and } i \text{ such that } A_i.y^0 < 0 \right\} \tag{11.42}$$

where we adopt the convention that the minimum in the empty set is $+\infty$. If $\theta = +\infty$ in (11.42), then the drop continues to move indefinitely along the half-line $\{\overline{x} + \lambda y^0 : \lambda \geqq 0\}$, and $z(x)$ is unbounded below on this feasible half-line, terminate. If θ is finite in (11.42), at the end of this move, the drop will be in a position with its center at $\overline{x} + \theta y^0$, touching the boundary of $\mathbf{K}$, and it will either halt (see the conditions for this, discussed later on) or change direction into the gravitational direction at $\overline{x} + \theta y^0$ and move in that direction.

When $\overline{x}$ is such that $\mathbf{J}(\overline{x}) \neq \emptyset$, that is,

$$\min\{(A_i.\overline{x} - b_i)/\|A_i.\| : i = 1 \text{ to } m\} = \varepsilon \tag{11.43}$$

the direction that the drop will move next, called the **gravitational direction** at $\overline{x}$, can be defined using many different principles. One principle to define the gravitational direction at $\overline{x}$, where $\overline{x}$ is an interior point of $\mathbf{K}$ satisfying (11.43) is by the following procedure, which may take several steps.

Step 1 : If the drop moves in the direction y^0 from $\overline{x}$, the position of its center will be $\overline{x} + \lambda y^0$ for some $\lambda > 0$. Since (11.39) holds, the ith constraint will block the movement of the drop in the direction y^0, only if $i \in \mathbf{J}(\overline{x})$ and $A_i.y^0 < 0$. Define

$$\mathbf{J}_1 = \{i : i \in \mathbf{J}(\overline{x}), \text{ and } A_i.y^0 < 0\}.$$

Case 1 : $\mathbf{J}_1 = \emptyset$: If $\mathbf{J}_1 = \emptyset$, y^0 is the gravitational direction at $\overline{x}$, and the distance it can move in this direction is determined as in (11.42).

Case 2 : $\mathbf{J}_1 \neq \emptyset$: If $\mathbf{J}_1 \neq \emptyset$, each of the constraints $A_i.x \geqq b_i$ for $i \in \mathbf{J}_1$, is currently blocking the movement of the drop in the direction y^0.

Define $\mathbf{T}_1 = \mathbf{J}_1$, and let D_1 be the matrix of order $|\mathbf{T}_1| \times n$ whose rows are $A_i.$ for $i \in \mathbf{T}_1$. Let E_1 be the submatrix of D_1 of order (rank of D_1) $\times\ n$, whose set of rows is a maximal linearly independent subset of row vectors of D_1. Let $\mathbf{I}_1 = \{i : A_i.$ is a row vector of $E_1\}$. So $\mathbf{I}_1 \subset \mathbf{T}_1$. Let $\mathbf{F}_1$ be the subspace $\{x : D_1x = 0\} = \{x : E_1x = 0\}$, F_1 is the subspace corresponding to the set of all constraints which are blocking the movement of the drop in the direction y^0. Let ξ^1 be the orthogonal projection of y^0 in the subspace F_1, that is

$$\xi^1 = (I - E_1^T(E_1E_1^T)^{-1}E_1)y^0. \tag{11.44}$$

Subcase 2.1 : $\xi^1 \neq 0$: If $\xi^1 \neq 0$, let $y^1 = \xi^1/\|\xi^1\|$, go to Step 2.

Subcase 2.2 : $\xi^1 = 0$: If $\xi^1 = 0$, let the row vector $\mu = (\mu_i : i \in \mathbf{I}_1) = -\|c\|((E_1E_1^T)^{-1}E_1y^0)^T$. Then $\mu E = c$.

Subcase 2.2.1 : $\xi^1 = 0$ and $\mu \geqq 0$: If $\mu \geqq 0$, define the row vector $\overline{\pi} = (\overline{\pi}_i)$ by

$$\begin{aligned}\overline{\pi}_i &= 0, \text{ if } i \notin \mathbf{I}_1\\ &= \mu_i, \text{ if } i \in \mathbf{I}_1.\end{aligned}$$

Then $\overline{\pi}$ is a basic feasible solution to the dual of (11.36). In this case, as will be shown later on, the drop halts in the current position, it cannot roll any further, under the gravitational force.

Subcase 2.2.2 : $\xi^1 = 0$, $\mu \not\geqq 0$: If $\xi^1 = 0$ and $\mu \not\geqq 0$, delete the i corresponding to the most negative μ_i from the set $\mathbf{I}_1$ (any other commonly used rule for deleting one or more of the i associated with negative μ_i from $\mathbf{I}_1$ can be applied in this case). Redefine the matrix E_1 to be the one whose rows are $A_i.$ for i in the new set $\mathbf{I}_1$, compute the new orthogonal projection ξ^1 as in (11.44) using the new E_1 and repeat Subcase 2.1 or 2.2 as appropriate with the new ξ^1.

General Step r : Let y^{r-1} be the direction determined in the previous step. Define

$$\mathbf{J}_r = \{i : i \in \mathbf{J}(\overline{x}) \text{ and } A_i.y^{r-1} < 0\}.$$

Case 1 : $\mathbf{J}_r = \emptyset$: If $\mathbf{J}_r = \emptyset$, y^{r-1} is the gravitational direction at $\overline{x}$, and the distance the drop can move in this direction is determined as in (11.42) with y^{r-1} replacing y^0.

Case 2 : $\mathbf{J}_r \neq \emptyset$: Define $\mathbf{T}_r = \bigcup_{s=1}^{r} \mathbf{J}_s$ and let D_r be the matrix of order $|\mathbf{T}_r| \times n$ whose rows are $A_i.$ for $i \in \mathbf{T}_r$. Let E_r be the submatrix of D_r of order (rank of D_r) $\times n$, whose set of rows is a maximal linearly independent subset of row vectors of D_r. Let $\mathbf{I}_r = \{i : A_i. \text{ is a row vector of } E_r\}$. Let $\mathbf{F}_r$ be the subspace $\{x : D_rx = 0\} = \{x : E_rx = 0\}$. Let ξ^r be the orthogonal projection of y^0 in the subspace $\mathbf{F}_r$, that is

$$\xi^r = (I - E_r^T(E_rE_r^T)^{-1}E_r)y^0.$$

Subcase 2.1 : $\xi^r \neq 0$: Let $y^r = \xi^r/\|\xi^r\|$, go to Step $r+1$.

Subcase 2.2 : $\xi^r = 0$: Let $\mu = (\mu_i : i \in \mathbf{I}_r) = -\|c\|((E_rE_r^T)^{-1}E_ry^0)^T$.

Subcase 2.2.1 : $\xi^r = 0$, and $\mu \geqq 0$: Define $\overline{\pi} = (\overline{\pi}_i)$ by

$$\begin{aligned}\overline{\pi}_i &= 0, \text{ for } i \notin \mathbf{I}_r\\ &= \mu_i, \text{ for } i \in \mathbf{I}_r.\end{aligned}$$

$\overline{\pi}$ is a basic feasible solution to the dual of (11.36). In this case the drop halts, it cannot roll any further under the gravitational force.

Subcase 2.2.2 : $\xi^r = 0$, and $\mu \not\geqq 0$: If $\xi^r = 0$ and $\mu \not\geqq 0$, proceed exactly as under Subcase 2.2.2 described under Step 1, with $\mathbf{I}_r$ replacing P_1.

It can be shown that this procedure does produce the gravitational direction at $\overline{x}$, finitely, if the drop can move at all. Currently work is being carried out on developing efficient methods for choosing the index set $\mathbf{I}_r$ of maximal linearly independent subset of row vectors of D_r, in Case 2, and on the best strategies for deleting a subset of constraints associated with negative μ_i in Subcase 2.2.2. Other principles for defining the gravitational direction at the interior point $\overline{x}$ of $\mathbf{K}$, are also being investigated.

Conditions for the Halting of the Drop

Let ε be the radius of the drop and $\overline{x} \in \mathbf{K}$ satisfy (11.39). We have the following theorem.

Theorem 11.6 *When the center of the drop is at $\overline{x}$, it halts iff $\mathbf{J}(\overline{x})$ defined in (11.40) is $\neq \emptyset$, and there exists a dual feasible solution $\overline{\pi} = (\overline{\pi}_i)$ for the dual of (11.36) satisfying*

$$\overline{\pi}_i = 0, \text{ for all } i \notin \mathbf{J}(\overline{x}). \tag{11.45}$$

Proof. The drop will halt when its center is at $\overline{x}$, iff there exists no direction at $\overline{x}$ along which the drop could move within the interior of $\mathbf{K}$, that will slide its center on a line of decreasing objective value for some positive length. That is, iff there exists no y satisfying

$$cy < 0$$
$$(A_i.(\overline{x} + \lambda y) - b_i)/\|A_i.\| \geqq \varepsilon, \ i = 1 \text{ to } m$$

for $0 \leqq \lambda < \alpha$, for some $\alpha > 0$. Since $\overline{x}$ satisfies (11.39), and from the definition of $\mathbf{J}(\overline{x})$ in (11.40), this implies that the drop will halt when its center is at $\overline{x}$ iff the system

$$A_i.y \geqq 0, \text{ for all } i \in \mathbf{J}(\overline{x})$$
$$cy < 0$$

has no solution y. By the well known Farkas' lemma, Theorem 3 in Appendix 1, this holds iff there exists a $\overline{\pi} = (\overline{\pi}_i : i = 1 \text{ to } m)$ feasible to the dual of (11.36) satisfying (11.45).

□

What to Do When the Drop Halts?

Theorem 11.7 *Suppose the drop of radius ε halts with its center at $\overline{x} \in \mathbf{K}$. Then the LP (11.36) has a finite optimum solution. Let z^* be the optimum objective value in (11.36). Let $\overline{\pi} = (\overline{\pi}_i)$ be the dual feasible solution satisfying (11.45) guaranteed to exist by Theorem 11.6. Then*

$$c\overline{x} = \overline{\pi}b + \varepsilon \sum_{i \in \mathbf{J}(\overline{x})} \overline{\pi}_i \tag{11.46}$$

and

$$c\overline{x} \leqq z^* + \varepsilon \sum_{i \in \mathbf{J}(\overline{x})} \overline{\pi}_i. \tag{11.47}$$

Proof. If the drop halts, by Theorem 11.6, the dual of (11.36) is feasible. So, the LP (11.36) has a finite optimum solution by the duality theory of LP. Consider the perturbed LP

$$\begin{array}{ll} \text{minimize} & z(x) = cx \\ \text{subject to} & A_{i.}x \geqq \begin{cases} b_i, & \text{for } i \notin \mathbf{J}(\overline{x}) \\ b_i + \varepsilon, & \text{for } i \in \mathbf{J}(\overline{x}). \end{cases} \end{array} \tag{11.48}$$

The hypothesis in the theorem implies that $\overline{x}$, $\overline{\pi}$, together satisfy the primal, dual feasibility and the complementary slackness optimality conditions for (11.48) and its dual. Hence, by the duality theorem of LP, (11.46) holds. Also, by the weak duality theorem of LP, (11.47) holds.

Hence, if the drop halts with its center at position $\overline{x}$, and a $\overline{\pi}$ satisfying (11.45) is found, and $\varepsilon \sum_{i \in \mathbf{J}(\overline{x})} \overline{\pi}_i$ is small, then $\overline{x}$ can be taken as a near optimum solution to (11.36) and the algorithm terminated. Also, in this case $\overline{\pi}$ is an optimum solution for the dual of (11.36), and the true optimum solution of (11.36) can be obtained by well known pivotal methods that move from $\overline{x}$ to an extreme point without increasing the objective value (see Subsection 5 in Section 11.4.1).

Theorem 11.8 *Suppose the drop of radius ε halts with its center at $\overline{x} \in \mathbf{K}$. If the system of equations*

$$A_{i.}x = b_i, \ i \in \mathbf{J}(\overline{x}) \tag{11.49}$$

has a solution $\tilde{x}$ which is feasible to (11.36), then $\tilde{x}$ is an optimum feasible solution of (11.36).

Proof. Let $\overline{\pi}$ be the dual feasible solution satisfying (11.45) guaranteed by Theorem 11.6. It can be verified that $\tilde{x}$, $\overline{\pi}$ together satisfy the complementary slackness optimality conditions for (11.36) and its dual, so $\tilde{x}$ is an optimum solution for (11.36). In this case $\overline{\pi}$ is optimum to the dual of (11.36).

□

If the drop of radius ε halts with its center at $\overline{x} \in \mathbf{K}$, and there exists no solution to the system of equations (11.49) which is feasible to (11.36), then this drop is unable to move any further down in $\mathbf{K}$ under the gravitational force, even though it is not close to an optimum solution for (11.36). See Figure 11.8.

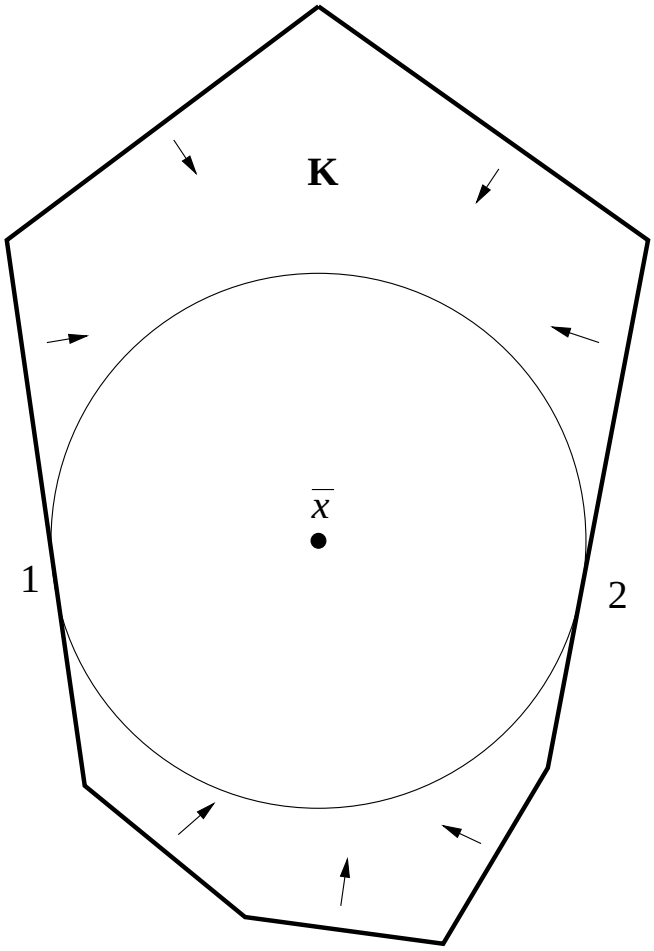


Figure 11.8 The set **K** is on the side of the arrow marked on each constraint. The gravitational force is pulling the drop straight down, but it cannot move any further, because it is squeezed between hyperplanes 1 and 2.

Suppose the drop of radius ε halts with its center at $\overline{x}$. If the system

$$A_i.x = b_i,\ i \in \mathbf{J}(x) \tag{11.50}$$

has no feasible solution, the gravitational method reduces the radius of the drop, see below, keeping the center at $\overline{x}$, and continues.

On the other hand, suppose the drop of radius ε halts with its center at $\overline{x}$, and the system (11.50) is feasible. Let E be the matrix whose rows form a maximal linearly independent subset of rows of $\{A_i. : i \in \mathbf{J}(\overline{x})\}$. Then the nearest point to $\overline{x}$ in the flat $\{x : A_i.x = b_i, i \in \mathbf{J}(\overline{x})\}$ is $\hat{x} = \overline{x} + E^T(EE^T)^{-1}(d - E\overline{x})$ where d is the column vector of b_i for i such that $A_i.$ is a row of E. If $\hat{x}$ is feasible to (11.36), then by Theorem 11.8, $\hat{x}$ is an optimum feasible solution for (11.36) and the method terminates. Otherwise, at this stage the gravitational method reduces the radius of the drop (for example, replace ε by $\varepsilon/2$), keeping the center at $\overline{x}$, and traces the locus of the center of the new

drop as it now begins to fall under the influence of gravity again. The same process is repeated when the new drop halts.

See Figure 11.9 for an illustration of the path of the drop in a convex polyhedron in $\mathbf{R}^3$.

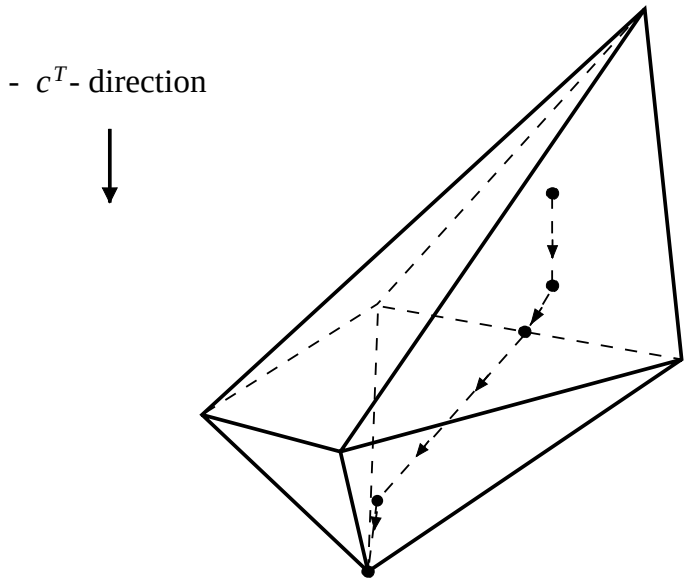


Figure 11.9 Path of the drop in the gravitational method in a convex polyhedron in $\mathbf{R}^3$.

The theoretical worst case computational complexity of this algorithm is currently under investigation. Initial computational trials with the method are very encouraging. The practical efficienty of this algorithm is also being studied via a computational project.

11.5 References

11.1 R. Chandrasekaran, S. N. Kabadi and K. G. Murty, "Some $\mathcal{NP}$-complete problems in linear programming", *Operations Research Letters*, 1, 3 (July 1982) 101–104.

11.2 N. Karmarkar, "A new polynomial-time algorithm for linear programming", *Combinatorica*, 4, no. 4 (1984) 373–395.

11.3 K. G. Murty, "Algorithm for finding all the feasible complementary bases for a linear complementarity problem", Technical Report 70-7, Department of Industrial and Operations Engineering, University of Michigan, Ann Arbor, MI 48109–2117.

11.4 K. G. Murty and Y. Fathi, "A feasible direction method for linear programming", *Operations Research Letters*, 3, no. 3 (August 1984) 121–127.

11.5 K. G. Murty, "A new interior variant of the gradient projection method for linear programming", Technical Report 85-18, Department of IOE, University of Michigan, Ann Arbor, MI 48109–2117, USA, 1985.

11.6 K. G. Murty, "The gravitational method for linear programming", *Opsearch*, 23, no. 4 (1986) 206–214.

11.7 E. Tardos, "A strongly polynomial minimum cost circulation algorithm", *Combinatorica*, 5 (1985) 247–255.

Appendix

PRELIMINARIES

1. THEOREMS OF ALTERNATIVES FOR SYSTEMS OF LINEAR CONSTRAINTS

Here we consider systems of linear constraints, consisting of equations or inequalities or both. A feasible solution of a system is a vector which satisfies all the constraints in the system. If a feasible solution exists, the **system** is said to be **feasible**. The system is said to be **infeasible** if there exists no feasible solution for it. A typical theorem of alternatives shows that corresponding to any given system of linear constraints, system I, there is another associated system of linear constraints, system II, based on the same data, satisfying the property that one of the systems among I, II is feasible iff the other is infeasible. These theorems of alternatives are very useful for deriving optimality conditions for many optimization problems.

First consider systems consisting of linear equations only. The **fundamental inconsistent equation** is

$$0 = 1 \tag{1}$$

consider the following system of equations

$$\begin{aligned} x_1 + x_2 + x_3 &= \quad 2 \\ -x_1 - x_2 - x_3 &= -1. \end{aligned} \tag{2}$$

When we add the two equations in (2), the coefficients of all the variables on the left hand side of the sum are zero, and the right hand side constant is 1. Thus the

fundamental inconsistent equation (1) can be obtained as a linear combination of the two equations in (2). This clearly implies that there exists no feasible solution for (2). Now consider the general system of linear equations $Ax = b$, written out in full as

$$\sum_{j=1}^{n} a_{ij}x_j = b_i,\ i = 1 \text{ to } m. \tag{3}$$

A linear combination of system (3) with coefficients $\pi = (\pi_1, \ldots, \pi_m)$ is

$$\sum_{j=1}^{n}\Big(\sum_{i=1}^{m} \pi_i a_{ij}\Big)x_j = \Big(\sum_{i=1}^{m} \pi_i b_i\Big) \tag{4}$$

(4) is the same as $(\pi A)x = (\pi b)$. (4) becomes the fundamental inconsistent equation (1) if

$$\begin{aligned} \sum_{i=1}^{m} \pi_i a_{ij} &= 0,\ j = 1 \text{ to } n \\ \sum_{i=1}^{m} \pi_i b_i &= 1 \end{aligned} \tag{5}$$

and in this case, (3) is clearly infeasible. The system of linear equations (3) is said to be **inconsistent** iff the fundamental inconsistent equation (1) can be obtained as a linear combination of the equations in (3), that is, iff there exists $\pi = (\pi_1, \ldots, \pi_m)$ satisfying (5). Clearly an inconsistent system of equations is infeasible. The converse of this statement is also true. So a system of linear equations is infeasible iff it is inconsistent. This is implied by the following theorem of alternatives for systems of linear equations.

Theorem 1 *Let $A = (a_{ij})$, $b = (b_i)$ be given matrices of orders $m \times n$ and $m \times 1$. Let $x = (x_1, \ldots, x_n)^T$ and $\pi = (\pi_1, \ldots, \pi_m)$. Exactly one of the two following systems (I) and (II) has a solution and the other has no solution.*

(I)	(II)
$Ax = b$	$\pi A = 0$
	$\pi b = 1$

Proof. If (I) has a solution $\overline{x}$ and (II) has a solution $\overline{\pi}$, then $A\overline{x} = b$, and so $\overline{\pi}A\overline{x} = \overline{\pi}b$, but $(\overline{\pi}A)\overline{x} = 0$, $\overline{\pi}b = 1$, so this is impossible. So it is impossible for both (I) and (II) to have solutions.

Put (I) in detached coefficient tabular form and introduce the unit matrix of order m on the left hand side of this tableau. The tableau at this stage is

		x	
I		A	b

Perform Gauss-Jordan pivot steps in this tableau to put A into row echelon normal form. For this, perform Gauss-Jordan pivot steps in rows 1 to m, in that order. Consider the step in which the rth row is the pivot row. Let the entries in the current tableau in the rth row at this stage be

$\beta_{r1} \ldots \beta_{rm}$		$\overline{a}_{r1} \ldots \overline{a}_{rn}$	$\overline{b}_r$

Let $\beta_r. = (\beta_{r1}, \ldots, \beta_{rm})$. Then, $(\overline{a}_{r1}, \ldots, \overline{a}_{rn}) = \beta_r.A$ and $\overline{b}_r = \beta_r.b$. If $(\overline{a}_{r1}, \ldots, \overline{a}_{rn}) = 0$ and $\overline{b}_r = 0$, this row at this stage represents a redundant constraint, erase it from the tableau and continue. If $(\overline{a}_{r1}, \ldots, \overline{a}_{rn}) = 0$ and $\overline{b}_r \neq 0$, define $\overline{\pi} = \beta_r./\overline{b}_r$. Then we have $\overline{\pi}A = 0$, $\overline{\pi}b = 1$, so $\overline{\pi}$ is a feasible solution of system (II) and (I) has no feasible solution, terminate. If $(\overline{a}_{r1}, \ldots, \overline{a}_{rn}) \neq 0$, select a j such that $\overline{a}_{rj} \neq 0$, and perform a pivot step with row r as the pivot row and column j as the pivot column, make x_j the basic variable in the rth row, and continue. If the conclusion that (I) is infeasible is not made at any stage in this process, make the basic variable in each row equal to the final updated right hand side constant in that row, and set all the nonbasic variables equal to zero; this is a solution for system (I). Since (II) cannot have a solution when (I) does, (II) has no solution in this case.

□

Example 1

Let

$$A = \begin{pmatrix} 1 & -2 & 2 & -1 & 1 \\ -1 & 0 & 4 & -7 & 7 \\ 0 & -2 & 6 & -8 & 8 \end{pmatrix}, \quad b = \begin{pmatrix} -8 \\ 16 \\ 6 \end{pmatrix}.$$

So, system (I) in Theorem 1 corresponding to this data is

x_1	x_2	x_3	x_4	x_5	b
1	−2	2	−1	1	−8
−1	0	4	−7	7	16
0	−2	6	−8	8	6

We introduce the unit matrix of order 3 on the left hand side and apply the Gauss-Jordan method on the resulting tableau. This leads to the following work. Pivot elements are inside a box.

			x_1	x_2	x_3	x_4	x_5	
1	0	0	$\boxed{1}$	−2	2	−1	1	− 8
0	1	0	−1	0	4	−7	7	16
0	0	1	0	−2	6	−8	8	6
1	0	0	1	−2	2	−1	1	− 8
1	1	0	0	$\boxed{-2}$	6	−8	8	8
0	0	1	0	−2	6	−8	8	6
0	−1	0	1	0	−4	7	−7	−16
$-\frac{1}{2}$	$-\frac{1}{2}$	0	0	1	−3	4	−4	− 4
−1	−1	1	0	0	0	0	0	− 2

From the last row in the last tableau, we conclude that this system is inconsistent. Defining $\overline{\pi} = (-1, -1, 1)/(-2) = (1/2, 1/2, -1/2)$, we verify that $\overline{\pi}$ is a solution for system (II) in Theorem 1 with data given above.

Now consider a system of linear inequalities. The **fundamental inconsistent inequality** is

$$0 \geqq 1 \tag{6}$$

Consider the following system of inequalities.

$$\begin{aligned} x_1 + x_2 + x_3 &\geqq 2 \\ -x_1 - x_2 - x_3 &\geqq -1. \end{aligned} \tag{7}$$

Adding the two inequalities in (7) yields the fundamental inconsistent inequality (6), this clearly implies that no feasible solution exists for (7).

Given the system of linear inequalities

$$\sum_{j=1}^{n} a_{ij} x_j \geqq b_i, \ i = 1 \text{ to } m \tag{8}$$

a **valid linear combination** of (8) is a linear combination of the constraints in (8) with nonnegative coefficients, that is

$$\sum_{j=1}^{n} \Big(\sum_{i=1}^{m} \pi_i a_{ij} \Big) x_j \geqq \sum_{i=1}^{m} \pi_i b_i \tag{9}$$

where $\pi = (\pi_1, \ldots, \pi_m) \geqq 0$. (9) is the fundamental inconsistent equation (6) iff

$$\begin{aligned} \sum_{j=1}^{n} \pi_i a_{ij} &= 0, \ j = 1 \text{ to } m \\ \sum_{i=1}^{m} \pi_i b_i &= 1 \end{aligned} \tag{10}$$

and if (10) has a solution $\pi \geqq 0$, (8) is clearly infeasible. The system of linear inequalities (8) is said to be **inconsistent** iff the fundamental inconsistent inequality (6) can be obtained as a valid linear combination of it. We will prove below, that a system of linear inequalities is infeasible iff it is inconsistent. In fact, given any system of linear constraints (consisting of equations and/or inequalities) we will prove that it has no feasible solution iff the fundamental inconsistent inequality (6) can be obtained as a valid linear combination of it. This leads to a theorem of alternatives for that system. These theorems of alternatives can be proven in several ways. One way is by using the duality theorem of linear programming (see [2.26]). Another way is to prove them directly using a lemma proved by A. W. Tucker. We first discuss this Tucker's lemma [see A 10].

Theorem 2 (Tucker's Lemma). *If A is a given $m \times n$ real matrix, the systems*

$$Ax \geqq 0 \tag{11}$$

$$\pi A = 0, \ \pi \geqq 0 \tag{12}$$

where $x = (x_1, \ldots, x_n)^T$ and $\pi = (\pi_1, \ldots, \pi_m)$, have feasible solutions $\overline{x}$, $\overline{\pi}$ respectively, satisfying

$$(\overline{\pi})^T + A\overline{x} > 0. \tag{13}$$

Proof. We will first prove that there exist feasible solutions x^1, $\pi^1 = (\pi_1^1, \ldots, \pi_m^1)$ to (11), (12) respectively, satisfying

$$A_1.x^1 + \pi_1^1 > 0. \tag{14}$$

The proof is by induction on m, the number of rows in the matrix A. If $m = 1$ let

$$\begin{aligned} &\pi^1 = (\pi_1^1) = (1), \ x^1 = 0, \ \text{ if } A_1. = 0 \\ &\pi^1 = 0, \ x^1 = (A_1.)^T, \ \text{ if } A_1. \neq 0 \end{aligned}$$

and verify that these solutions satisfy (14). So the theorem is true if $m = 1$. We now set up an induction hypothesis.

Induction Hypothesis. If D is any real matrix of order $(m-1) \times n$, there exist vectors $x = (x_j) \in \mathbf{R}^n$, $u = (u_1, \ldots, u_{m-1})$ satisfying: $Dx \geqq 0$; $uD = 0$, $u \geqq 0$; $u_1 + D_1.x > 0$.

Under the induction hypothesis we will now prove that this result also holds for the matrix A of order $m \times n$. Let A' be the $(m-1) \times n$ matrix obtained by deleting the last row $A_m.$ from A. Applying the induction hypothesis on A', we know that there exist $x' \in \mathbf{R}^n$, $u' = (u'_1, \ldots, u'_{m-1})$ satisfying

$$A'x' \geqq 0;\ u'A' = 0,\ u' \geqq 0;\ u'_1 + A_1.x' > 0. \tag{15}$$

If $A_m.x' \geqq 0$, define $x^1 = x'$, $\pi^1 = (u', 0)$, and verify that x^1, π^1 are respectively feasible to (11), (12) and satisfy (14), by (15). On the other hand, suppose $A_m.x' < 0$. We now attempt to find a vector $\tilde{x} \in \mathbf{R}^n$ and real number α such that

$$x^1 = \tilde{x} + \alpha x' \text{ and vector } \pi^1$$

together satisfy (14). We have to determine $\tilde{x}$, α, π^1 so that this will be true. For this we require

$$A_m.x^1 = A_m.\tilde{x} + \alpha A_m.x' \geqq 0 \text{ that is}$$
$$\alpha \leqq (A_m.\tilde{x})/(-A_m.x').$$

So it suffices if we define $\alpha = (A_m.\tilde{x})/(-A_m.x')$. We still have to determine $\tilde{x}$ and π^1 appropriately. The vector x^1 should also satisfy for $i = 1$ to $m-1$

$$A_i.x^1 = A_i.\tilde{x} + \alpha A_i.x' = (A_i. + \lambda_i A_m.)\tilde{x} \geqq 0$$

where $\lambda_i = (A_i.x')/(-A_m.x')$. Now define $B_i. = A_i. + \lambda_i A_m.$, for $i = 1$ to $m-1$ and let B be the $(m-1) \times n$ matrix whose rows are $B_i.$, $i = 1$ to $m-1$. By applying the induction hypothesis on B, we know that there exists $x'' \in \mathbf{R}^n$, $u'' = (u''_1, \ldots, u''_{m-1})$ satisfying

$$Bx'' \geqq 0,\ u''B = 0,\ u'' \geqq 0,\ u''_1 + B_1.x'' > 0. \tag{16}$$

We take this vector x'' to be the $\tilde{x}$ we are looking for, and therefore define

$$x^1 = x'' - x'(A_m.x'')/(A_m.x')$$
$$\pi^1 = \Big(u'', \sum_{i=1}^{m} \lambda_i u''_i\Big).$$

Using (15), (16) and the fact that $A_m.x' < 0$ in this case, verify that x^1, π^1 are respectively feasible to (11) and (12) and satisfy (14). So under the induction hypothesis, the result in the induction hypothesis also holds for the matrix A of order $m \times n$. The result in the induction hypothesis has already been verified to be true for matrices with 1 row only. So, by induction, we conclude that there exist feasible solutions x^1, π^1 to (11), (12) respectively, satisfying (14).

For any $i = 1$ to m, the above argument can be used to show that there exist feasible solutions x^i, $\pi^i = (\pi^i_1, \ldots, \pi^i_m)$ to (11) and (12) respectively satisfying

$$\pi^i_i + A_i.x^i > 0. \tag{17}$$

Define $\overline{x} = \sum_{i=1}^m x^i$, $\overline{\pi} = \sum_{i=1}^m \pi^i$, and verify that $\overline{x}$, $\overline{\pi}$ together satisfy (11) and (12) and (13).

□

Corollary 1. *Let A, D be matrices of orders $m_1 \times n$ and $m_2 \times n$ respectively with $m_1 \geqq 1$. Then there exist $x = (x_1, \ldots, x_n)^T$, $\pi = (\pi_1, \ldots, \pi_{m_1})$. $\mu = (\mu_1, \ldots, \mu_{m_2})$ satisfying*

(18)	(19)	(20)
$Ax \geqq 0$	$\pi A + \mu D = 0$	$\pi^T + Ax > 0\,.$
$Dx = 0$	$\pi \geqq 0$	

Proof. Applying Tucker's lemma to the systems

(21)	(22)
$Ax \geqq 0$	$\pi A + \gamma D - \nu D = 0$
$Dx \geqq 0$	$\pi, \gamma, \nu \geqq 0$
$-Dx \geqq 0$	

we know that there exist $\overline{x}$, $\overline{\pi}$, $\overline{\gamma}$, $\overline{\nu}$ feasible to them, satisfying $\overline{\pi}^T + A\overline{x} > 0$. Verify that $\overline{x}$, $\overline{\pi}$, $\overline{\mu} = \overline{\gamma} - \overline{\nu}$ satisfy (18), (19) and (20).

□

We will now discuss some of the most useful theorems of alternatives for linear systems of constraints.

Theorem 3 (Farkas' Theorem). *Let A, b be given matrices of orders $m \times n$ and $m \times 1$ respectively. Let $x = (x_1, \ldots, x_n)^T$, $\pi = (\pi_1, \ldots, \pi_m)$. Exactly one of the following two systems (I), (II) is feasible.*

(I)	(II)
$Ax = b$	$\pi A \leqq 0$
$x \geqq 0$	$\pi b \ > 0.$

Proof. Suppose both systems are feasible. Let $\overline{x}$ be feasible to (I) and $\overline{\pi}$ be feasible to (II). Then $(\overline{\pi}A)\overline{x} \leqq 0$ since $\overline{\pi}A \leqq 0$ and $\overline{x} \geqq 0$. Also $\overline{\pi}(A\overline{x}) = \overline{\pi}b > 0$. So there is a contradiction. So it is impossible for both systems (I) and (II) to be feasible.

Suppose (II) is infeasible. Let $y = \pi^T$. So this implies that in every solution of

$$\begin{pmatrix} b^T \\ -A^T \end{pmatrix} \quad y \geqq 0 \tag{23}$$

the first constraint always holds as an equation. By Tucker's lemma (Theorem 2) there exists a $\overline{y}$ feasible to (23) and $(\overline{\delta}, \overline{\mu}_1, \ldots, \overline{\mu}_n) \geqq 0$ feasible to

$$(\overline{\delta}, \overline{\mu}_1, \ldots, \overline{\mu}_n) \begin{pmatrix} b^T \\ -A^T \end{pmatrix} = 0 \tag{24}$$

which together satisfy $b^T\overline{y} + \overline{\delta} > 0$. But since $\overline{y}$ is feasible to (23) we must have $b^T\overline{y} = 0$ as discussed above (since (II) is infeasible) and so $\overline{\delta} > 0$. Define $\overline{x}_j = \overline{\mu}_j/\overline{\delta}$ for $j = 1$ to n and let $\overline{x} = (\overline{x}_1, \ldots, \overline{x}_n)^T$. From (24) we verify that $\overline{x}$ is feasible to (I). So if (II) is infeasible, (I) is feasible. Thus exactly one of the two systems (I), (II) is feasible.

□

Note 1. Given A, b, the feasibility of system (I) in Farkas' theorem can be determined using Phase I of the Simplex Method for linear programming problems. If (I) is feasible, Phase I terminates with a feasible solution of (I), in this case system (II) has no feasible solution. If Phase I terminates with the conclusion that (I) is infeasible, the Phase I dual solution at termination provides a vector π which is feasible to system (II).

Note 2. Theorem 3, Farkas' theorem, is often called Farkas' lemma in the literature.

An Application of Farkas' Theorem to Derive Optimality Conditions for LP

To illustrate an application of Farkas' theorem, we will now show how to derive the necessary optimality conditions for a linear program using it. Consider the LP

$$\begin{array}{ll} \text{minimize} & f(x) = cx \\ \text{subject to} & Ax \geqq b \end{array} \tag{25}$$

where A is a matrix of order $m \times n$. The constraints in (25) include all the conditions in the problem, including any bound restrictions, lower or upper, on individual variables. If there are any equality constraints in the problem, each of them can be represented by the corresponding pair of opposing inequality constraints and expressed in the form given in (25) (for example, the equality constraint $x_1 + x_2 - x_3 = 1$ is equivalent to the pair of inequality constraints $x_1 + x_2 - x_3 \geqq 1$, $-x_1 - x_2 + x_3 \geqq -1$). Thus every linear program can be expressed in this form. We now state the necessary optimality conditions for a feasible solution $\overline{x}$ to be optimal to this LP, and prove it using Farkas' theorem.

Theorem 4. *If $\overline{x}$ is a feasible solution for (25), and $\overline{x}$ is optimal to (25), there must exist a vector $\overline{\Pi} = (\overline{\Pi}_1, \ldots, \overline{\Pi}_m)$ which together with $\overline{x}$ satisfies*

$$\begin{aligned} c - \overline{\Pi}A &= 0 \\ \overline{\Pi} &\geqq 0 \\ \overline{\Pi}_i(A_{i.}\overline{x} - b_i) &= 0, \ i = 1 \text{ to } m. \end{aligned} \tag{26}$$

Proof. Consider the case $c = 0$ first. In this case the objective value is a constant, zero, and hence every feasible solution of (25) is optimal to it. It can be verified that $\overline{\Pi} = 0$ satisfies (26) together with any feasible solution $\overline{x}$ for (25).

Now consider the case $c \neq 0$. We claim that the fact that $\overline{x}$ is optimal to (25) implies that $A\overline{x} \not> b$ in this case. To prove this claim, suppose $A\overline{x} > b$. For any $y \in \mathbf{R}^n$, $A(\overline{x} + \alpha y) = A\overline{x} + \alpha Ay \geqq b$ as long as α is sufficiently small, since $A\overline{x} > b$.

Take $y = -c^T$. Then, for $\alpha > 0$, $c(\overline{x} + \alpha y) < c\overline{x}$ and $\overline{x} + \alpha y$ is feasible to (25) as long as α is positive and sufficiently small, contradicting the optimality of $\overline{x}$ to (25). So, if $\overline{x}$ is optimal to (25) in this case ($c \neq 0$) at least one of the constraints in (25) must hold as an equation at $\overline{x}$.

Rearrange the rows of A, and let A_1 be the matrix of order $m_1 \times n$ consisting of all the rows in A corresponding to constraints in (25) which hold as equations in (25) when $x = \overline{x}$, and let A_2, of order $m_2 \times n$, be the matrix consisting of all the other rows of A.

By the above argument A_1 is nonempty, that is, $m_1 \geqq 1$. Let b^1, b^2 be the corresponding partition of b. So

$$A = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix}, \quad b = \begin{pmatrix} b^1 \\ b^2 \end{pmatrix} \tag{27}$$

and

$$\begin{aligned} A_1\overline{x} &= b^1 \\ A_2\overline{x} &> b^2. \end{aligned} \tag{28}$$

We now show that if $\overline{x}$ is optimal to (25), the system

$$\begin{aligned} A_1 y &\geqq 0 \\ cy &< 0 \end{aligned} \tag{29}$$

cannot have a solution y. Suppose not. Let $\overline{y}$ be a solution for (29). Then for $\alpha > 0$, $A_1(\overline{x} + \alpha\overline{y}) = A_1\overline{x} + \alpha A_1\overline{y} \geqq b^1$; and $A_2(\overline{x} + \alpha\overline{y}) = A_2\overline{x} + \alpha A_2\overline{y} \geqq b^2$ as long as α is sufficiently small, since $A_2\overline{x} > b^2$. So when α is positive but sufficiently small, $\overline{x} + \alpha\overline{y}$ is feasible to (25) and since $c(\overline{x} + \alpha\overline{y}) = c\overline{x} + \alpha c\overline{y} < c\overline{x}$, since $c\overline{y} < 0$, we have a contradiction to the optimality of $\overline{x}$ for (25).

So, (29) has no solution y. By taking transposes, we can put (29) in the form of system (II) under Theorem 3 (Farkas' theorem). Writing the corresponding system (I) and taking transposes again, we conclude that since (29) has no solution, there exists a row vector Π^1 satisfying

$$\begin{aligned} \Pi^1 A_1 &= c \\ \Pi^1 &\geqq 0 \end{aligned} \tag{30}$$

Define $\Pi^2 = 0$ and let $\overline{\Pi} = (\Pi^1, \Pi^2)$. From the fact that A_1, A_2 is a partition of A as in (27), and using (30), (28), we verify that $\overline{\Pi} = (\Pi^1, \Pi^2)$ satisfies (26) together with $\overline{x}$.

□

Example 2

Consider the LP

$$\begin{array}{lrcl} \text{minimize} & f(x) = -3x_1 + x_2 + 3x_3 + 5x_5 & & \\ \text{subject to} & x_1 + x_2 - x_3 + 2x_4 - x_5 & \geqq & 5 \\ & -2x_1 + 2x_3 - x_4 + 3x_5 & \geqq & -8 \\ & x_1 & \geqq & 6 \\ & -3x_2 + 3x_4 & \geqq & -5 \\ & 5x_3 - x_4 + 7x_5 & \geqq & 7. \end{array}$$

Let $\overline{x} = (6, 0, -1, 0, 2)^T$. Verify that $\overline{x}$ satisfies constraints 1,2 and 3 in the problem as equations and the remaining as strict inequalities. We have

$$A_1 = \begin{pmatrix} 1 & 1 & -1 & 2 & -1 \\ -2 & 0 & 2 & -1 & 3 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix},$$

$$b^1 = \begin{pmatrix} 5 \\ -8 \\ 6 \end{pmatrix}$$

$$A_2 = \begin{pmatrix} 0 & -3 & 0 & 3 & 0 \\ 0 & 0 & 5 & -1 & 7 \end{pmatrix},$$

$$b^2 = \begin{pmatrix} -5 \\ 7 \end{pmatrix}$$

$$c = (-3, 1, 3, 0, 5), \quad f(x) = cx$$

and $A_1\overline{x} = b^1$, $A_2\overline{x} > b^2$. If we take $\Pi^1 = (1, 2, 0)$ then $\Pi^1 A_1 = c$, $\Pi^1 \geqq 0$. Let $\Pi^2 = (0, 0)$, and $\overline{\Pi} = (\Pi^1, \Pi^2) = (1, 2, 0, 0, 0)$. These facts imply that $\overline{\Pi}$, $\overline{x}$ together satisfy the necessary optimality conditions (26) for this LP.

We leave it to the reader to verify that if $\overline{x}$ is feasible to (25), and there exists a vector $\overline{\Pi}$ such that $\overline{x}$, $\overline{\Pi}$ together satisfy (26), then $\overline{x}$ is in fact optimal to (25), from first principles. Thus the conditions (26) and feasibility are together necessary and sufficient optimality conditions for the LP (25). It can also be verified that any $\overline{\Pi}$ satisfying (26) is an optimum dual solution associated with the LP (25); and that (26) are in fact the dual feasibility and complementary slackness optimality conditions for the LP (25). See [2.26, A10]. Thus Farkas' theorem leads directly to the optimality conditions for the LP (25). Later on, in Appendix 4, we will see that Theorems of alternatives like Farkas' theorem and others discussed below are very useful for deriving optimality conditions in nonlinear programming too. We will now discuss some more theorems of alternatives.

Some Other Theorems of Alternatives

Theorem 5 (Motzkin's Theorem of the Alternatives). *Let $m \geqq 1$, and let A, B, C be given matrices of orders $m \times n$, $m_1 \times n$, $m_2 \times n$. Let $x = (x_1, \ldots, x_n)^T$, $\pi = (\pi_1, \ldots, \pi_m)$, $\mu = (\mu_1, \ldots, \mu_{m_1})$, $\gamma = (\gamma_1, \ldots, \gamma_{m_2})$. Then exactly one of the following two systems (I), (II) is feasible.*

(I)	(II)
$Ax > 0$	$\pi A + \mu B + \gamma C = 0$
$Bx \geqq 0$	$\pi \geq 0,\ \mu \geqq 0$
$Cx = 0$	

Proof. As in the proof of Theorem 3, it can be verified that if both (I), (II) are feasible, there is a contradiction. Suppose system (I) is infeasible. This implies that

every feasible solution of

$$\begin{aligned} Ax &\geqq 0 \\ Bx &\geqq 0 \\ Cx &= 0 \end{aligned} \tag{31}$$

satisfies $A_i.x = 0$ for at least one $i = 1$ to m. By Corollary 1, there exists $\overline{x}$ feasible to (31) and $\overline{\pi}$, $\overline{\mu}$, $\overline{\gamma}$ feasible to

$$\begin{aligned} \overline{\pi}A + \overline{\mu}B + \overline{\gamma}C &= 0 \\ \overline{\pi} \geqq 0,\ \overline{\mu} &\geqq 0 \end{aligned} \tag{32}$$

satisfying $(\overline{\pi})^T + A\overline{x} > 0$. But since $\overline{x}$ is feasible to (31), $A_i.\overline{x} = 0$ for at least one i as discussed above. This implies that for that i, $\overline{\pi}_i > 0$, that is, $\overline{\pi} \geq 0$. So $(\overline{\pi}, \overline{\mu}, \overline{\gamma})$ satisfies (II). So if (I) is infeasible, (II) is feasible. Thus exactly one of the two systems (I), (II) is feasible.

□

Theorem 6 (Gordan's Theorem of the Alternatives). *Give a matrix A of order $m \times n$, exactly one of the following systems (I) and (II) is feasible.*

(I)	(II)
$Ax > 0$	$\pi A = 0$
	$\pi \quad \geq 0$

Proof. Follows from Theorem 5 by selecting B, $C = 0$ there.

□

Theorem 7 (Tucker's Theorem of the Alternatives). *Let $m \geqq 1$, and let A, B, C be given matrices of orders $m \times n$, $m_1 \times n$, $m_2 \times n$ respectively. Let $x = (x_1, \ldots, x_n)^T$, $\pi = (\pi_1, \ldots, \pi_m)$, $\mu = (\mu_1, \ldots, \mu_{m_1})$, $\gamma = (\gamma_1, \ldots, \gamma_{m_2})$. Exactly one of the following systems (I), (II) is feasible.*

(I)	(II)
$Ax \geq 0$	$\pi A + \mu B + \gamma C = 0$
$Bx \geqq 0$	$\pi > 0,\ \mu \geqq 0$
$Cx = 0$	

Proof. As in the proof of Theorem 3, it can be verified that if both (I), (II) are feasible, there is a contradiction. Suppose that (I) is infeasible. This implies that every feasible solution of

$$\begin{aligned} Ax &\geqq 0 \\ Bx &\geqq 0 \\ Cx &= 0 \end{aligned} \tag{33}$$

must satisfy $Ax = 0$. By Corollary 1, there exists $\overline{x}$ feasible to (33) and $\overline{\pi}$, $\overline{\mu}$, $\overline{\gamma}$ feasible to

$$\begin{aligned} \overline{\pi}A + \overline{\mu}B + \overline{\gamma}C &= 0 \\ \pi \geqq 0,\ \mu &\geqq 0 \end{aligned} \tag{34}$$

satisfying $(\overline{\pi})^T + A\overline{x} > 0$. But since $\overline{x}$ is feasible to (33), $A\overline{x} = 0$ as discussed above; so $(\overline{\pi})^T > 0$. So $(\overline{\pi}, \overline{\mu}, \overline{\gamma})$ satisfies (II). So if (I) is infeasible, (II) is feasible. Thus exactly one of the two systems (I), (II) is feasible.

□

Theorem 8 (Gale's Theorem of Alternatives). *Let A, b be given matrices of orders $m \times n$, $m \times 1$ respectively. Let $x = (x_1, \ldots, x_n)^T$, $\pi = (\pi_1, \ldots, \pi_m)$. Exactly one of the following systems (I), (II) is feasible.*

(I)	(II)
$Ax \geqq b$	$\pi A = 0$
	$\pi b = 1$
	$\pi \geqq 0$

Proof. System (I) is equivalent to

$$\begin{aligned} (\,A \quad -b\,) \begin{pmatrix} x \\ x_{n+1} \end{pmatrix} &\geqq 0 \\ d \begin{pmatrix} x \\ x_{n+1} \end{pmatrix} &> 0 \end{aligned} \tag{35}$$

where $d = (0, 0, \ldots, 0, 1) \in \mathbf{R}^{n+1}$. (I) is equivalent to (35) in the sense that if a solution of one of these systems is given, then a solution of the other system in the pair can be constructed from it. For example if $\overline{x}$ is a feasible solution of (I), then $(\overline{x}, \overline{x}_{n+1} = 1)$ is a feasible solution of (35). Conversely, if $(\hat{x}, \hat{x}_{n+1})$ is a feasible solution of (35), then $\hat{x}_{n+1} > 0$ and $(1/\hat{x}_{n+1})\hat{x}$ is a feasible solution of (I).

This theorem follows from Theorem 5 applied to (35).

□

For a complete discussion of several other Theorems of alternatives for linear systems and their geometric interpretation, see O. L. Mangasarian's book [A10].

Exercises

---

1. Let $\mathbf{K}$ be the set of feasible solutions of

$$\sum_{j=1}^{n} a_{ij}x_j \geqq b_i,\ i = 1 \text{ to } m. \tag{36}$$

Assume that $\mathbf{K} \neq \emptyset$. Prove that all $x \in \mathbf{K}$ satisfy

$$\sum_{j=1}^{n} c_j x_j \geqq d \tag{37}$$

iff, for some $\alpha \geqq d$, the inequality $\sum_{j=1}^{n} c_j x_j \geqq \alpha$ is a valid linear combination of the constraints in (36), that is, iff there exists $\overline{\pi} = (\overline{\pi}_1, \ldots, \overline{\pi}_m) \geqq 0$, satisfying $c_j = \sum_{i=1}^{m} \overline{\pi}_i a_{ij}$, $j = 1$ to n, and $\alpha = \sum_{i=1}^{m} \overline{\pi}_i b_i$.

2. Let M be a square matrix of order n. Prove that for each $q \in \mathbf{R}^n$, the system "$Mx + q \geqq 0$, $x \geqq 0$" has a solution $x \in \mathbf{R}^n$ iff the system "$My > 0$, $y \geqq 0$" has a solution y. (O. L. Mangasarian [3.42])

3. Let M be a square matrix of order n and $q \in \mathbf{R}^n$. Prove that the following are equivalent

i) the system $Mx + q > 0$, $x \geqq 0$ has a solution $x \in \mathbf{R}^n$,
ii) the system $Mx + q > 0$, $x > 0$ has a solution $x \in \mathbf{R}^n$,
iii) the system $M^T u \leqq 0$, $q^T u \leqq 0$, $0 \leq u$ has no solution $u \in \mathbf{R}^n$.

(O. L. Mangasarian [3.42])

4. Prove that (36) is infeasible iff it is inconsistent (that is, the fundamental inconsistent inequality (6) can be obtained as a valid linear combination of it) as a corollary of the result in Exercise 1.

5. Let A be an $m \times n$ matrix, and suppose the system: $Ax = b$, has at least one solution; and the equation $cx = d$ holds at all solutions of the system $Ax = b$. Then prove that the equation $cx = d$ can be obtained as a linear combination of equations from the system $Ax = b$. That is, there exists $\pi = (\pi_1, \ldots, \pi_m)$, such that $c = \pi A$ and $d = \pi b$.

---

2. CONVEX SETS

A subset $\mathbf{K} \subset \mathbf{R}^n$ is said to be convex if $x^1, x^2 \in \mathbf{K}$ implies that $\alpha x^1 + (1 - \alpha)x^2 \in \mathbf{K}$ for all $0 \leqq \alpha \leqq 1$. Thus, a subset of $\mathbf{R}^n$ is convex iff given any pair of points in it, the entire line segment connecting these two points is in the set. See Figures 1, 2.

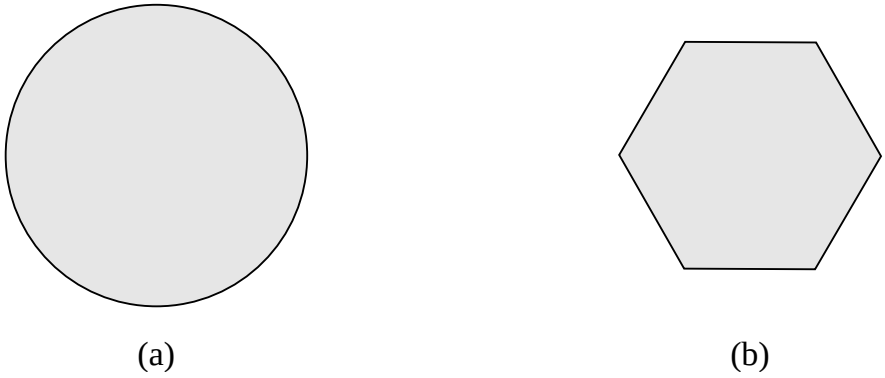


Figure 1 Convex sets. (a) All points inside or on the circle. (b) All points inside or on the polygon.

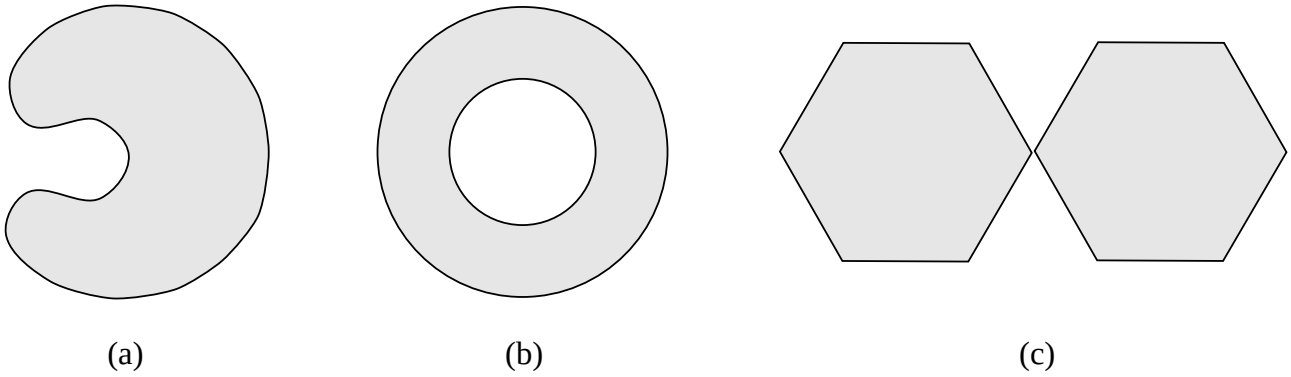


Figure 2 Non-convex sets. (a) All points inside or on the cashew nut. (b) All points on or between two circles. (c) All points on at least one of the two polygons.

2.1 Convex Combinations, Convex Hull

Let $(x^1, \ldots, x^r\}$ be any finite set of points in $\mathbf{R}^n$. A convex combination of this set is a point of the form

$$\alpha_1 x^1 + \ldots + \alpha_r x^r, \text{ where } \alpha_1 + \ldots + \alpha_r = 1 \text{ and } \alpha_1, \ldots, \alpha_r \geqq 0.$$

The set of all convex combinations of $\{x^1, \ldots, x^r\}$ is known as the convex hull of $\{x^1, \ldots, x^r\}$.

Given $\boldsymbol{\Delta} \subset \mathbf{R}^n$, the convex hull of $\boldsymbol{\Delta}$ is the set consisting of all convex combinations of all finite sets of points from $\boldsymbol{\Delta}$.

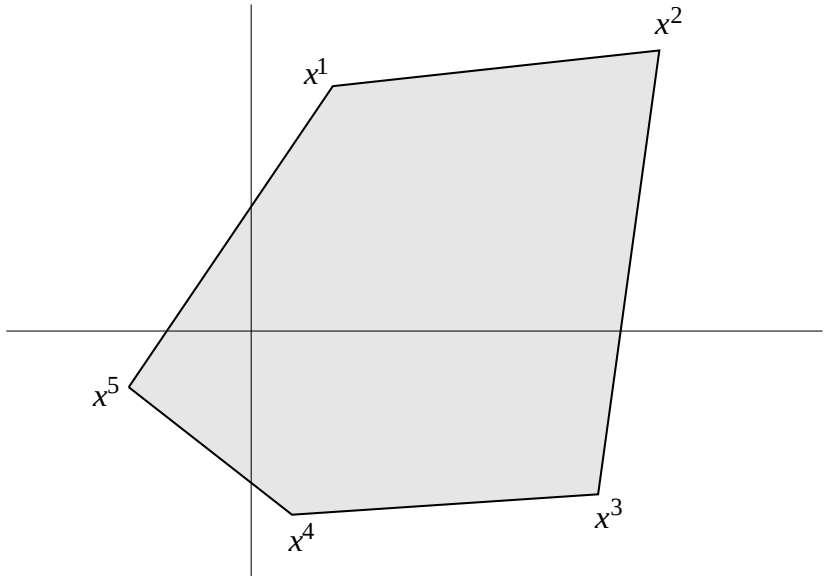


Figure 3 Convex hull of $\{x^1, \ldots, x^5\}$ in $\mathbf{R}^2$.

The following results can be verified to be true:

1. $\mathbf{K} \subset \mathbf{R}^n$ is convex iff for any finite number r, given $x^1, \ldots, x^r \in \mathbf{K}$, $\alpha_1 x^1 + \ldots + \alpha_r x^r \in \mathbf{K}$ for all $\alpha_1, \ldots, \alpha_r$ satisfying $\alpha_1 + \ldots + \alpha_r = 1$, $\alpha_1 \geqq 0, \ldots, \alpha_r \geqq 0$.
2. The intersection of any family of convex subsets of $\mathbf{R}^n$ is convex. The union of two convex sets may not be convex.
3. The set of feasible solutions of a system of linear constraints

$$\begin{aligned} A_{i.}x &= b_i, \ i = 1 \text{ to } m \\ &\geqq b_i, \ i = m+1 \text{ to } m+p \end{aligned}$$

   is convex. A convex set like this is known as a **convex polyhedron**. A bounded convex polyhedron is called a **convex polytope**.
4. The set of feasible solutions of a homogeneous system of linear inequalities in $x \in \mathbf{R}^n$,

$$Ax \geqq 0 \tag{38}$$

   is known as a **convex polyhedral cone**. Given a convex polyhedral cone, there exists a finite number of points $x^1, \ldots, x^s$ such that the cone is $\{x : x = \alpha_1 x^1 + \ldots + \alpha_s x^s, \alpha_1 \geqq 0, \ldots, \alpha_s \geqq 0\} = \text{Pos}\{x^1, \ldots, x^s\}$. The polyhedral cone which is the set of feasible solutions of (38) is said to be a **simplicial cone** if A is a nonsingular square matrix. Every simplicial cone of dimension n is of the form $\text{Pos}\{B_{.1}, \ldots, B_{.n}\}$ where $\{B_{.1}, \ldots, B_{.n}\}$ is a basis for $\mathbf{R}^n$.
5. Given two convex subsets of $\mathbf{R}^n$, $\mathbf{K}_1$, $\mathbf{K}_2$, their sum, denoted by $\mathbf{K}_1 + \mathbf{K}_2 = \{x+y : x \in \mathbf{K}_1, y \in \mathbf{K}_2\}$ is also convex.

Separating Hyperplane Theorems

Given two nonempty subsets $\mathbf{K}_1$, $\mathbf{K}_2$ of $\mathbf{R}^n$, the hyperplane $\mathbf{H} = \{x : cx = \alpha\}$ is said to separate $\mathbf{K}_1$ and $\mathbf{K}_2$ if $cx - \alpha$ has the same sign for all $x \in \mathbf{K}_1$, say $\geqq 0$, and the opposite sign for all $x \in \mathbf{K}_2$, that is, if

$$\begin{aligned} cx &\geqq \alpha \text{ for all } x \in \mathbf{K}_1 \\ &\leqq \alpha \text{ for all } x \in \mathbf{K}_2. \end{aligned}$$

Here we will prove that if two convex subsets of $\mathbf{R}^n$ are disjoint, there exists a separating hyperplane for them. See Figures 4, 5.

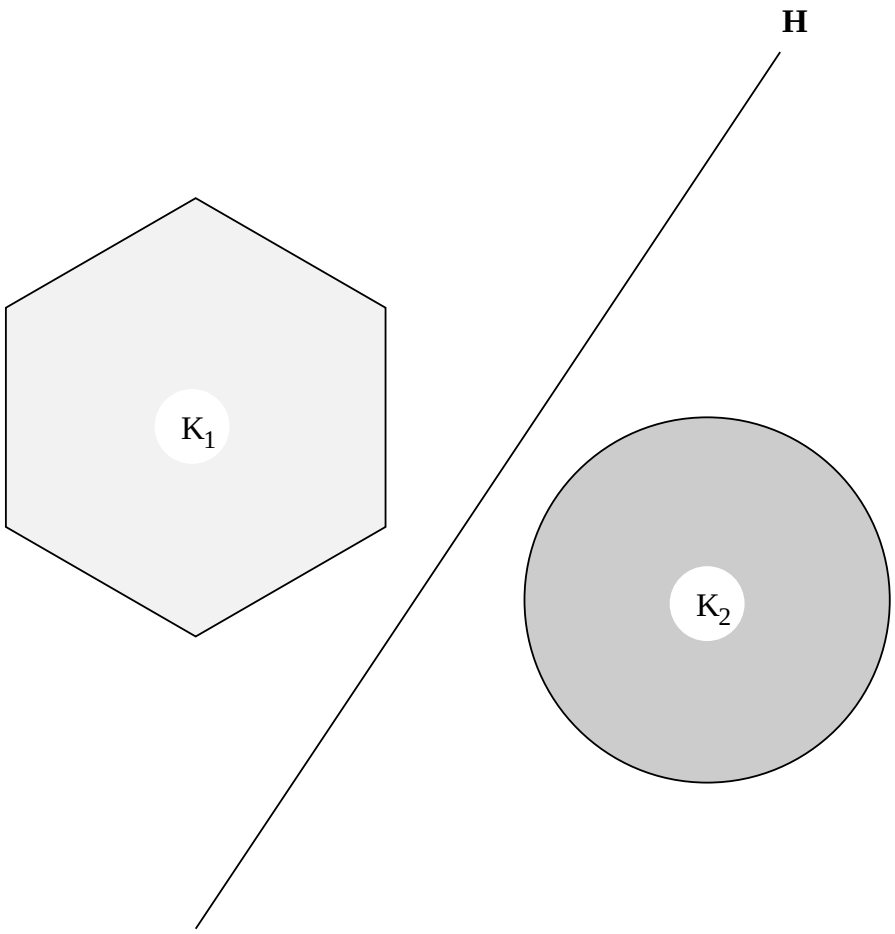


Figure 4 The hyperplane $\mathbf{H}$ separates the two disjoint convex sets $\mathbf{K}_1$ and $\mathbf{K}_2$.

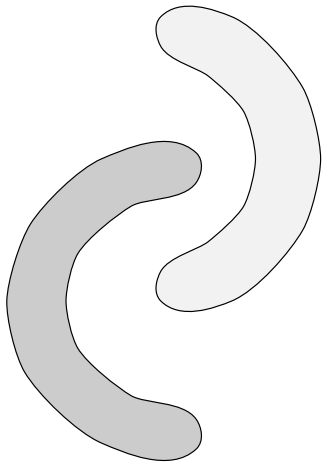


Figure 5 Even though the two cashew nuts (both nonconvex) are disjoint, they cannot be separated by a hyperplane.

Theorem 9. *Let* $\mathbf{K}$ *be a nonempty closed convex subset of* $\mathbf{R}^n$ *and* $0 \notin \mathbf{K}$. *Then there exists a hyperplane containing the origin separating it from* $\mathbf{K}$.

Proof. Take any point $\hat{x} \in \mathbf{K}$, and let $\mathbf{E} = \{x : \|x\| \leqq \|\hat{x}\|\}$. Since $0 \notin \mathbf{K}$, $\hat{x} \neq 0$, and hence $\mathbf{E}$ is a nonempty ball. Let $\mathbf{\Gamma} = \mathbf{E} \cap \mathbf{K}$. $\mathbf{\Gamma}$ is a bounded closed convex subset of $\mathbf{R}^n$, not containing the origin. The problem: minimize $\|x\|$ over $x \in \mathbf{\Gamma}$, has an optimum solution, since a continuous function attains its minimum on a compact set. We will show that this problem has a unique optimum solution. Suppose not. Let $x^1, x^2 \in \mathbf{\Gamma}$, $x^1 \neq x^2$, minimize $\|x\|$ over $x \in \mathbf{\Gamma}$. Let $x^3 = (x^1 + x^2)/2$. By Cauchy-Schwartz inequality $|(x^1)^T x^2| \leqq \|x^1\| \cdot \|x^2\|$ with equality holding iff $x^1 = \lambda x^2$ for some real number λ. So $\|x^3\|^{\mathbf{2}} = (\|x^1\|^{\mathbf{2}} + \|x^2\|^{\mathbf{2}} + 2|(x^1)^T x^2|)/4 \leqq (\|x^1\|^{\mathbf{2}} + \|x^2\|^{\mathbf{2}} + 2\|x^1\| \cdot \|x^2\|)/4$ by Cauchy-Schwartz inequality. Let $\|x^1\| = \gamma$. So $\|x^2\| = \gamma$ also, since both x^1, x^2 minimize $\|x\|$ over $x \in \mathbf{\Gamma}$. So, from the above, we have $\|x^3\|^{\mathbf{2}} \leqq \gamma^{\mathbf{2}}$. Since $x^3 \in \mathbf{\Gamma}$ and $\gamma^{\mathbf{2}}$ is the minimum of $\|x\|^{\mathbf{2}}$ over $x \in \mathbf{\Gamma}$, $\|x^3\|^{\mathbf{2}} \leqq \gamma^{\mathbf{2}}$ implies that $\|x^3\|^{\mathbf{2}} = \gamma^{\mathbf{2}}$. By Cauchy-Schwartz inequality, this equality holds iff $x^1 = \lambda x^2$ for some scalar λ. But since $\|x^1\| = \|x^2\|$, we must have $\lambda = +1$ or -1. If $\lambda = -1$, $x^3 = 0$, and this contradicts the fact that $0 \notin \mathbf{\Gamma}$. So $\lambda = +1$, that is, $x^1 = x^2$. So the problem of minimizing $\|x\|$ over $x \in \mathbf{\Gamma}$, has a unique optimum solution, say $\overline{x}$. We will now prove that

$$(x - \overline{x})^T \overline{x} \geqq 0 \text{ for all } x \in \mathbf{K}. \tag{39}$$

$\overline{x}$ minimizes $\|x\|$ over $x \in \mathbf{\Gamma}$, and from the definition of $\mathbf{\Gamma}$, it is clear that $\overline{x}$ also minimizes $\|x\|$ over $x \in \mathbf{K}$. Let $x \in \mathbf{K}$. By convexity of $\mathbf{K}$, $\overline{x} + \alpha(x - \overline{x}) \in \mathbf{K}$ for all $0 \leqq \alpha \leqq 1$. So $\|\overline{x} + \alpha(x - \overline{x})\|^{\mathbf{2}} \geqq \|\overline{x}\|^{\mathbf{2}}$ for all $0 \leqq \alpha \leqq 1$. That is, $\alpha^{\mathbf{2}}\|x - \overline{x}\|^{\mathbf{2}} + 2\alpha(x - \overline{x})^T \overline{x} \geqq 0$ for all $0 \leqq \alpha \leqq 1$. So for $0 < \alpha \leqq 1$, we have $\alpha\|x - \overline{x}\|^{\mathbf{2}} + 2(x - \overline{x})^T \overline{x} \geqq 0$. Making α approach zero through positive values, this implies (39).

Conversely, if $\overline{x} \in \mathbf{K}$ satisfies (39), then for any $x \in \mathbf{K}$, $\|x\|^2 = \|(x - \overline{x}) + \overline{x}\|^2 = \|x - \overline{x}\|^2 + \|\overline{x}\|^2 + 2(x - \overline{x})^T\overline{x} \geqq \|\overline{x}\|^2$ (by (39)), and this implies that $\overline{x}$ minimizes $\|x\|$ over $x \in \mathbf{K}$. Thus (39) is a necessary and sufficient optimality condition for the problem of minimizing $\|x\|$ over $x \in \mathbf{K}$.

Since $0 \notin \mathbf{K}$, $\overline{x} \neq 0$. From (39) we have $(\overline{x})^T x \geqq \|\overline{x}\|^2 > 0$ for all $x \in \mathbf{K}$. So the hyperplane $\{x : (\overline{x})^T x = 0\}$ through the origin separates $\mathbf{K}$ from 0.

□

Theorem 10. *Let $\mathbf{K}$ be a nonempty convex subset of $\mathbf{R}^n$, $b \notin \mathbf{K}$. Then $\mathbf{K}$ can be separated from b by a hyperplane.*

Proof. If $\mathbf{K}$ is a closed convex subset, by translating the origin to b and using Theorem 9 we conclude that $\mathbf{K}$ and b can be separated by a hyperplane.

If $\mathbf{K}$ is not closed, let $\overline{\mathbf{K}}$ be the closure of $\mathbf{K}$. If $b \notin \overline{\mathbf{K}}$, then again by the previous result b and $\overline{\mathbf{K}}$ can be separated by a hyperplane, which also separates b and $\mathbf{K}$.

So assume that $b \in \overline{\mathbf{K}}$. Since $b \in \overline{\mathbf{K}}$ but $\notin \mathbf{K}$, b must be a boundary point of $\mathbf{K}$. So every open neighborhood of b contains a point not in $\overline{\mathbf{K}}$. So we can get a sequence of points $\{b^r : r = 1 \text{ to } \infty\}$ such that $b^r \notin \overline{\mathbf{K}}$ for all r, and b^r converges to b as r tends to ∞. Since $b^r \notin \overline{\mathbf{K}}$, by the previous result, there exists c^r such that $c^r(x - b^r) \geqq 0$ for all $x \in \mathbf{K}$, with $\|c^r\| = 1$. The sequence of row vectors $\{c^r : r = 1, \ldots\}$ all lying on the unit sphere in $\mathbf{R}^n$ (which is a closed bounded set) must have a limit point. Let c be a limit point of $\{c^r : r = 1, 2, \ldots\}$. So $\|c\| = 1$. Let $\mathbf{S}$ be a monotonic increasing sequence of positive integers such that c^r converges to c as r tends to ∞ through $r \in \mathbf{S}$. But $c^r(x - b^r) \geqq 0$ for all $x \in \mathbf{K}$. Taking the limit in this inequality, as r tends to ∞ through $r \in \mathbf{S}$ we conclude that $c(x - b) \geqq 0$ for all $x \in \mathbf{K}$. So the hyperplane $\{x : cx = cb\}$ separates $\mathbf{K}$ from b.

□

Corollary 2. *Let $\mathbf{K}$ be a convex subset of $\mathbf{R}^n$, and let b be a boundary point of $\mathbf{K}$. Then there exists a row vector $c \neq 0$, $c \in \mathbf{R}^n$ such that $cx \geqq cb$ for all $x \in \mathbf{K}$.*

Proof. Follows from the arguments in the proof of Theorem 10.

□

The hyperplane $\{x : cx = cb\}$ in Corollary 2 is known as a **supporting hyperplane** for the convex set $\mathbf{K}$ at its boundary point b.

Theorem 11. *If $\mathbf{K}_1$, $\mathbf{K}_2$ are two mutually disjoint convex subsets of $\mathbf{R}^n$, there exists a hyperplane separating $\mathbf{K}_1$ from $\mathbf{K}_2$.*

Proof. Let $\mathbf{\Gamma} = \mathbf{K}_1 - \mathbf{K}_2 = \{x - y : x \in \mathbf{K}_1, y \in \mathbf{K}_2\}$. Since $\mathbf{K}_1$, $\mathbf{K}_2$ are convex, $\mathbf{\Gamma}$ is a convex subset of $\mathbf{R}^n$. Since $\mathbf{K}_1 \cap \mathbf{K}_2 = \emptyset$, $0 \notin \mathbf{\Gamma}$. So by Theorem 10, there exists a row vector $c \neq 0$, $c \in \mathbf{R}^n$, satisfying

$$cz \geqq 0 \text{ for all } z \in \mathbf{\Gamma}. \tag{40}$$

Let $\alpha =$ Infimum $\{cx : x \in \mathbf{K}_1\}$, $\beta =$ Supremum $\{cx : x \in \mathbf{K}_2\}$. By (40), we must

have $\alpha \geqq \beta$. So if $\gamma = (\alpha + \beta)/2$, we have

$$\begin{aligned} cx &\geqq \gamma \text{ for all } x \in \mathbf{K}_1 \\ &\leqq \gamma \text{ for all } x \in \mathbf{K}_2. \end{aligned}$$

So $x : \{cx = \gamma\}$ is a hyperplane that separates $\mathbf{K}_1$ from $\mathbf{K}_2$.

□

The theorems of alternatives discussed in Appendix 1, can be interpreted as separating hyperlane theorems about separating a point from a convex polyhedral cone not containing the point.

Exercises

---

6. Let $\mathbf{K}$ be a closed convex subset of $\mathbf{R}^n$ and $x \in \mathbf{R}^n$ and let y be the nearest point (in terms of the usual Euclidean distance) in $\mathbf{K}$ to x. Prove that $(x - y)^T(y - z) \geqq 0$ for all $z \in \mathbf{K}$. Also prove that $\|y - z\| \leqq \|x - z\|$ for all $z \in \mathbf{K}$.

7. Given sets $\mathbf{\Gamma}$, $\mathbf{\Delta}$ define $\alpha\mathbf{\Gamma} = \{\alpha x : x \in \mathbf{\Gamma}\}$ and $\mathbf{\Gamma} + \mathbf{\Delta} = \{x + y : x \in \mathbf{\Gamma}, y \in \mathbf{\Delta}\}$. Is $\mathbf{\Gamma} + \mathbf{\Gamma} = 2\mathbf{\Gamma}$? Also, when $\mathbf{\Gamma} = \{(x_1, x_2)^T : (x_1 - 1)^{\mathbf{2}} + (x_2 - 1)^{\mathbf{2}} \leqq 1\}$, $\mathbf{\Delta} = \{(x_1, x_2)^T : (x_1 + 4)^{\mathbf{2}} + (x_2 + 4)^{\mathbf{2}} \leqq 4\}$, find $\mathbf{\Gamma} + \mathbf{\Delta}$, $2\mathbf{\Gamma}$, $\mathbf{\Gamma} + \mathbf{\Gamma}$ and draw a figure in $\mathbf{R}^2$ illustrating each of these sets.

8. Prove that a convex cone in $\mathbf{R}^n$ is either equal to $\mathbf{R}^n$ or is contained in a half-space generated by a hyperplane through the origin.

9. Let $\mathbf{\Delta}_1 = \{x^1, \ldots, x^r\} \subset \mathbf{R}^n$. If $y^1, y^2 \in \mathbf{R}^n$, $y^1 \neq y^2$ are such that

$$\begin{aligned} y^1 &\in \text{ convex hull of } \{y^2\} \cup \mathbf{\Delta}_1 \\ y^2 &\in \text{ convex hull of } \{y^1\} \cup \mathbf{\Delta}_1 \end{aligned}$$

prove that both y^1 and y^2 must be in the convex hull of $\mathbf{\Delta}_1$. Using this and an induction argument, prove that if $\{y^1, \ldots, y^m\}$ is a set of distinct points in $\mathbf{R}^n$ and for each $j = 1$ to m

$$y^j \in \text{ convex hull of } \mathbf{\Delta}_1 \cup \{y^1, \ldots, y^{j-1}, y^{j+1}, \ldots, y^m\}$$

then each $y^j \in$ convex hull of $\mathbf{\Delta}_1$.

On Computing a Separating Hyperplane

Given a nonempty convex subset $\mathbf{K} \subset \mathbf{R}^n$, and a point $b \in \mathbf{R}^n$, $b \notin \mathbf{K}$, Theorem 10 guarantees that there exists a hyperplane $\mathbf{H} = \{x : cx = \alpha, c \neq 0\}$ which separates b from $\mathbf{K}$. It is a fundamental result, in mathematics such results are called **existence theorems**. This result can be proved in many different ways, and most books on convexity or optimization would have a proof for it. However, no other book seems to discuss how such a separating hyperplane can be computed, given b and $\mathbf{K}$ in some form (this essentially boils down to determining the vector c in the definition of the separating hyperplane $\mathbf{H}$), or how difficult the problem of computing it may be. For this reason, the following is very important. In preparing this, I benefitted a lot from discussions with R. Chandrasekaran.

However elegant the proof may be, an existence theorem cannot be put to practical use unless an efficient algorithm is known for computing **the thing** whose existence the theorem establishes. In order to use Theorem 10 in practical applications, we should be able to compute the separating hyperplane $\mathbf{H}$ given b and $\mathbf{K}$. Procedures to be used for constructing an algorithm to compute $\mathbf{H}$ depend very critically on the form in which the set $\mathbf{K}$ is made available to us. In practice, $\mathbf{K}$ may be specified either as the set of feasible solutions of a given system of constraints, or as the set of points satisfying a well specified set of properties, or as the convex hull of a set of points satisfying certain specified properties or constraints or those that can be obtained by a well defined constructive procedure. The difficulty of computing a separating hyperplane depends on the form in which $\mathbf{K}$ is specified.

$\mathbf{K}$ *Represented by a System of Linear Inequalities*

Consider the case, $\mathbf{K} = \{x : A_i.x \geqq d_i, i = 1 \text{ to } m\}$, where $A_i.$, d_i are given for all $i = 1$ to m. If $b \notin \mathbf{K}$, there must exist an i between 1 to m satisfying $A_i.b < d_i$. Find such an i, suppose it is r. Then the hyperplane $\{x : A_r.x = d_r\}$ separates $\mathbf{K}$ from b in this case.

$\mathbf{K}$ *Represented by a System of Linear Equations and Inequalities*

Consider the case, $\mathbf{K} = \{x : A_i.x = d_i, i = 1 \text{ to } m, \text{ and } A_i.x \geqq d_i, i = m+1 \text{ to } m+p\}$ where $A_i.$, d_i are given for all $i = 1$ to $m + p$. Suppose $b \notin \mathbf{K}$. If one of the inequality constraints $A_i.x \geqq d_i$, $i = m + 1$ to $m + p$, is violated by b, a hyperplane separating $\mathbf{K}$ from b, can be obtained from it as discussed above. If b satisfies all the inequality constraints in the definition of $\mathbf{K}$, it must violate one of the equality constraints. In this case, find an i, $1 \leqq i \leqq m$, satisfying $A_i.b \neq d_i$, suppose it is r, then the hyperplane $\{x : A_r.x = d_r\}$ separates $\mathbf{K}$ from b.

K *Represented as a Nonnegative Hull of a Specified Set of Points*

Consider the case, $\mathbf{K} =$ nonnegative hull of $\{A_{.j} : j = 1 \text{ to } t\} \subset \mathbf{R}^n$, t finite. Let A be the $n \times t$ matrix consisting of column vectors $A_{.j}$, $j = 1$ to t. Then $\mathbf{K} = \text{Pos}(A)$, a convex polyhedral cone, expressed as the nonnegative hull of a given finite set of points from $\mathbf{R}^n$. In this special case, the separating hyperplane theorem becomes exactly Farkas' theorem (Theorem 3). See Section 4.6.7 of [2.26] or [1.28]. Since $b \notin \text{Pos}(A)$, system (I) of Farkas' theorem, Theorem 3, has no feasible solution, and hence system (II) has a solution π. Then the hyperplane $\{x : \pi x = 0\}$ separates b from $\text{Pos}(A)$. The solution π for system (II) can be computed efficiently using Phase I of the simplex method, as discussed in Note 1 of Appendix 1. Given any point $b \in \mathbf{R}^n$, this provides an efficient method to check whether $b \in \text{Pos}(A)$ (which happens when system (I) of Farkas' theorem, Theorem 3, with this data, has a feasible solution); and if not, to compute a hyperplane separating b from $\text{Pos}(A)$, as long as the number of points in the set $\{A_{.j} : j = 1 \text{ to } t\}$, t is not too large. If t is very large, the method discussed here for computing a separating hyperplane, may not be practically useful, this is discussed below using some actual examples.

K *Represented as the Convex Hull of a Specified Set of Points*

Consider the case where $\mathbf{K}$ is specified as the convex hull of a given set of points $\{A_{.j} : j = 1 \text{ to } t\} \subset \mathbf{R}^n$. So, in this case, $b \notin \mathbf{K}$, iff the system

$$\sum_{j=1}^{t} A_{.j} x_j = b$$

$$\sum_{j=1}^{t} x_j = 1$$

$$x_j \geqq 0, \ j = 1 \text{ to } t$$

has no feasible solution $x = (x_j)$. This system is exactly in the same form as system (I) of Farkas' theorem, Theorem 3, and a separating hyperplane in this case can be computed using this theorem, as discussed above, as long as t is not too large.

K *Represented by a System of "$\leqq$" Inequalities Involving Convex Functions*

Now consider the case where $\mathbf{K}$ is represented as the set of feasible solutions of a system of inequalities

$$f_i(x) \leqq 0, \ i = 1 \text{ to } m$$

where each $f_i(x)$ is a differentiable convex function defined on $\mathbf{R}^n$. See the following section, Appendix 3, for definitions of convex functions and their properties. In this case, $b \notin \mathbf{K}$, iff there exists an i satisfying $f_i(b) > 0$. If $b \notin \mathbf{K}$, find such an i, say r. Then the hyperplane $H = \{x : f_r(b) + \nabla f_r(b)(x - b) = 0\}$ separates b from $\mathbf{K}$, by Theorem 15 of Appendix 3 (see Exercise 11 in Appendix 3).

$\mathbf{K}$ *Represented by a System of "$\leqq$" Inequalities Involving General Functions*

Now consider the case in which the convex set $\mathbf{K}$ is represented by a system of constraints

$$g_i(x) \leqq 0,\ i = 1 \text{ to } m$$

where the functions $g_i(x)$ are not all convex functions. It is possible for the set of feasible solutions of such a system to be convex set. As an example let $n = 2$, $x = (x_1, x_2)^T$, and consider the system

$$\begin{aligned} -x_1 - x_2 + 2 &\leqq 0 \\ x_1 - 1 &\leqq 0 \\ x_2 - 1 &\leqq 0 \\ -x_1^4 - x_2^4 + 2 &\leqq 0. \end{aligned}$$

This system has the unique solution $x = (1, 1)^T$, and yet, not all the functions in the system are convex functions. As another example, let M be a P-matrix of order n which is not a PSD matrix, and $q \in \mathbf{R}^n$. Consider the system in variables $z = (z_1, \ldots, z_n)^T$

$$\begin{aligned} -z &\leqq 0 \\ -q - Mz &\leqq 0 \\ z^T(q + Mz) &\leqq 0. \end{aligned}$$

This system has the unique solution $\overline{z}$ ($\overline{z}$ is the point which leads to the unique solution of the LCP (q, M)), so the set of feasible solutions of this system is convex, being a singleton set, and yet the constraint function $z^T(q + Mz)$ is not convex, since M is not PSD.

In general, when the functions $g_i(x)$, $i = 1$ to m are not all convex, even though the set $\mathbf{K} = \{x : g_i(x) \leqq 0,\ i = 1 \text{ to } m\}$ may be convex, and $b \notin \mathbf{K}$, there is no efficient method known for computing a hyperplane separating b from $\mathbf{K}$. See Exercise 40.

Now we consider some cases in which $\mathbf{K}$ is the convex hull of a set of points specified by some properties.

$\mathbf{K}$ *Is the Convex Hull of the Tours of a Traveling Salesman Problem*

Consider the famous traveling salesman problem in cities $1, 2, \ldots, n$. See [1.28]. In this problem, a salesman has to start in some city, say city 1, visit each of the other cities

exactly once in some order, and in the end return to the starting city, city 1. If he travels to cities in the order i to $i+1$, $i = 1$ to $n-1$ and then from city n to city 1, this route can be represented by the order "$1, 2, \ldots, n; 1$". Such an order is known as a **tour**. So, a tour is a circuit spanning all the cities, that leaves each city exactly once. From the starting city, city 1, he can go to any of the other $(n-1)$ cities. So there are $(n-1)$ different ways in which he can pick the city that he travels from the starting city, city 1. From that city he can travel to any of the remaining $(n-2)$ cities, etc. Thus the total number of possible tours in an n city traveling salesman problem is $(n-1)(n-2)\ldots 1 = (n-1)!$ Given a tour, define a $0-1$ matrix $x = (x_{ij})$ by

$$x_{ij} = \begin{cases} 1 & \text{if the salesman goes from city } i \text{ to city } j \text{ in the tour} \\ 0 & \text{otherwise.} \end{cases}$$

Such a matrix $x = (x_{ij})$ is called the **tour assignment** corresponding to the tour. An **assignment** (of order n) is any $0-1$ square matrix $x = (x_{ij})$ of order n satisfying

$$\sum_{j=1}^{n} x_{ij} = 1, \ i = 1 \text{ to } n$$

$$\sum_{i=1}^{n} x_{ij} = 1, \ j = 1 \text{ to } n$$

$$x_{ij} = 0 \text{ or } 1 \text{ for all } i, j.$$

Every tour assignment is an assignment, however not all assignments may be tour assignments. For example, if $n = 5$

$$x^1 = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix}$$

is a tour assignment representing the tour 1, 4, 2, 5, 3; 1 covering all the cities 1 to 5. But the assignment

$$x^2 = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

is not a tour assignment, since it consists of two subtours 1, 2, 3; 1 and 4, 5; 4 each spanning only a proper subset of the original set of cities.

Let $\mathbf{K}_T$ be the convex hull of all the $(n-1)!$ tour assignments of order n. $\mathbf{K}_T$ is well defined, it is the convex hull of a finite set of points in $\mathbf{R}^{n \times n}$. However, if n is large (even $n \geqq 10$), the number of tour assignments, $(n-1)!$ is very large. $\mathbf{K}_T$ is of course a convex polytope. It can be represented as the set of feasible solutions of a system of linear constraints, but that system is known to contain a very large number

of constraints. Deriving a linear constraint representation of $\mathbf{K}_T$ remains an unsolved problem. In this case, if $b = (b_{ij})$ is a given square matrix of order n satisfying the conditions

$$\begin{aligned} b_{ii} &= 0, \ i = 1 \text{ to } n \\ \sum_{j=1}^{n} b_{ij} &= 1, \ i = 1 \text{ to } n \\ \sum_{i=1}^{n} b_{ij} &= 1, \ i = 1 \text{ to } n \\ 0 \leqq b_{ij} &\leqq 1, \text{ for all } i, j = 1 \text{ to } n \end{aligned}$$

even to check whether $b \in \mathbf{K}_T$ is a hard problem for which no efficient algorithm is known. Ideally, given such a b, we would like an algorithm which

either determines that $b \in \mathbf{K}_T$

or determines that $b \notin \mathbf{K}_T$ and produces in this case a hyperplane separating b from $\mathbf{K}_T$

and for which the computational effort in the worst case is bounded above by a polynomial in n. No such algorithm is known, and the problem of constructing such an algorithm, or even establishing whether such an algorithm exists, seems to be a very hard problem. If such an algorithm exists, using it we can construct efficient algorithms for solving the traveling salesman problem, which is the problem of finding a minimum cost tour assignment that minimizes $\sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} x_{ij}$ for given cost matrix $c = (c_{ij})$.

K *Is the Convex Hull of Feasible Solutions of an Integer Linear System*

Let A, d be given integer matrices of orders $m \times n$ and $m \times 1$ respectively. Consider the following systems: $x = (x_j) \in \mathbf{R}^n$

$$\begin{aligned} Ax &= d \\ x &\geqq 0 \\ x &\text{ an integer vector} \end{aligned}$$

or the system

$$\begin{aligned} Ax &= d \\ x &\geqq 0 \\ 0 \leqq x_j &\leqq 1, \ j = 1 \text{ to } n \\ x_j &\text{ integer for all } j. \end{aligned}$$

Let $\mathbf{K}_I$ denote the convex hull of all feasible solutions of such a system. Again, $\mathbf{K}_I$ is a well defined set, it is the convex hull of integer feasible solutions to a specified system of linear constraints. Given a point $b \in \mathbf{R}^n$, ideally we would like an algorithm which

either determines that $b \in \mathbf{K}_I$

or determines that $b \notin \mathbf{K}_I$ and produces in this case a hyperplane separating b from $\mathbf{K}_I$

and for which the computational effort in the worst case is bounded above by a polynomial in the size of (A, b). No such algorithm is known.

$\mathbf{K}$ *Is the Convex Hull of Extreme Points of an Unbounded Convex Polyhedron*

Let A, d be given integer matrices of orders $m \times n$ and $m \times 1$ respectively, with rank $(A) = m$. Let $\mathbf{\Gamma}$ be the set of feasible solutions of the system

$$\begin{aligned} Ax &= d \\ x &\geqq 0. \end{aligned}$$

Suppose it is known that $\mathbf{\Gamma}$ is an unbounded convex polyhedron. $\mathbf{\Gamma}$ has a finite set of extreme points, each of these is a BFS of the above system. Let $\mathbf{K}$ be the convex hull of all these extreme points $\mathbf{\Gamma}$. Here again $\mathbf{K}$ is a well defined convex polytope, but it is the convex hull of extreme points of $\mathbf{\Gamma}$, and the number of these extreme points may be very large. See Section 3.7 of [2.26]. In general, given a point $b \in \mathbf{\Gamma}$, the problem of determining whether $b \in \mathbf{K}$, and the problem of determining a separating hyperplane separating b and $\mathbf{K}$ when $b \notin \mathbf{K}$, are very hard problems for which no efficient algorithms are known (the special case when $n = m + 2$ or $m + 1$ are easy, because in this case the dimension of $\mathbf{\Gamma}$ is at most two).

Summary

This discussion clearly illustrates the fact that even though we have proved the existence of separating planes, at the moment algorithms for computing one of them efficiently are only known when $\mathbf{K}$ can be represented in very special forms.

3. CONVEX, CONCAVE FUNCTIONS, THEIR PROPERTIES

Let $\mathbf{\Gamma}$ be a convex subset of $\mathbf{R}^n$ and let $f(x)$ be a real valued function defined on $\mathbf{\Gamma}$. $f(x)$ is said to be a **convex function** iff for any $x^1, x^2 \in \mathbf{\Gamma}$, and $0 \leqq \alpha \leqq 1$, we have

$$f(\alpha x^1 + (1 - \alpha)x^2) \leqq \alpha f(x^1) + (1 - \alpha) f(x^2). \tag{41}$$

This inequality is called **Jensen's inequality** after the Danish mathematician who first discussed it. The important property of convex functions is that when you join two points on the surface of the function by a chord, the function itself lies underneath the chord on the interval joining these points, see Figure 6.

Similarly, if $g(x)$ is a real valued function defined on the convex set $\mathbf{\Gamma} \subset \mathbf{R}^n$, it is said to be a **concave function** iff for any $x^1, x^2 \in \mathbf{\Gamma}$ and $0 \leqq \alpha \leqq 1$, we have

$$g(\alpha x^1 + (1-\alpha)x^2) \geqq \alpha g(x^1) + (1-\alpha)g(x^2). \tag{42}$$

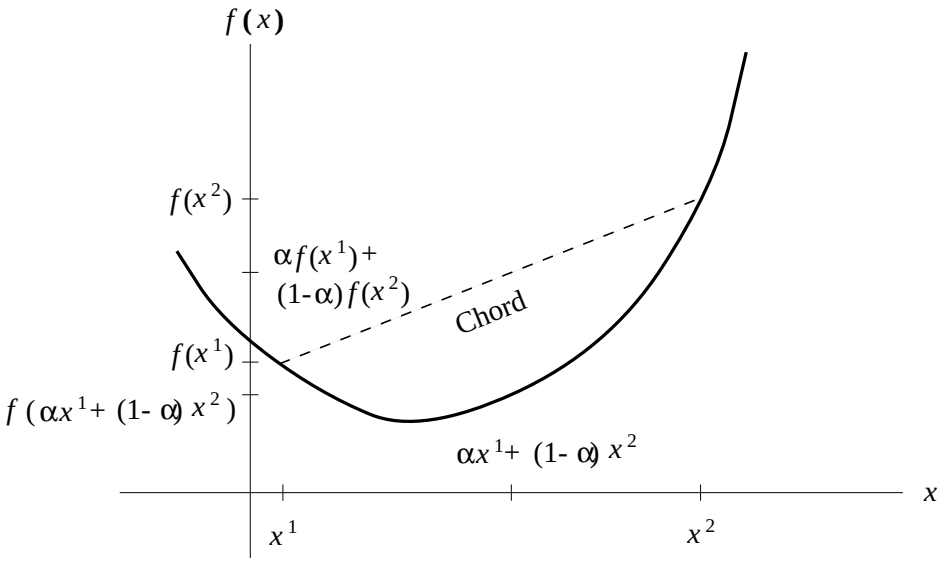


Figure 6 A convex function defined on the real line.

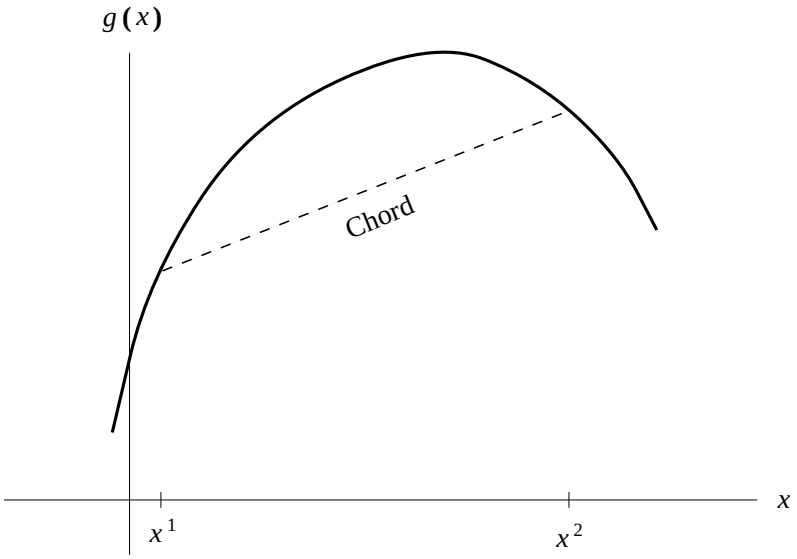


Figure 7 A concave function defined on the real line.

Clearly, a function is concave iff its negative is convex. Also, a concave function lies above the chord on any interval, see Figure 7. Convex and concave functions figure prominently in optimization. In mathematical programming literature, the problem of

either minimizing a convex function, or maximizing a concave function, on a convex set, are known as convex programming problems. For a convex programming problem, a local optimum solution is a global optimum solution (see Theorem 12 below) and hence any techniques for finding a local optimum will lead to a global optimum on these problems.

The function $f(x)$ defined above is said to be **strictly convex**, if (41) holds as a strict inequality for $0 < \alpha < 1$ and for all $x^1, x^2 \in \mathbf{\Gamma}$. Likewise $g(x)$ is said to be a **strictly concave function** if (42) holds as a strict inequality for $0 < \alpha < 1$ and for all $x^1, x^2 \in \mathbf{\Gamma}$.

The following results can be verified to be true.

1. A nonnegative combination of convex functions is convex. Likewise a nonnegative combination of concave functions is concave.
2. If $f(x)$ is a convex function defined on the convex set $\mathbf{\Gamma} \subset \mathbf{R}^n$, $\{x : f(x) \leqq \alpha\}$ is a convex set for all real numbers α. Likewise, if $g(x)$ is a concave function defined on the convex set $\mathbf{\Gamma} \subset \mathbf{R}^n$, $\{x : g(x) \geqq \alpha\}$ is a convex set for all real numbers α.
3. If $f_1(x), \ldots, f_r(x)$ are all convex functions defined on the convex set $\mathbf{\Gamma} \subset \mathbf{R}^n$, the pointwise supremum function $f(x) = \text{maximum } \{f_1(x), \ldots, f_r(x)\}$ is convex.
4. If $g_1(x), \ldots, g_r(x)$ are all concave functions defined on the convex set $\mathbf{\Gamma} \subset \mathbf{R}^n$, the pointwise infimum function $g(x) = \text{minimum } \{g_1(x), \ldots, g_r(x)\}$ is concave.
5. A convex or concave function defined on an open convex subset of $\mathbf{R}^n$ is continuous (see [A10] for a proof of this).
6. Let $f(x)$ be a real valued function defined on a convex subset $\mathbf{\Gamma} \subset \mathbf{R}^n$. In $\mathbf{R}^{n+1}$, plot the objective value of $f(x)$ along the x_{n+1}-axis. The subset of $\mathbf{R}^{n+1}$, $\mathbf{F} = \{X = (x_1, \ldots, x_n, x_{n+1}) : x = (x_1, \ldots, x_n) \in \mathbf{\Gamma}, x_{n+1} \geqq f(x)\}$ is known as the **epigraph** of the function $f(x)$. It is the set of all points in $\mathbf{R}^{n+1}$ lying above (along the x_{n+1}-axis) the surface of $f(x)$. See Figure 8 for an illustration of the epigraph of a convex function defined on an interval of the real line $\mathbf{R}^1$. It can be shown that $f(x)$ is convex iff its epigraph is a convex set, from the definitions of convexity of a function and of a set. See Figures 8, 9.
7. Let $g(x)$ be a real valued function defined on a convex subset $\mathbf{\Gamma} \subset \mathbf{R}^n$. In $\mathbf{R}^{n+1}$, plot the objective value of $f(x)$ along the x_{n+1}-axis. The subset of $\mathbf{R}^{n+1}$, $\mathbf{G} = \{X = (x_1, \ldots, x_n, x_{n+1}) : x = (x_1, \ldots, x_n) \in \mathbf{\Gamma}, x_{n+1} \leqq g(x)\}$ is known as the **hypograph** of the function $g(x)$. It is the set of all points in $\mathbf{R}^{n+1}$ lying below (along the x_{n+1}-axis) the surface of $g(x)$. See Figure 10. It can be shown from the definitions, that $g(x)$ is concave, iff its hypograph is a convex set.

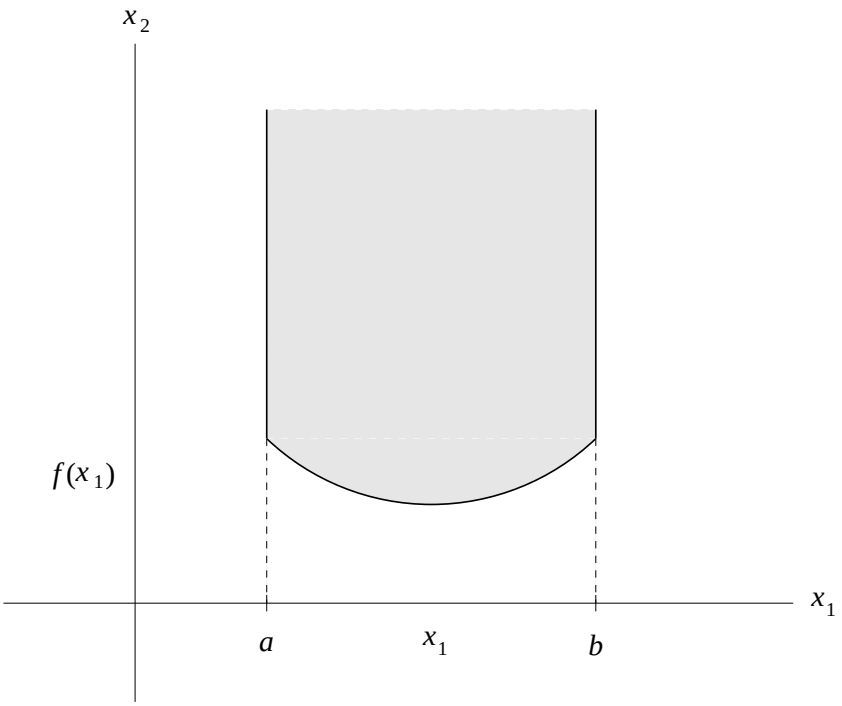


Figure 8 The epigraph of a convex function defined on the interval $a \leqq x_1 \leqq b$ is a convex subset of $\mathbf{R}^2$.

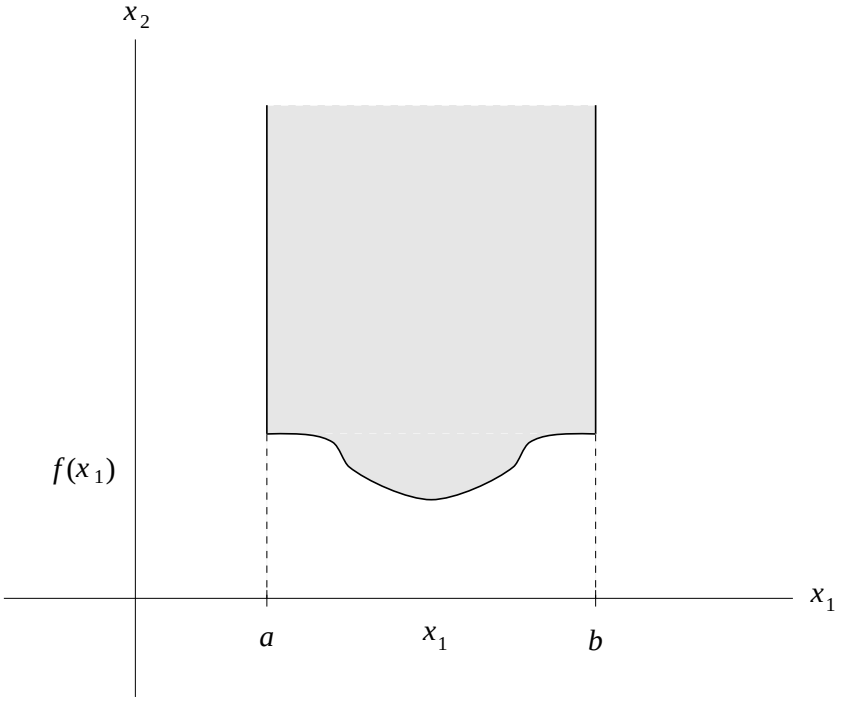


Figure 9 The epigraph of a nonconvex function $f(x_1)$ defined on the interval $a \leqq x_1 \leqq b$, is not a convex subset of $\mathbf{R}^2$.

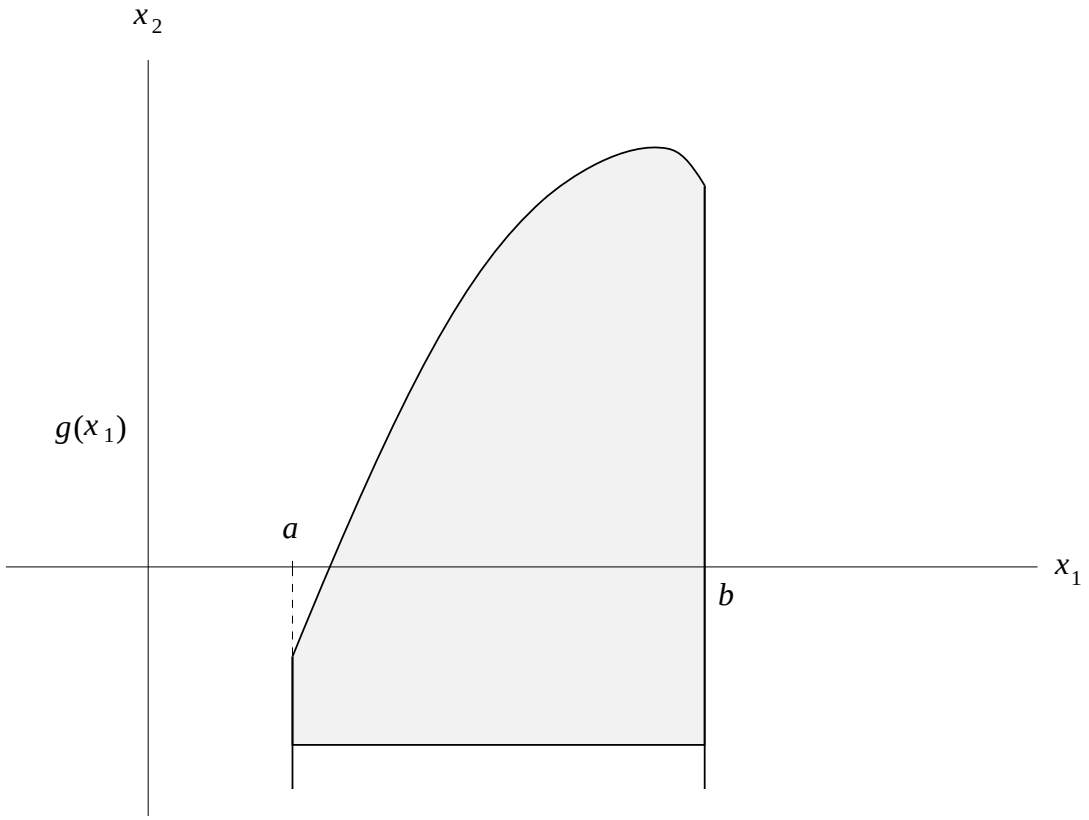


Figure 10 The hypograph of a concave function defined on the interval $a \leqq x_1 \leqq b$ is a convex subset of $\mathbf{R}^2$.

Theorem 12. *For the problem of minimizing a convex function $f(x)$ on a convex set $\mathbf{\Gamma} \subset \mathbf{R}^n$, every local minimum is a global minimum.*

Proof. Let x^1 be a local minimum for this problem. Suppose there exists an $x^2 \in \mathbf{\Gamma}$ such that $f(x^2) < f(x^1)$. Then, by convexity, for $0 < \alpha < 1$,

$$f(x^1 + \alpha(x^2 - x^1)) = f(\alpha x^2 + (1-\alpha)x^1) \leqq (1-\alpha)f(x^1) + \alpha f(x^2) < f(x^1). \quad (43)$$

So when α is positive but sufficiently small, the point $x^1 + \alpha(x^2 - x^1)$ contained in the neighborhood of x^1 satisfies (43), contradicting the local minimum property of x^1. So we cannot have an $x^2 \in \mathbf{\Gamma}$ satisfying $f(x^2) < f(x^1)$, that is, x^1 is in fact the global minimum for $f(x)$ in $\mathbf{\Gamma}$.

□

Theorem 13. *Let f be a real valued convex function defined on the convex set $\mathbf{\Gamma} \subset \mathbf{R}^n$. The set of optimum solutions for the problem of minimizing $f(x)$ over $x \in \mathbf{\Gamma}$ is a convex set.*

Proof. Let $\mathbf{L}$ denote the set of optimum solutions for the problem: minimize $f(x)$ over $x \in \mathbf{\Gamma}$. Let $x^1, x^2 \in \mathbf{L}$. So $f(x^1) = f(x^2) = \lambda =$ minimum value of $f(x)$ over $x \in \mathbf{\Gamma}$. Let $0 \leqq \alpha \leqq 1$. By convexity of $f(x)$, $f(\alpha x^1 + (1-\alpha)x^2) \leqq \alpha f(x^1) + (1-\alpha)f(x^2) = \lambda$

and $\alpha x^1 + (1-\alpha)x^2 \in \mathbf{\Gamma}$, since $\mathbf{\Gamma}$ is convex. Since λ is the minimum value of $f(x)$ over $x \in \mathbf{\Gamma}$, the above inequality must hold as an equation, that is, $f(\alpha x^1 + (1-\alpha)x^2) = \lambda$, which implies that $\alpha x^1 + (1-\alpha)x^2 \in \mathbf{L}$ also. So $\mathbf{L}$ is a convex set.

□

Theorem 14. *For the problem of maximizing a concave function $g(x)$ on a convex set $\mathbf{\Gamma} \subset \mathbf{R}^n$, every local maximum is a global maximum.*

Proof. Similar to Theorem 12.

□

A real valued function $\theta(x)$ defined on an open set $\mathbf{\Gamma} \subset \mathbf{R}^n$ is said to be differentiable at a point $\overline{x} \in \mathbf{\Gamma}$ if the partial derivative vector $\nabla\theta(\overline{x})$ exists, and for each $y \in \mathbf{R}^n$, limit $((\theta(\overline{x}+\lambda y) - \theta(\overline{x}) - \lambda(\nabla\theta(\overline{x}))y)/\lambda)$ as λ tends to zero is zero. $\theta(x)$ is said to be twice differentiable at $\overline{x}$ if the Hessian matrix $H(\theta(\overline{x}))$ exists and for each $y \in \mathbf{R}^n$, limit $[(\theta(\overline{x}+\lambda y) - \theta(\overline{x}) - \lambda(\nabla\theta(\overline{x}))y - (\lambda^{\mathbf{2}}/2)y^T H(\theta(\overline{x}))y]/\lambda^{\mathbf{2}})$ as λ tends to zero is zero.

The real valued function $\theta(x)$ defined on an open set $\mathbf{\Gamma} \subset \mathbf{R}^n$ is said to be continuously differentiable at a point $\overline{x} \in \mathbf{\Gamma}$ if it is differentiable at $\overline{x}$ and the partial derivatives $\frac{\partial\theta(x)}{\partial x_j}$ are all continuous at $\overline{x}$. The function $\theta(x)$ is said to continuously differentiable at a point $\overline{x} \in \mathbf{\Gamma}$ if it is twice differentiable at $\overline{x}$ and the second order partial derivatives $\frac{\partial^2\theta(x)}{\partial x_i \partial x_j}$ are all continuous at $\overline{x}$. The function is said to be differentiable, continuously differentiable, etc., over the set $\mathbf{\Gamma}$, if it satisfies the corresponding property for each point in $\mathbf{\Gamma}$.

Theorem 15 (Gradient Support Inequality): *Let $f(x)$ be a real valued convex function defined on an open convex set $\mathbf{\Gamma} \subset \mathbf{R}^n$. If $f(x)$ is differentiable at $\overline{x} \in \mathbf{\Gamma}$,*

$$f(x) - f(\overline{x}) \geqq (\nabla f(\overline{x}))(x - \overline{x}) \text{ for all } x \in \mathbf{\Gamma}. \tag{44}$$

Conversely, if $f(x)$ is a real valued differentiable function defined on $\mathbf{\Gamma}$ and (44) holds for all $x, \overline{x} \in \mathbf{\Gamma}$, $f(x)$ is convex.

Proof. Suppose $f(x)$ is convex. Let $x \in \mathbf{\Gamma}$. By convexity of $\mathbf{\Gamma}$, $\alpha x + (1-\alpha)\overline{x} = \overline{x} + \alpha(x - \overline{x}) \in \mathbf{\Gamma}$ for all $0 \leqq \alpha \leqq 1$. Since $f(x)$ is convex we have $f(\overline{x} + \alpha(x - \overline{x})) \leqq \alpha f(x) + (1-\alpha)f(\overline{x})$. So for $0 < \alpha \leqq 1$, we have

$$f(x) - f(\overline{x}) \geqq (f(\overline{x} + \alpha(x - \overline{x})) - f(\overline{x}))/\alpha. \tag{45}$$

By definition of differentiability, the right hand side of (45) tends to $\nabla f(\overline{x})(x - \overline{x})$ as α tends to zero through positive values. Since (45) holds for all $0 < \alpha \leqq 1$, this implies (44) as α tends to zero through positive values in (45).

Conversely, suppose $f(x)$ is a real valued differentiable function defined on $\mathbf{\Gamma}$ and suppose (44) holds for all, $x, \overline{x} \in \mathbf{\Gamma}$. Given $x^1, x^2 \in \mathbf{\Gamma}$, from (44) we have, for $0 < \alpha < 1$,

$$f(x^1) - f((1-\alpha)x^1 + \alpha x^2) \geqq \alpha(\nabla f(1-\alpha)x^1 + \alpha x^2)(x^1 - x^2)$$
$$f(x^2) - f((1-\alpha)x^1 + \alpha x^2) \geqq -(1-\alpha)(\nabla f((1-\alpha)x^1 + \alpha x^2))(x^1 - x^2).$$

Multiply the first inequality by $(1-\alpha)$ and the second by α and add. This leads to

$$(1-\alpha)f(x^1)+\alpha f(x^2)-f((1-\alpha)x^1+\alpha x^2)\geqq 0. \tag{46}$$

Since (46) holds for all $x^1, x^2 \in \mathbf{\Gamma}$ and $0<\alpha<1$, $f(x)$ is convex.

□

Theorem 16. *Let $g(x)$ be a concave function defined on an open convex set $\mathbf{\Gamma} \subset \mathbf{R}^n$. If $g(x)$ is differentiable at $\overline{x} \in \mathbf{\Gamma}$,*

$$g(x) \leqq g(\overline{x}) + (\nabla g(\overline{x}))(x-\overline{x}), \text{ for all } x \in \mathbf{\Gamma}. \tag{47}$$

Conversely, if $g(x)$ is a differentiable function defined on $\mathbf{\Gamma}$ and (47) holds for all $x, \overline{x} \in \mathbf{\Gamma}$, $g(x)$ is concave.

Proof. Similar to the proof of Theorem 15.

□

Figures 11, 12 provide illustrations of gradient support inequalities for convex and concave functions.

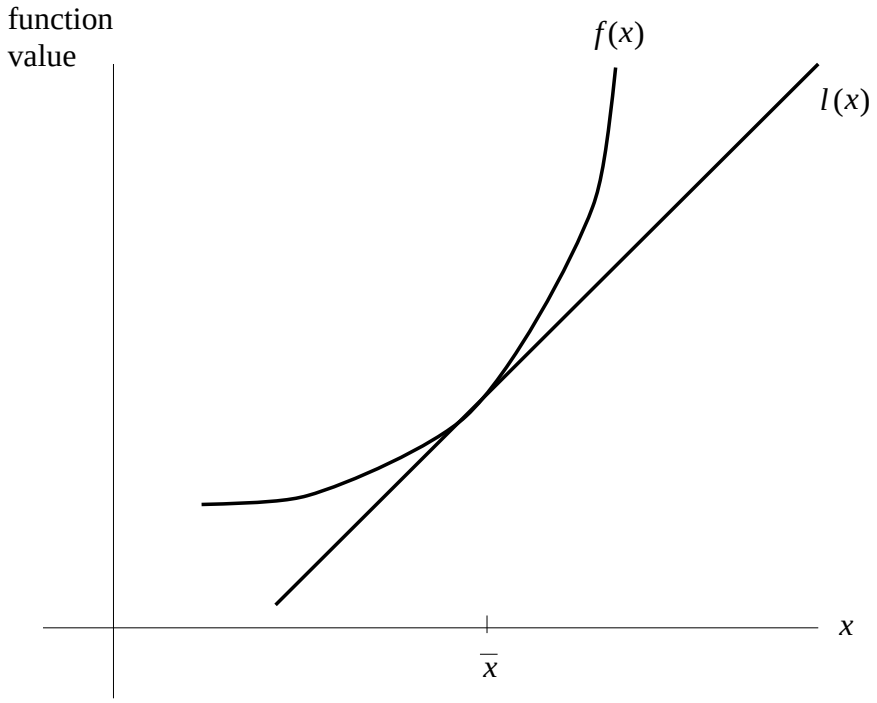


Figure 11 $f(x)$ is a differentiable convex function. $l(x) = f(\overline{x}) + (\nabla f(\overline{x}))(x-\overline{x})$, an affine function (since $\overline{x}$ is a given point), is the first order Taylor series approximation for $f(x)$ around $\overline{x}$. It underestimates $f(x)$ at each point.

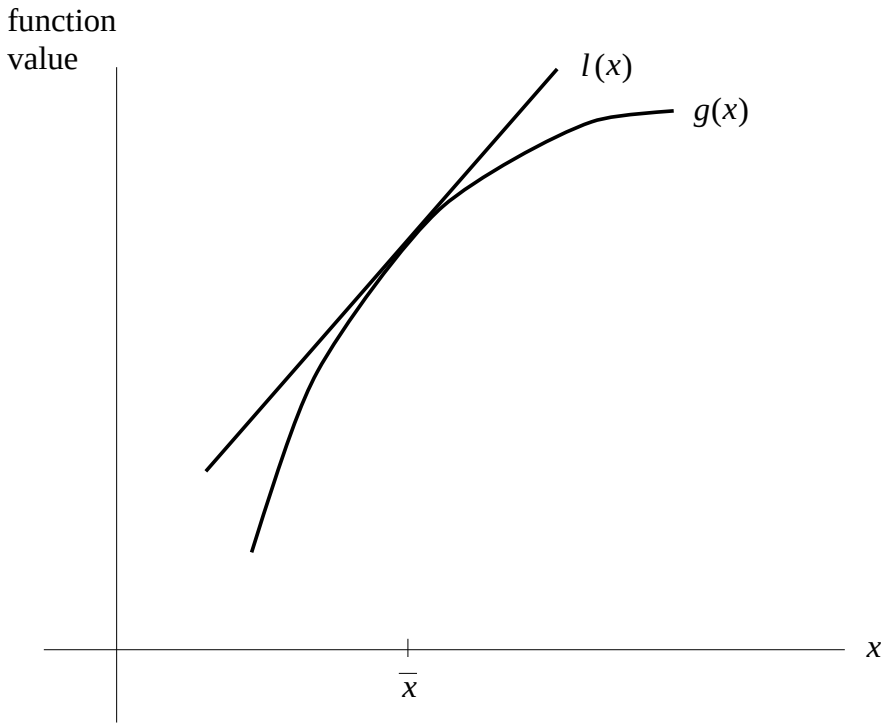


Figure 12 $g(x)$ is a differentiable concave function. $l(x) = g(\overline{x}) + (\nabla g(\overline{x})(x - \overline{x}))$ is the first order Taylor series approximation for $g(x)$ around $\overline{x}$. It overestimates $g(x)$ at each point.

Theorem 17. *Let $f(x)$ be a real valued convex function defined on an open convex subset $\mathbf{\Gamma} \subset \mathbf{R}^n$. If $f(x)$ is twice differentiable at $\overline{x} \in \mathbf{\Gamma}$, $H(f(\overline{x}))$ is* PSD. *Conversely, if $f(x)$ is a twice differentiable real valued function defined on $\mathbf{\Gamma}$ and $H(f(\overline{x}))$ is* PSD *for all $\overline{x} \in \mathbf{\Gamma}$, $f(x)$ is convex.*

Proof. Let $\overline{x} \in \mathbf{\Gamma}$ and $y \in \mathbf{R}^n$. Suppose $f(x)$ is convex. For $\alpha > 0$ and sufficiently small, by Theorem 15 we have

$$(f(\overline{x} + \alpha y) - f(\overline{x}) - \alpha(\nabla f(\overline{x}))y)/\alpha \geqq 0. \tag{48}$$

Taking the limit as α tends to zero through positive values, from (48) we have $y^T H(f(\overline{x}))y \geqq 0$, and since this holds for all $y \in \mathbf{R}^n$, $H(f(\overline{x}))$ is PSD.

Suppose $f(\overline{x})$ is twice differentiable on $\mathbf{\Gamma}$ and $H(f(\overline{x}))$ is PSD for all $\overline{x} \in \mathbf{\Gamma}$. By Taylor's theorem of calculus we have, for $x^1, x^2 \in \mathbf{\Gamma}$, $f(x^2) - f(x^1) - (\nabla f(x^1))(x^2 - x^1) = (x^2 - x^1)^T H(f(x^1 + \alpha(x^2 - x^1)))(x^2 - x^1)/2$ for some $0 < \alpha < 1$. But the latter expression is $\geqq 0$ since $H(f(\overline{x}))$ is PSD for all $\overline{x} \in \mathbf{\Gamma}$. So $f(x^2) - f(x^1) - (\nabla f(x^1))(x^2 - x^1) \geqq 0$ for all $x^1, x^2 \in \mathbf{\Gamma}$. By Theorem 15, this implies that $f(x)$ is convex.

□

Given a general twice continuously differentiable real valued function $f(x)$ defined on $\mathbf{R}^n$, it may be hard to check whether it is convex. For some $\overline{x} \in \mathbf{R}^n$, if $H(f(\overline{x}))$ is PD, we know that in a small convex neighborhood of $\overline{x}$, $H(f(x))$ is PSD, and hence $f(x)$ is locally convex in this neighborhood.

Theorem 18. *Let $g(x)$ be a real valued concave function defined on an open convex subset $\mathbf{\Gamma} \subset \mathbf{R}^n$. If $g(x)$ is twice differentiable at $\overline{x} \in \mathbf{\Gamma}$, $H(g(\overline{x}))$ is* NSD. *Conversely, if $g(x)$ is a twice differentiable real valued function defined on $\mathbf{\Gamma}$ and $H(g(\overline{x}))$ is* NSD *for all $\overline{x} \in \mathbf{\Gamma}$, $g(x)$ is concave.*

Proof. Similar to the proof of Theorem 17.

□

Exercises

10. Let $X^r = (x_1^r, \ldots, x_n^r, x_{n+1}^r)^T$ $r = 1$ to m be given points in $\mathbf{R}^{n+1}$. Let $x^r = (x_1^r, \ldots, x_n^r)^T$ $r = 1$ to m. It is required to check whether there exists a convex function $\theta(x)$ defined on $\mathbf{R}^n$ (with the objective value plotted along the x_{n+1}-axis in $\mathbf{R}^{n+1}$) satisfying the property $\theta(x^r) = x_{n+1}^r$ for $r = 1$ to m. Formulate this as a linear programming problem.

11. Let $f(x)$ be a real valued continuously differentiable convex function defined on $\mathbf{R}^n$. Let α be a real number and $\mathbf{K} = \{x : f(x) \leqq \alpha\}$. Given a point $x^0 \notin \mathbf{K}$, develop an efficient method for finding a separating hyperplane separating x^0 from $\mathbf{K}$. Generalize this to the case where $f(x) = (f_1(x), \ldots, f_m(x))$, each $f_i(x)$ being a real valued continuously differentiable function defined on $\mathbf{R}^n$, and $\alpha \in \mathbf{R}^m$.

12. Let $\theta(x)$ be a differentiable convex function defined over a convex set $\mathbf{K} \subset \mathbf{R}^n$. Let $\overline{x}$ be a given point in $\mathbf{K}$. If $\overline{x}$ satisfies the property that it minimizes the linear function $(\nabla\theta(\overline{x}))x$ over $x \in \mathbf{K}$, prove that $\overline{x}$ also minimizes $\theta(x)$ over $x \in \mathbf{K}$.

Convexity, Concavity of a Vector Function

Let $f(x)$ be the vector $(f_i(x))$ where each $f_i(x)$ is a real valued function defined on the convex set $\mathbf{\Gamma} \subset \mathbf{R}^n$. $f(x)$ is said to be convex or concave on $\mathbf{\Gamma}$, iff each $f_i(x)$ has the same property.

Subgradients, and Subdifferential Sets

Let $f(x)$ be a real valued convex function defined on $\mathbf{R}^n$. As defined in Section 2.7.1, the vector $d = (d_1, \ldots, d_n)^T$ is said to be a **subgradient of** $f(x)$ at a point $x^0 \in \mathbf{R}^n$, if

$$f(x) \geqq f(x^0) + d^T(x - x^0), \quad \text{for all } x \in \mathbf{R}^n.$$

The set of all such vectors d satisfying this condition is known as the **subdifferential set** for $f(x)$ at x^0, and denoted by the symbol $\partial f(x^0)$.

By Theorem 15, if $f(x)$ is differentiable at x^0, the gradient vector $(\nabla f(x^0))^T \in \partial f(x^0)$, and in fact it can be shown that in this case $\partial f(x^0) = \{\nabla f(x^0)^T\}$. Also, as mentioned in Section 2.7.1, if $f(x) = \max\{f_1(x), \ldots, f_r(x)\}$ where each $f_i(x)$ is a differentiable convex function defined on $\mathbf{R}^n$, then for any $\overline{x} \in \mathbf{R}^n$,

$$\partial f(\overline{x}) = \text{ convex hull of } \{\nabla f_i(\overline{x}) : i \text{ such that } f(\overline{x}) = f_i(\overline{x})\}.$$

See Section 2.7.1 for figures illustrating the subgradient property. The definition implies that if $f(x)$ is convex and $d \in \partial f(\overline{x})$, then the affine function $l(x) = f(\overline{x}) + d^T(x - \overline{x})$ is equal to $f(x)$ at $x = \overline{x}$, and is an underestimate for $f(x)$ at all points x.

So the error $f(x) - l(x) = f(x) - (f(\overline{x}) + d^T(x - \overline{x})) \geqq 0$ for all x and $d \in \partial f(\overline{x})$. See Section 2.7.1 for figures illustrating this property. The affine function $l(x)$ defined above is known as a **linearization of** $f(x)$ at $\overline{x}$.

If $h(x)$ is a concave function defined on $\mathbf{R}^n$, the vector d is said to be a subgradient of $h(x)$ at $\overline{x}$ if

$$h(x) \leqq h(\overline{x}) + d^T(x - \overline{x}) \text{ for all } x \in \mathbf{R}^n$$

and the set of all subgradients to $h(x)$ at $\overline{x}$ is denoted by $\partial h(\overline{x})$. With this definition, analogous results to those stated above, can be constructed for concave functions.

Let $g(x)$ be a real valued function defined on $\mathbf{R}^n$ which is neither convex nor concave. If $g(x)$ is differentiable at a point $\overline{x} \in \mathbf{R}^n$, the affine function $l(x) = g(\overline{x}) + \nabla g(\overline{x})(x - \overline{x})$ is known as the **linearization of** $g(x)$ at $\overline{x}$. However, since $g(x)$ is neither convex nor concave, it is possible for the error $g(x) - l(x)$ to take both positive and negative values over $\mathbf{R}^n$. See Figure 13.

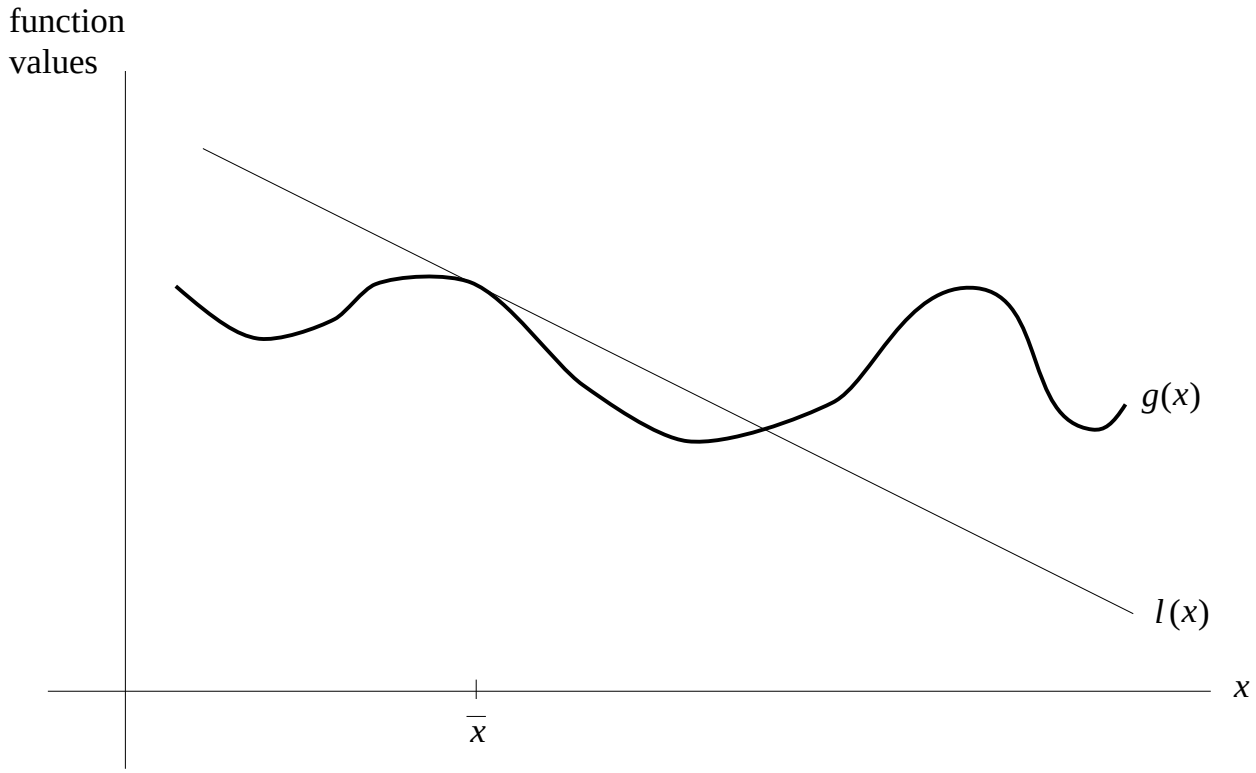


Figure 13 The linearization at $\overline{x}$, $l(x)$, of a differentiable function $g(x)$ which is neither convex nor concave may be $> g(x)$ at some points x, and $< g(x)$ at other points.

For this general function $g(x)$, if it is differentiable at $\overline{x}$, we define $\partial g(\overline{x}) = \{\nabla g(\overline{x})\}$. If $g(x)$ is not differentiable at $\overline{x}$, we let $\partial f(\overline{x})$ denote the convex hull of all limits of sequences of the form $\{\nabla g(x^r) : \{x^r\}$ is a sequence converging to $\overline{x}$, such that $g(x)$ is differentiable at each x^r in the sequence$\}$. In this case, vectors in the set $\partial g(\overline{x})$ are called generalized gradients or subgradients of $g(x)$ at $\overline{x}$. See F. H. Clarke [A1]. With this definition, it can be shown that if $g(x) = \max\{g_1(x), \ldots, g_m(x)\}$, where each $g_i(x)$ is a continuously differentiable function, then $\partial g(\overline{x}) =$ convex hull of $\{\nabla g_i(\overline{x}) : i$ such that $g(\overline{x}) = g_i(\overline{x})\}$. If $g(x)$ is convex, the set $\partial g(x)$ defined here equals the subdifferential set of $g(x)$ at x as defined earlier. Also, it can be shown under fairly general conditions on $g(x)$ (for example, if $g(x)$ is a locally Lipschitz function, that is, if there exists an $\alpha > 0$ such that $|g(x) - g(y)| \leqq \alpha\|x - y\|$ for all x, y) that the following mean value result holds: there exists an $\hat{x}$ on the line segment joining x and y and a $\hat{d} \in \partial g(\hat{x})$, satisfying

$$g(x) - g(y) = \hat{d}^T(x - y).$$

This definition of subgradients or generalized gradients for general functions is used in Section 10.7.9 in constructing an algorithm for constrained line minimization. Also see N. Z. Shor [A13] for a detailed treatment of various types of generalized gradients, and their applications in subgradient algorithms for nondifferentiable minimization.

4. OPTIMALITY CONDITIONS FOR SMOOTH OPTIMIZATION PROBLEMS

Here we briefly survey the known optimality conditions for NLPs in which the objective and constraint functions are continuously differentiable.

The Principles on Which Optimality Conditions are Based

Let $\mathbf{K}$ denote the set of feasible solutions for an optimization problem in which the objective function $\theta(x)$ is to be minimized. Let $\overline{x} \in \mathbf{K}$ be a feasible solution. A feasible direction at $\overline{x}$ for $\mathbf{K}$ is a direction y satisfying the property that beginning at $\overline{x}$, you can move a positive length along a straight line in the direction y, without leaving $\mathbf{K}$. Necessary optimality conditions for this optimization problem are derived, based on two very simple principles. These are the following:

1. If $\overline{x}$ is a local minimum for this optimization problem, then, as you move from $\overline{x}$ straight along any feasible direction at $\overline{x}$ for $\mathbf{K}$, in a small neighborhood of $\overline{x}$, the objective value cannot decrease.

2. Take a one dimensional, nonlinear, differentiable curve in the feasible region $\mathbf{K}$, passing through $\overline{x}$. If $\overline{x}$ is a local minimum for this optimization problem, then, as you move from $\overline{x}$ along this curve, in a small neighborhood of $\overline{x}$, the objective value cannot decrease (in effect this says that if $\overline{x}$ is a local minimum for $\theta(x)$ in $\mathbf{K}$, then $\overline{x}$ must be a local minimum for the one dimensional optimization problem of minimizing $\theta(x)$ on the curve).

Of course 1 is a special case of 2, since a straight line is a differentiable curve. These principles make it possible for us to derive necessary conditions for local minimality in higher dimensional feasible regions using well known necessary conditions for local minimality in one-dimensional optimization problems.

All the necessary optimality conditions are derived using the above principles. Even though the principles are the same, their application leads to optimality conditions which depend on the structure of the problem.

We will now derive optimality conditions for different types of nonlinear programming problems.

Unconstrained Minimization

First consider the unconstrained minimization problem

$$\begin{aligned} &\text{minimize } \theta(x) \\ &\text{over } x \in \mathbf{R}^n. \end{aligned} \tag{49}$$

Given $\overline{x} \in \mathbf{R}^n$, $y \in \mathbf{R}^n$, $y \neq 0$, by differentiability of $\theta(x)$, we know that limit of $(\theta(\overline{x} + \alpha y) - \theta(\overline{x}) - \alpha(\nabla\theta(\overline{x}))y)/\alpha$ as α tends to zero is zero. So, if $(\nabla\theta(\overline{x}))y < 0$ by choosing α positive and sufficiently small, we will have $\theta(\overline{x} + \alpha y) < \theta(\overline{x})$. Similarly, if $(\nabla\theta(\overline{x}))y > 0$, by choosing α negative with sufficiently small absolute value we will have again $\theta(\overline{x} + \alpha y) < \theta(\overline{x})$. So if $\overline{x}$ is a local minimum for (49), we must have $(\nabla\theta(\overline{x}))y = 0$ for all $y \in \mathbf{R}^n$, that is

$$\nabla\theta(\overline{x}) = 0 \tag{50}$$

(50) is the **first order necessary condition** for $\overline{x}$ to be a local minimum for (49).

If $\theta(x)$ is twice continuously differential at $\overline{x}$, we know that the limit of $(\theta(\overline{x}+\alpha y) - \theta(\overline{x}) - \alpha(\nabla\theta(\overline{x}))y - (\alpha^{\mathbf{2}}/2)y^T H(\theta(\overline{x}))y)/\alpha^{\mathbf{2}}$ as α tends to zero is zero, where $H(\theta(\overline{x}))$ is the Hessian matrix (the matrix of second order partial derivatives) of $\theta(x)$ at $\overline{x}$. So if $\overline{x}$ is such that (50) is satisfied, and y is such that $y^T H(\theta(\overline{x}))y < 0$ then for $\alpha \neq 0$ and sufficiently small, we will have $\theta(\overline{x} + \alpha y) < \theta(\overline{x})$. So, if $\overline{x}$ is a local minimum for (49) we must have $y^T H(\theta(\overline{x}))y \geqq 0$ for all $y \in \mathbf{R}^n$, when $\overline{x}$ satisfies (50), that is

$$H(\theta(\overline{x})) \text{ must be PSD.} \tag{51}$$

(50) and (51) together are the **second order necessary conditions** for $\overline{x}$ to be a local minimum to (49).

We now state a sufficient optimality condition for (49) in the form of a theorem.

Theorem 19. *Suppose $\theta(x)$ is twice continuously differentiable, and $\overline{x}$ is a point satisfying*

$$\nabla\theta(\overline{x}) = 0, \quad \textit{and } H(\theta(\overline{x})) \text{is PD} \tag{52}$$

then $\overline{x}$ is a local minimum for (49).

Proof. Since $H(\theta(\overline{x}))$ is PD, all its principal subdeterminants are > 0. Since $\theta(x)$ is twice continuously differentiable, all principal subdeterminants of the Hessian matrix $H(\theta(x))$ are continuous functions. These facts imply that there exists an $\varepsilon > 0$, such that if $\mathbf{\Gamma} = \{x : \|x - \overline{x}\| < \varepsilon\}$, all principal subdeterminants of $H(\theta(x))$ are > 0 for all $x \in \mathbf{\Gamma}$. Being a Hessian matrix $H(\theta(x))$ is also symmetric, by Theorem 1.9 of Section 1.3.1, these facts imply that $H(\theta(x))$ is PSD for all $x \in \mathbf{\Gamma}$. By Theorem 17 of Appendix 3, this implies that $\theta(x)$ is convex over $x \in \mathbf{\Gamma}$. So by Theorem 15 of Appendix 3 (the gradient support inequality)

$$\begin{aligned} \theta(x) - \theta(\overline{x}) &\geqq (\nabla\theta(\overline{x}))(x - \overline{x}) \text{ for all } x \in \mathbf{\Gamma} \\ &\geqq 0, \text{ since } \nabla\theta(\overline{x}) = 0 \text{ by (52).} \end{aligned}$$

This proves that $\overline{x}$ is a local minimum for $\theta(x)$.

□

Thus a **sufficient condition** for $\overline{x}$ to be a local minimum for (49) is (52).

Example 3

Consider the problem

$$\text{minimize } \theta(x) = 2x_1^2 + x_2^2 + x_3^2 + x_1x_2 + x_2x_3 + x_3x_1 - 9x_1 - 9x_2 - 8x_3$$
$$\text{over } x \in \mathbf{R}^3.$$

From the necessary optimality conditions, we know that every local minimum for this problem must satisfy

$$\begin{aligned}
\frac{\partial\theta(x)}{\partial x_1} &= 4x_1 + \ x_2 + \ x_3 - 9 = 0 \\
\frac{\partial\theta(x)}{\partial x_2} &= \ x_1 + 2x_2 + \ x_3 - 9 = 0 \\
\frac{\partial\theta(x)}{\partial x_3} &= \ x_1 + \ x_2 + 2x_3 - 8 = 0.
\end{aligned}$$

This system of equations has the unique solution $\overline{x} = (1, 3, 2)^T$. The Hessian matrix is

$$H(\theta(\overline{x})) = \begin{pmatrix} 4 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}.$$

This matrix is PD. So $\overline{x}$ satisfies the sufficient conditions for a local minimum. Clearly, here, $\theta(x)$ is convex and hence $\overline{x}$ is a global minimum for $\theta(x)$.

Example 4

Consider the problem

$$\text{minimize } \theta(x) = 2x_1^2 + x_3^2 + 2x_1x_2 + 2x_1x_3 + 4x_2x_3 + 4x_1 - 8x_2 + 2x_3$$
$$\text{over } x \in \mathbf{R}^3.$$

The first order necessary conditions for a local minimum are

$$\begin{aligned}
\frac{\partial\theta(x)}{\partial x_1} &= 4x_1 + 2x_2 + 2x_3 + 4 = 0 \\
\frac{\partial\theta(x)}{\partial x_2} &= 2x_1 \qquad\quad + 4x_3 - 8 = 0 \\
\frac{\partial\theta(x)}{\partial x_3} &= 2x_1 + 4x_2 + 2x_3 + 2 = 0.
\end{aligned}$$

This system has the unique solution $\tilde{x} = (-2, -1, 3)^T$. The Hessian matrix is

$$H(\theta(\tilde{x})) = 2 \begin{pmatrix} 2 & 1 & 1 \\ 1 & 0 & 2 \\ 1 & 2 & 1 \end{pmatrix}$$

which is not PSD. So $\tilde{x}$ violates the second order necessary conditions for a local minimum. So the function $\theta(x)$ here does not have a local minimum. It can be verified that in fact $\theta(x)$ is unbounded below on $\mathbf{R}^3$.

Example 5

Let $\theta(x) = -2x_1^2 - x_2^2 + x_1x_2 - 10x_1 + 6x_2$ and consider the problem of minimizing $\theta(x)$ over $x \in \mathbf{R}^2$. The first order necessary conditions for a local minimum are

$$\begin{aligned} \frac{\partial \theta(x)}{\partial x_1} &= -4x_1 + \ x_2 - 10 = 0 \\ \frac{\partial \theta(x)}{\partial x_2} &= \quad x_1 - 2x_2 + \ 6 = 0 \end{aligned}$$

which has the unique solution $\hat{x} = (-2, 2)^T$. The Hessian matrix is

$$H(\theta(\hat{x})) = \begin{pmatrix} -4 & 1 \\ 1 & -2 \end{pmatrix}.$$

Since $H(\theta(\hat{x}))$ is not PSD, $\hat{x}$ violates the second order necessary conditions for being a local minimum of $\theta(x)$. So $\theta(x)$ has no local minimum. In fact, it can be verified that the Hessian matrix is ND, so $\hat{x}$ satisfies the sufficient condition for being a local maximum for $\theta(x)$ (a local maximum for $\theta(x)$ is a local minimum for $-\theta(x)$). Actually, $\theta(x)$ here is concave and $\hat{x}$ is a global maximum point for $\theta(x)$. It can be verified that $\theta(x)$ is unbounded below on $\mathbf{R}^2$.

An Important Caution for NLP Users

These examples point out one important aspect of using nonlinear programming algorithms in practical applications. One should not blindly accept any solution of the first order necessary optimality conditions as a solution to the problem, if it is a nonconvex programming problem (this caution can be ignored if the problem being solved is a linear or other convex programming problem). An effort should be made to check whether the solution is at least a local minimum by using second order necessary optimality conditions, or the sufficient optimality conditions, or at least through a local search in the neighborhood of the point.

Stationary Point Necessary Optimality Conditions for Constrained Minima

Consider the problem

$$\begin{array}{ll} \text{minimize} & \theta(x) \\ \text{subject to} & x \in \mathbf{\Gamma} \end{array} \tag{53}$$

where $\mathbf{\Gamma}$ is a specified subset of $\mathbf{R}^n$, and $\theta(x)$ is a real valued continuously differentiable function defined on $\mathbf{R}^n$.

Given $\overline{x} \in \mathbf{\Gamma}$, $y \neq 0$, $y \in \mathbf{R}^n$ is said to be a **feasible direction** for $\mathbf{\Gamma}$ at $\overline{x}$ if $\overline{x} + \alpha y \in \mathbf{\Gamma}$ for all $0 \leqq \alpha \leqq \lambda$ for some positive λ. As an example, if $\mathbf{\Gamma} = \{x : x = (x_1, x_2)^T, x_1 \geqq 0, x_2 \geqq 0\}$ and $\overline{x} = (1, 0)^T$, then $\{y : y = (y_1, y_2), y_2 \geqq 0\}$ is the set of feasible directions at $\overline{x}$.

Using the definition of differentiability, it follows that if $\overline{x} \in \mathbf{\Gamma}$ is a local minimum for (53), and $\theta(x)$ is continuously differentiable at $\overline{x}$, then we must have

$$(\nabla\theta(\overline{x}))y \geqq 0 \text{ for all feasible directions } y \text{ at } \overline{x} \text{ to } \mathbf{\Gamma}. \tag{54}$$

(54) are the **first order necessary conditions** for $\overline{x}$ to be a local minimum for (53). If $\theta(x)$ is twice continuously differentiable at $\overline{x} \in \mathbf{\Gamma}$, and $\overline{x}$ is a local minimum for (53), we must have

$$(54), \text{ and } y^T H(\theta(\overline{x}))y \geqq 0 \text{ for all feasible directions } y \text{ satisfying } (\nabla\theta(\overline{x}))y = 0. \tag{55}$$

The conditions (54), (55) become simplified if $\mathbf{\Gamma}$ is a convex set. In this case, a feasible direction y at $\overline{x}$ to $\mathbf{\Gamma}$ is $y = x - \overline{x}$ for any $x \in \mathbf{\Gamma}$. See Figure 14. So in case $\mathbf{\Gamma}$ is convex, the necessary conditions for $\overline{x} \in \mathbf{\Gamma}$ to be a local minimum is that (54), (55) hold for all $y = x - \overline{x}$, $x \in \mathbf{\Gamma}$.

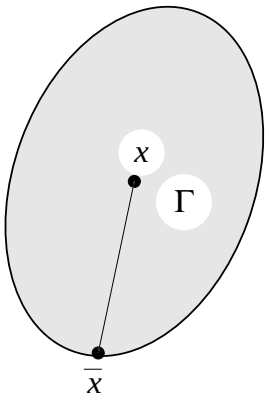


Figure 14 If $\mathbf{\Gamma}$ is a convex set, feasible directions at $\overline{x}$ to $\mathbf{\Gamma}$ are of the form $x - \overline{x}$ for any $x \in \mathbf{\Gamma}$, $x \neq \overline{x}$.

Example 6

Consider the problem

$$\begin{array}{lll} \text{minimize} & \theta(x) = 3x_1x_2 - x_1 - x_2 & \\ \text{subject to} & x_1 & \geqq 1 \\ & x_2 & \geqq 1. \end{array}$$

The set of feasible solutions, **K**, is marked in Figure 15.

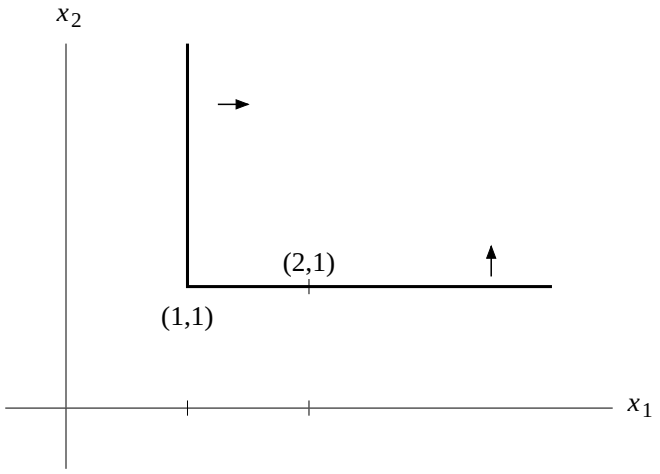


Figure 15

We have

$$\nabla\theta(x) = (3x_2 - 1, 3x_1 - 1)$$
$$H(\theta(x)) = \begin{pmatrix} 0 & 3 \\ 3 & 0 \end{pmatrix}.$$

Let $\overline{x} = (2,1)^T$. The set of feasible directions at $\overline{x}$ to **K** is clearly $\{y : y = (y_1, y_2)^T, y_2 \geqq 0\}$. $\nabla\theta(\overline{x}) = (2,5)$. $\overline{y} = (-1,0)^T$ is a feasible direction to **K** at $\overline{x}$, and yet $(\nabla\theta(\overline{x}))\overline{y} = -2 < 0$ and hence the necessary condition (54) is violated at $\overline{x}$.

Let $\hat{x} = (1,1)^T$. The set of feasible directions to **K** at $\hat{x}$ is clearly $\{y : y \geqq 0\}$. $\nabla\theta(\hat{x}) = (2,2)$ and we verify that both the necessary optimality conditions (54) and (55) are satisfied at $\hat{x}$. Acutally, $\hat{x}$ is the global minimum for this problem.

The conditions (54), (55) are respectively the **first and second order stationary point necessary optimality conditions** for the NLP (53).

Variational Inequality Problem

The stationary point necessary optimality conditions discussed above, lead to a problem commonly known as the variational inequality problem. In this problem we are given a

real vector function $f(x) = (f_1(x), \ldots, f_n(x))^T$ defined over $\mathbf{R}^n$, and a subset $\mathbf{K} \subset \mathbf{R}^n$. The variational inequality problem with this data, is to find a point $x^* \in \mathbf{K}$ satisfying

$$(x - x^*)^T f(x^*) \geqq 0 \text{ for all } x \in \mathbf{K}.$$

Suppose $\mathbf{K} = \{x : Ax \geqq b, x \geqq 0\}$ where A, b are given matrices of orders $m \times n$ and $m \times 1$, the above variational inequality problem is equivalent to the nonlinear complementarity problem: find $z \in \mathbf{R}^{n+m}$ satisfying

$$z \geqq 0,\ g(z) \geqq 0,\ z^T g(z) = 0$$

where $z = (x_1, \ldots, x_n; y_1, \ldots, y_m)^T$, $y = (y_1, \ldots, y_m)^T$ and

$$g(z) = \begin{pmatrix} f(x) & -A^T y \\ Ax & -b \end{pmatrix}.$$

Optimality Conditions for Equality Constrained Minimization

Consider the NLP

$$\begin{array}{ll} \text{minimize} & \theta(x) \\ \text{subject to} & h_i(x) = 0,\ i = 1 \text{ to } m \end{array} \tag{56}$$

where $\theta(x)$, $h_i(x)$ are all real valued continuously differentiable functions defined on $\mathbf{R}^n$. Let $h(x) = (h_1(x), \ldots, h_m(x))^T$. The set of feasible solutions is a surface in $\mathbf{R}^n$, and it is smooth if each $h_i(x)$ is a smooth function (i. e., continuously differentiable). If $\overline{x}$ is a feasible point, when some of the $h_i(x)$ are nonlinear, there may be no feasible direction at $\overline{x}$. In order to retain feasibility while moving from $\overline{x}$, one has to follow a nonlinear curve through $\overline{x}$ which lies on the feasible surface. See Figure 16.

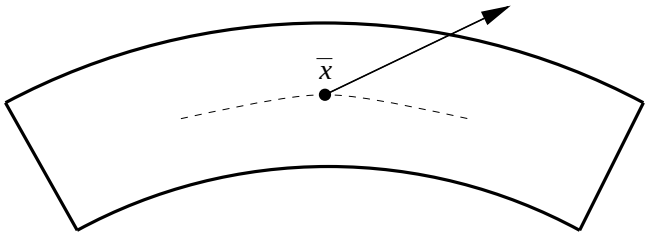


Figure 16 Feasible surface $\mathbf{\Gamma} = \{x : h_1(x) = 0\}$ satisfying a nonlinear equation. At $\overline{x} \in \mathbf{\Gamma}$, the direction marked by the arrow is not a feasible direction, since any move of positive length in that direction takes the point out of $\mathbf{\Gamma}$. To move from $\overline{x}$ and remain inside $\mathbf{\Gamma}$ one has to follow a curve like the dashed curve.

A curve in $\mathbf{R}^n$ is the locus of a point $x(\lambda) = (x_j(\lambda))$, where each $x_j(\lambda)$ is a real valued function of the real parameter λ, as the parameter varies over some interval of the real line. See Figure 17.

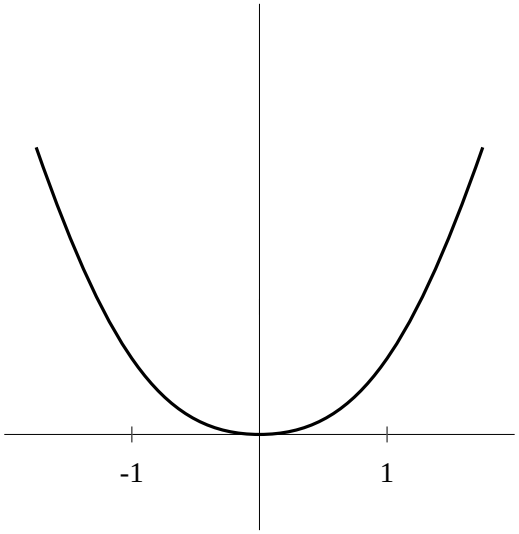


Figure 17 A curve in $\mathbf{R^2}$. $\{x(\lambda) = (\lambda, \lambda^2) : -1 \leqq \lambda \leqq 1\}$ is a piece of a curve (parabola) in $\mathbf{R}^2$ through the origin $x = (0, 0)^T$.

The curve $x(\lambda) = (x_j(\lambda)$ is said to be differentiable at λ if $\frac{dx_j(\lambda)}{d\lambda}$ exists for all j, and twice differentiable if $\frac{d^2x_j(\lambda)}{d\lambda^2}$ exists for all j. The curve $x(\lambda)$ is said to pass through the point $\overline{x}$ if $\overline{x} = x(\overline{\lambda})$ for some $\overline{\lambda}$.

If the curve $x(\lambda)$ defined over $a < \lambda < b$ is differentiable at $\overline{\lambda}$, $a < \overline{\lambda} < b$, then the line $\{x = x(\overline{\lambda}) + \delta\frac{dx}{d\lambda}(\overline{\lambda}) : \delta \text{ real}\}$ is the tangent line to the curve at the point $x(\overline{\lambda})$ on it. See Figure 18.

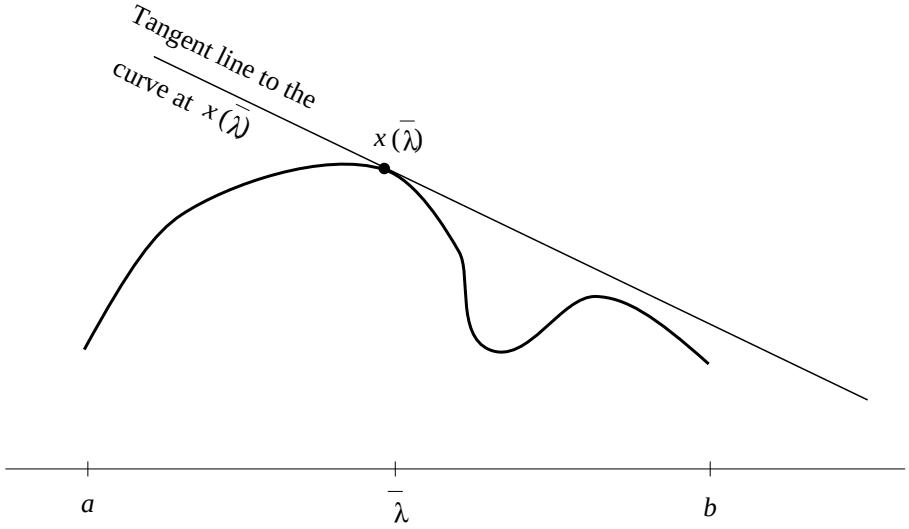


Figure 18

The **tangent plane** at a feasible point $\overline{x}$ to (56) is defined to be the set of all directions $(\frac{dx(\lambda)}{d\lambda})_{\lambda=0}$, where $x(\lambda)$ is a differential curve in the feasible region with $x(0) = \overline{x}$. See Figure 19.

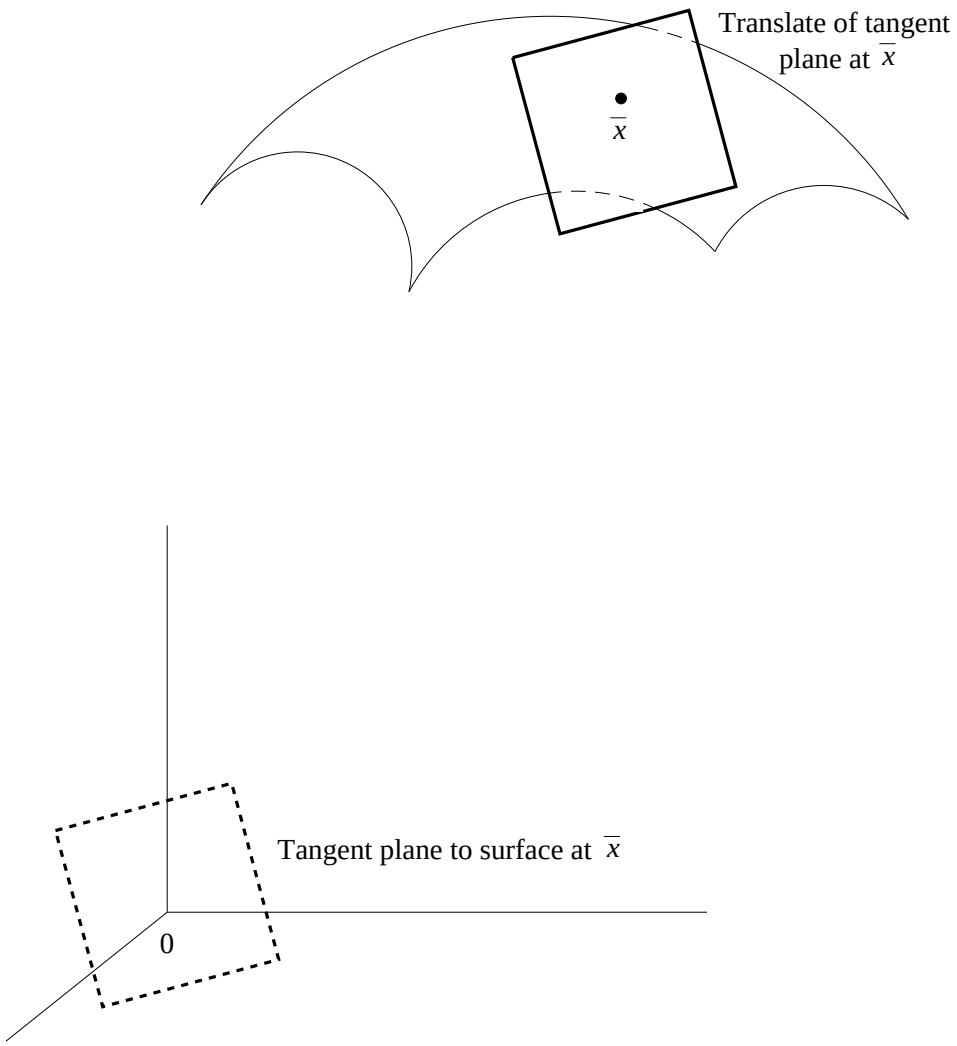


Figure 19 The tangent plane to surface $\{x : h_1(x) = 0\}$ at a point $\overline{x}$ on it is the collection of all directions of tangent lines to differentiable curves lying in surface and passing through $\overline{x}$.

We need the following results to study these tangent planes.

The Implicit Function Theorem

Consider the system of m equations in n variables $x_1, \ldots, x_n$

$$f_i(x_1, \ldots, x_n) = 0, \ i = 1 \text{ to } m \tag{57}$$

where each $f_i(x)$ is continuously differentiable in some open subset $\mathbf{D} \subset \mathbf{R}^n$. Let $\overline{x} \in \mathbf{D}$ be feasible to (57) and let the subset of m variables, $x_1, \ldots, x_m$, say, be such that the $m \times m$ Jacobian $(\frac{\partial f_i(\overline{x})}{\partial x_j} : i = 1 \text{ to } m, \ j = 1 \text{ to } m)$ is nonsingular. Then in a neighborhood of $\overline{x}$, we can use the equations in (57) to express $x_1, \ldots, x_m$ as functions of $x_{m+1}, \ldots, x_n$ on the set of feasible solutions of (57). That is, there exists a neighborhood $\mathcal{D}$ of $(\overline{x}_{m+1}, \ldots, \overline{x}_n)$ in $\mathbf{R}^{n-m}$ and real valued differentiable functions $\psi_i(x_{m+1}, \ldots, x_n)$, $i = 1$ to m; such that for $(x_{m+1}, \ldots, x_n) \in \mathcal{D}$, (57) is equivalent to

$$x_i = \psi_i(x_{m+1}, \ldots, x_n), \ i = 1 \text{ to } m$$

i. e.,

$$f_i(\psi_1(x_{m+1}, \ldots, x_n), \ldots, \psi_m(x_{m+1}, \ldots, x_n), x_{m+1}, \ldots, x_n) = 0, \ i = 1 \text{ to } m \tag{58}$$

holds for all $(x_{m+1}, \ldots, x_n) \in \mathcal{D}$. Further, the partial derivatives $\frac{\partial \psi_i(\overline{x}_{m+1}, \ldots, \overline{x}_n)}{\partial x_j}$, $i = 1$ to m, $j = m+1$ to n, are obtained by solving the system of equations

$$\sum_{r=1}^{m} \frac{\partial f_i(\overline{x})}{\partial x_r} \frac{\partial \psi_r(\overline{x}_{m+1}, \ldots, \overline{x}_n)}{\partial x_j} + \frac{\partial f_i(\overline{x})}{\partial x_j} = 0, \ j = m+1 \text{ to } n, \ i = 1 \text{ to } m. \tag{59}$$

It can be verified that (59) is just obtained by setting the derivative of the identity (58) at $\overline{x}$ with respect to x_j to zero for each $j = m+1$ to n and $i = 1$ to m. See references [10.33] for a proof of the implicit function theorem.

Example 7: An Illustration of the Implicit Function Theorem.

Here we provide a simple example to illustrate the implicit function theorem using a linear system of constraints. Consider the following system in the variables $x = (x_1, x_2, x_3, x_4, x_5)^T$.

$$\begin{aligned} f_1(x) &= \ \ x_1 + x_2 + \ \ x_3 + x_4 - \ \ x_5 - 12 = 0 \\ f_2(x) &= -x_1 + x_2 - 2x_3 - x_4 + 4x_5 - \ \ 2 = 0. \end{aligned}$$

Let $\overline{x} = (5, 7, 0, 0, 0)^T$. $\overline{x}$ is a feasible solution, and

$$\begin{bmatrix} \frac{\partial f_1(\overline{x})}{\partial x_1} & \frac{\partial f_1(\overline{x})}{\partial x_2} \\ \frac{\partial f_2(\overline{x})}{\partial x_1} & \frac{\partial f_2(\overline{x})}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

is nonsingular. Therefore, by the implicit function theorem, it is possible to express x_1, x_2 as functions of the remaining variables x_3, x_4, x_5 in a neighborhood of $\overline{x}$ in the

feasible region. Since the constraints are linear, we can do this explicitly by solving for x_1, x_2 in terms of x_3, x_4, x_5 using these two equations, and this leads to

$$x_1(x_3, x_4, x_5) = -\frac{3}{2}x_3 - x_4 - \frac{5}{2}x_5 + 10$$
$$x_2(x_3, x_4, x_5) = \quad \frac{1}{2}x_3 \qquad - \frac{3}{2}x_5 + 7$$

where $x_1(x_3, x_4, x_5)$ and $x_2(x_3, x_4, x_5)$ are the expressions for x_1, x_2 as functions of x_3, x_4, x_5, on the feasible region for this system. When the equations are nonlinear, it may not be possible to obtain these expressions explicitly, but the implicit function theorem guarantees the existence of them in a neighborhood of $\overline{x}$ in the feasible region.

We verify that the partial derivatives are

$$\begin{bmatrix} \frac{\partial x_1}{\partial x_3}, & \frac{\partial x_1}{\partial x_4}, & \frac{\partial x_1}{\partial x_5} \\ \frac{\partial x_2}{\partial x_3}, & \frac{\partial x_2}{\partial x_4}, & \frac{\partial x_2}{\partial x_5} \end{bmatrix} = \begin{bmatrix} -\frac{3}{2}, & -1, & -\frac{5}{2} \\ \frac{1}{2}, & 0, & -\frac{3}{2} \end{bmatrix}.$$

The equations corresponding to (59) for this system for $j = 3$ are

$$\frac{\partial f_1}{\partial x_1}\frac{\partial x_1}{\partial x_3} + \frac{\partial f_1}{\partial x_2}\frac{\partial x_2}{\partial x_3} + \frac{\partial f_1}{\partial x_3} = \quad \frac{\partial x_1}{\partial x_3} + \frac{\partial x_2}{\partial x_3} + 1 = 0$$
$$\frac{\partial f_2}{\partial x_1}\frac{\partial x_1}{\partial x_3} + \frac{\partial f_2}{\partial x_2}\frac{\partial x_2}{\partial x_3} + \frac{\partial f_2}{\partial x_3} = -\frac{\partial x_1}{\partial x_3} + \frac{\partial x_2}{\partial x_3} - 2 = 0$$

which together yield $\frac{\partial x_1}{\partial x_3} = -\frac{3}{2}$, $\frac{\partial x_2}{\partial x_3} = \frac{1}{2}$, same as the values obtained above. In a similar manner, writing the equations corresponding to (59) for this system for $j = 4, 5$, we can compute the values $\frac{\partial x_i}{\partial x_j}$ for $i = 1, 2$, $j = 4, 5$, and verify that they are the same as those obtained above.

Constraint Qualifications

In general, determining the tangent plane for (56) at the feasible point $\overline{x}$ is hard. However, if the constraint functions $h_i(x)$ satisfy ceratin conditions at $\overline{x}$, it becomes possible to obtain a simple characterization of the tangent plane for (56) at $\overline{x}$. So these conditions are called **constraint qualifications** because these conditions are specifically on the constraints in (56), not so much on the set of feasible solutions of (56). Several constraint qualifications have been developed, but for most of them, it is very hard to verify whether they hold in any given problem. We will only discuss one constraint qualification, which can be checked efficiently. It is called the **regularity condition**.

The regularity condition is said to hold for (56) at the feasible point $\overline{x}$ if the Jacobian matrix $(\frac{\partial h_i(\overline{x})}{\partial x_j} : i = 1 \text{ to } m, j = 1 \text{ to } n)$ has rank m, in this case the feasible point $\overline{x}$ is called a **regular point** for (56).

Definition. *We denote by* $\nabla h(x) = (\frac{\partial h_i(x)}{\partial x_j} : i = 1 \text{ to } m, j = 1 \text{ to } n)$, *the Jacobian matrix of order* $m \times n$; *the* i*th row vector of* $\nabla h(x)$ *is the gradient vector of* $h_i(x)$ *written as a row vector.*

Tangent Planes at Regular Points

Theorem 20. *If $\overline{x}$ is a regular point for (56), the tangent plane for (56) at $\overline{x}$ is* $\{y : (\nabla h(\overline{x}))y = 0\}$.

Proof. Let $x(\alpha)$ be a differentiable curve lying in the feasible region for α lying in an interval around zero, with $x(0) = \overline{x}$ and $\frac{dx(0)}{d\alpha} = y$. So $h(x(\alpha)) = 0$ for all values of α lying in an interval around zero, and hence $(\frac{dh(x(\alpha))}{d\alpha})_{\alpha=0} = 0$, that is $(\nabla h(\overline{x}))y = 0$. This implies that the tangent plane is a subset of $\{y : (\nabla h(\overline{x}))y\} = 0$.

Suppose $y \in \{y : (\nabla h(\overline{x}))y = 0\}$ and $y \neq 0$. Define new variables $u = (u_1, \ldots, u_m)^T$. Consider the following system of m equations in $m+1$ variables $u_1, \ldots, u_m$, α.

$$g_i(u, \alpha) = h_i(\overline{x} + \alpha y + (\nabla h(\overline{x}))^T u) = 0, \ i = 1 \text{ to } m. \tag{60}$$

It can be verified that $g(0, 0) = 0$ and the Jacobian matrix of $g(u, \alpha)$ with respect to u is nonsigular at $u = 0$, $\alpha = 0$ (since $\overline{x}$ is a regular point of (56)). So by applying the implicit function theorem on (60), we can express u as a differentiable function of α, say $u(\alpha)$, in an interval around $\alpha = 0$, and that (60) holds as an identity in this interval when u in (60) is replaced by $u(\alpha)$, and that $u(0) = 0$, and $\frac{du(0)}{d\alpha}$ is obtained by solving

$$\left(\frac{d}{d\alpha} h(\overline{x} + \alpha y + (\nabla h(\overline{x}))^T u(\alpha))\right)_{\alpha=0} = 0$$

which leads to $\frac{d}{d\alpha} u(0) = 0$ since $\nabla h(\overline{x})$ has rank m. So if we define

$$x(\alpha) = \overline{x} + \alpha y + (\nabla h(\overline{x}))^T u(\alpha)$$

this defines a differentiable curve lying in the feasible region for (56) for values of α in an interval around $\alpha = 0$, and that $\frac{dx}{d\alpha}(0) = y$, which implies that y is in the tangent plane for (56) at $\overline{x}$.

□

Example 8

Consider the system

$$h(x_1, x_2) = x_1 = 0$$
$$x = (x_1, x_2)^T \in \mathbf{R}^2.$$

The set of feasible solutions is the x_2-axis in $\mathbf{R}^2$, since $\nabla h(x) = (1, 0)$ every feasible point is a regular point, and the tangent plane at any feasible point x is again the x_2-axis $= \{y : (\nabla h(\overline{x}))y = 0\} = \{y : y = (y_1, y_2), y_1 = 0\}$. On the other hand the system

$$g(x_1, x_2) = x_1^{\mathbf{3}} = 0$$
$$x = (x_1, x_2)^T \in \mathbf{R}^2$$

has the same set of feasible solutions, namely the x_2-axis in $\mathbf{R}^2$. Since $\nabla g(x) = (3x_1^{\mathbf{2}}, 0)$ is zero whenever x is feasible, no feasible point is regular. The tangent plane at every feasible solution is again the x_2-axis in $\mathbf{R}^2$, but $\{y : \nabla g(x)y = 0\} = \mathbf{R}^2$ for every feasible solution x.

Optimality Conditions

Using Theorem 20 we can now derive optimality conditions for (56). If $\overline{x}$ is a feasible regular point for (56), and it is a local minimum, clearly along every differentiable curve $x(\alpha)$ lying in the feasible region for (56) for values of α in an interval around $\alpha = 0$, satisfying $x(0) = \overline{x}$; $\alpha = 0$ must be a local minimum for $\theta(x)$ on this curve. That is, for the problem of minimizing $\theta(x(\alpha))$ over this interval for α, $\alpha = 0$ must be a local minimum. Since $\alpha = 0$ is an interior point of this interval this implies that $\frac{d\theta}{d\alpha}(x(0))$ must be zero. Applying this to all such curves and using Theorem 20 we conclude that $(\nabla\theta(\overline{x}))y = 0$ for all y satisfying $(\nabla h(\overline{x}))y = 0$. By Theorem 1 (see Exercise 5) this implies that there must exist $\overline{\mu} = (\overline{\mu}_1, \ldots, \overline{\mu}_m)$ such that

$$\begin{gathered} \nabla\theta(\overline{x}) - \sum_{i=1}^{m} \overline{\mu}_i \nabla h_i(\overline{x}) = 0 \\ \text{and by feasibility } h(\overline{x}) = 0 \end{gathered} \tag{61}$$

the conditions (61) are the **first order necessary optimality conditions** for (56), the vector $\overline{\mu}$ is the vector of **Lagrange multipliers**. (61) is a system of $(n + m)$ equations in $(n + m)$ unknowns (including $\overline{x}$ and $\overline{\mu}$) and it may be possible to solve (61) using algorithms for solving nonlinear equations. If we define the Lagrangian for (56) to be $L(x, \mu) = \theta(x) - \mu h(x)$ where $\mu = (\mu_1, \ldots, \mu_m)$, $h(x) = (h_1(x), \ldots, h_m(x))^T$, (61) becomes: $(\overline{x}, \overline{\mu})$ satisfies

$$\begin{aligned} h(x) &= 0 \\ \nabla_x L(x, \mu) &= 0. \end{aligned} \tag{62}$$

We will now derive the second order necessary optimality conditions for (56). Suppose the functions $\theta(x)$, $h_i(x)$ are all twice continuously differentiable. Let $\overline{x}$ be a feasible solution for (56) which is a regular point. If $\overline{x}$ is a local minimum for (56), by the first order necessary optimality conditions (61), there must exist a row vector of Lagrange multipliers, $\overline{\mu} = (\overline{\mu}_1, \ldots, \overline{\mu}_m)$ such that $\nabla_x(L(\overline{x}, \overline{\mu})) = 0$, where $L(x, \overline{\mu}) = \theta(x) - \overline{\mu} h(x)$ is the Lagrangian. Since $\overline{x}$ is a regular point, the tangent plane to (56) at $\overline{x}$ is $\mathbf{T} = \{y : (\nabla h(\overline{x}))y = 0\}$. Suppose there exists a $\overline{y} \in \mathbf{T}$ satisfying $\overline{y}^T H_x(L(\overline{x}, \overline{\mu}))\overline{y} < 0$. Since $\overline{y} \in \mathbf{T}$, and all the functions are twice continuously differentiable, there exists a twice differentiable curve $x(\lambda)$ through $\overline{x}$ lying in the feasible region (i. e., $x(0) = \overline{x}$, and the curve is defined in an interval of λ with 0 as an interior point, with $h(x(\lambda)) = 0$ for all λ in this interval), such that $(\frac{dx(\lambda)}{d\lambda})_{\lambda=0} = \overline{y}$. Now,

$$\begin{aligned} \frac{d}{d\lambda} L(x(\lambda), \overline{\mu}) &= (\nabla_x L(x(\lambda), \overline{\mu}))\Big(\frac{dx(\lambda)}{d\lambda}\Big) \\ \frac{d^2}{d\lambda^2} L(x(\lambda), \overline{\mu}) &= \Big(\frac{dx(\lambda)}{d\lambda}\Big)^T H_x(L(x(\lambda), \overline{\mu}))\frac{dx(\lambda)}{d\lambda} + (\nabla_x L(x(\lambda), \overline{\mu}))\Big(\frac{d^2x(\lambda)}{d\lambda^2}\Big) \end{aligned}$$

where $\nabla_x(L(\overline{x}, \overline{\mu}))$, $H_x(L(\overline{x}, \overline{\mu}))$ are the row vector of partial derivatives with respect to x, and the Hessian matrix with respect to x of $L(x, \overline{\mu})$ at $x = \overline{x}$ respectively. At

$\lambda = 0$, we have $\nabla_x L(x(0), \overline{\mu}) = \nabla_x L(\overline{x}, \overline{\mu}) = 0$ by the first order necessary optimality conditions.

So, from the above

$$\Big(\frac{d}{d\lambda} L(x(\lambda), \overline{\mu})\Big)_{\lambda=0} = 0$$

$$\Big(\frac{d^2}{d\lambda^2} L(x(\lambda), \overline{\mu})\Big)_{\lambda=0} = \overline{y}^T H_x(L(\overline{x}, \overline{\mu}))\overline{y}.$$

Using these in a Taylor series expansion for $f(\lambda) = L(x(\lambda), \overline{\mu})$ up to second order around $\lambda = 0$ leads to

$$f(\lambda) = L(x(\lambda), \overline{\mu}) = L(\overline{x}, \overline{\mu}) + \frac{\lambda^{\mathbf{2}}}{2} \overline{y}^T H_x(L(\overline{x}, \overline{\mu}))\overline{y} + 0(\lambda)$$

where $0(\lambda)$ is a function of λ satisfying the property that limit of $(0(\lambda))/\lambda^{\mathbf{2}}$ as λ tends to zero, is zero. Since $h(x(\lambda)) = 0$ for every point on the curve, we have $f(\lambda) = L(x(\lambda), \overline{\mu}) = \theta(x(\lambda))$ for all λ in the interval of λ on which the curve is defined. So in the neighborhood of $\lambda = 0$ on the curve we have from the above

$$\frac{2(\theta(x(\lambda)) - \theta(\overline{x}))}{\lambda^{\mathbf{2}}} = \frac{2(f(\lambda) - f(0))}{\lambda^{\mathbf{2}}} = \overline{y}^T H_x(L(\overline{x}, \overline{\mu}))\overline{y} + \frac{2(0(\lambda))}{\lambda^{\mathbf{2}}}$$

and since $\overline{y}^T H_x(L(\overline{x}, \overline{\mu}))\overline{y} < 0$ and limit of $(0(\lambda)/\lambda^{\mathbf{2}})$ as λ tends to zero is zero, for all λ sufficiently small $\theta(x(\lambda)) - \theta(\overline{x}) < 0$. For all these λ, $x(\lambda)$ is a point on the curve in the feasible region in the neighborhood of $\overline{x}$, and this is a contradiction to the fact that $\overline{x}$ is a local minimum for (56).

In fact it can be verified that $\overline{y}^T H_x(L(\overline{x}, \overline{\mu}))\overline{y} = (\frac{d^2 f(\lambda)}{d\lambda^2})_{\lambda=0}$, and if this quantity is < 0, $\lambda = 0$ cannot be a local minimum for the one variable minimization problem of minimizing $f(\lambda) = \theta(x(\lambda))$ over λ; or equivalently, that $\overline{x} = x(0)$ is not a local minimum for $\theta(x)$ along the curve $x(\lambda)$.

These facts imply that if $\theta(x)$, $h_i(x)$ are all twice continuously differentiable, and $\overline{x}$ is a regular point which is a feasible solution and a local minimum for (56), there must exist a Lagrange multiplier vector $\overline{\mu}$ such that the following conditions hold.

$$\begin{aligned} h(\overline{x}) &= 0 \\ \nabla_x L(\overline{x}, \overline{\mu}) &= \nabla\theta(\overline{x}) - \overline{\mu}\nabla h(\overline{x}) = 0 \\ y^T H_x(L(\overline{x}, \overline{\mu}))y &\geqq 0 \text{ for all } y \in \mathbf{T} = \{y : (\nabla h(\overline{x}))y = 0\}, \\ \text{that is } H_x(L(\overline{x}, \overline{\mu})) &\text{ is PSD on the subspace } \mathbf{T}. \end{aligned} \tag{63}$$

These are the **second order necessary optimality conditions** for a regular feasible point $\overline{x}$ to be a local minimum for (56).

We now state a sufficient optimality condition for (56) in the form of a theorem.

Theorem 21. *Suppose $\theta(x)$, $h_i(x)$, $i = 1$ to m are all twice continuously differentiable functions, and $\overline{x}$ is a feasible point such that there exists a Lagrange multiplier vector $\overline{\mu} = (\overline{\mu}_1, \ldots, \overline{\mu}_m)$ which together satisfy*

$$\begin{aligned} h(\overline{x}) &= 0 \\ \nabla\theta(\overline{x}) - \overline{\mu}\nabla h(\overline{x}) &= 0 \\ y^T H_x(L(\overline{x}, \overline{\mu}))y > 0 \text{ for all } y \in \{y : (\nabla h(\overline{x}))y = 0\}, y \neq 0 \end{aligned} \tag{64}$$

where $L(x, \overline{\mu}) = \theta(x) - \overline{\mu}h(x)$ is the Lagrangian for (56). Then $\overline{x}$ is a local minimum for (56).

Proof. Suppose $\overline{x}$ is not a local minimum for (56). There must exist a sequence of distinct feasible points $\{x^r : r = 1, 2, \ldots\}$ converging to $\overline{x}$ such that $\theta(x^r) < \theta(\overline{x})$ for all r. Let $\delta_r = \|\overline{x} - x^r\|$, $y^r = (x^r - \overline{x})/\delta_r$. Then $\|y^r\| = 1$ for all r and $x^r = \overline{x} + \delta_r y^r$. Thus $\delta_r \to 0^+$ as $r \to \infty$. Since the sequence of points $\{y^r : r = 1, 2, \ldots)$ all lie on the surface of the unit sphere in $\mathbf{R}^n$, a compact set, the sequence has at least one limit point. Let $\overline{y}$ be a limit point of $\{y^r : r = 1, 2, \ldots\}$. There must exist a subsequence of $\{y^r : r = 1, 2, \ldots\}$ which converges to $\overline{y}$, eliminate all points other than those in this subsequence, and for simplicity call the remaining sequence by the same notation $\{y^r : r = 1, 2, \ldots\}$. So now we have a sequence of points $x^r = \overline{x} + \delta_r y^r$ all of them feasible, such that $\|y^r\| = 1$ for all $r, y^r \to \overline{y}$ and $\delta_r \to 0$ as $r \to \infty$. By feasibility $h(\overline{x} + \delta_r y^r) = 0$ for all r, and by the differentiability of $h(x)$ we have

$$\begin{aligned} 0 = h(\overline{x} + \delta_r y^r) &= h(\overline{x}) + \delta_r \nabla h(\overline{x}) y^r + 0(\delta_r) \\ &= \delta_r \nabla h(\overline{x}) y^r + 0(\delta_r) \end{aligned}$$

Dividing by $\delta_r > 0$, and taking the limit as $r \to \infty$ we see that $\nabla h(\overline{x})\overline{y} = 0$.

Since $L(x, \overline{\mu})$ is a twice continuously differentiable function in x, applying Taylor's theorem to it, we conclude that for each r, there exists a $0 \leqq \alpha_r \leqq \delta_r$ such that

$$L(\overline{x} + \delta_r y^r, \overline{\mu}) = L(\overline{x}, \overline{\mu}) + \delta_r \nabla_x L(\overline{x}, \overline{\mu}) y^r + (1/2)\delta_r^2 (y^r)^T H_x(L(\overline{x} + \alpha_r y^r, \overline{\mu}))y^r.$$

From the fact that $\overline{x} + \delta_r y^r = x^r$ and $\overline{x}$ are feasible, we have $L(x^r, \overline{\mu}) = \theta(x^r)$ and $L(\overline{x}, \overline{\mu}) = \theta(\overline{x})$. Also, from (64), $\nabla_x L(\overline{x}, \overline{\mu}) = 0$. So, from the above equation, we have

$$\theta(x^r) - \theta(\overline{x}) = (1/2)\delta_r^2 (y^r)^T H_x(L(\overline{x} + \alpha_r y^r, \overline{\mu}))y^r. \tag{65}$$

Since $0 \leqq \alpha_r \leqq \delta_r$, and $\delta_r \to 0$ as $r \to \infty$, and by continuity, we know that $H_x(L(\overline{x} + \alpha_r y^r, \overline{\mu}))$ converges to $H_x(L(\overline{x}, \overline{\mu}))$ as $r \to \infty$. Since $y^r \to \overline{y}$ as $r \to \infty$, and $\nabla h(\overline{x})\overline{y} = 0$, from the last condition in (64) and continuity we conclude that when r is sufficiently large, the right-hand side of (65) is ≥ 0, while the left-hand side is < 0, a contradiction. So, $\overline{x}$ must be a local minimum for (56).

□

Thus, (64) provides a **sufficient condition** for a feasible point $\overline{x}$ to be a local minimum for (56).

Example 9

Consider the problem

$$\begin{array}{ll} \text{minimize} & \theta(x) = x_1x_2 \\ \text{subject to} & x_1 + x_2 = 2. \end{array}$$

The Lagrangian is $L(x, \lambda) = x_1x_2 - \lambda(x_1 + x_2 - 2)$. So, the first order necessary optimality conditions are

$$\frac{\partial L(x, \lambda)}{\partial x} = (x_2 - \lambda, x_1 - \lambda) = 0$$

which together with the feasibility conditions lead to $\overline{x} = (1, 1)^T$. $\overline{x}$ is the unique solution for the first order necessary optimality conditions. $\overline{x}$, $\overline{\lambda} = 1$ together satisfy the first order necessary conditions for a local minimum. The Hessian of the Lagrangian is

$$H_x(L(\overline{x}, \overline{\lambda})) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

The tangent plane at $\overline{x}$ is $\{y : y_1 + y_2 = 0\}$. So on the tangent plane, $y^T H_x((\overline{x}, \overline{\lambda}))y = 2y_1y_2 = -2y_2^2 < 0$, whenever $y \neq 0$. So the second order necessary optimality conditions for a local minimum are violated at $\overline{x}$. In fact it can be verified that $\overline{x}$ satisfies the sufficient conditions for being a local maximum for $\theta(x)$ in the feasible region. $\theta(x)$ has no local minimum in the feasible region, it is unbounded below in the feasible region.

Example 10

Consider the problem

$$\begin{array}{ll} \text{minimize} & -x_1 - x_2 \\ \text{subject to} & x_1^2 + x_2^2 - 8 = 0. \end{array}$$

The Lagrangian is $L(x, \lambda) = -x_1 - x_2 - \lambda(x_1^2 + x_2^2 - 8)$. The first order necessary optimality conditions are

$$\frac{\partial L(x, \lambda)}{\partial x} = \begin{pmatrix} -1 & -2x_1\lambda \\ -1 & -2x_2\lambda \end{pmatrix}^T = 0$$

together with the constraint on the variables, this leads to the unique solution $\overline{x} = (2, 2)^T$, $\overline{\lambda} = -1/4$. The Hessian of the Lagrangian is

$$H_x(L(\overline{x}, \overline{\lambda})) = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix}$$

which is PD. Hence the point $\overline{x}$ satisfies the sufficient condition for being a local minimum in this problem.

Example 11

Consider the problem

$$\begin{array}{ll} \text{minimize } \theta(x) & = 2x_1^{\mathbf{3}} + (1/2)x_2^{\mathbf{2}} + x_1x_2 + (1/24)x_1 \\ \text{subject to} & x_1 + x_2 = 2. \end{array}$$

The Lagrangian is $L(x, \lambda) = 2x_1^{\mathbf{3}} + (1/2)x_2^{\mathbf{2}} + x_1x_2 + (1/24)x_1 - \lambda(x_1 + x_2 - 2)$. The first order necessary optimality conditions are

$$\frac{\partial L(x,\lambda)}{\partial x} = \begin{pmatrix} 6x_1^{\mathbf{2}} + x_2 + (1/24) - \lambda \\ x_2 + x_1 - \lambda \end{pmatrix}^T = 0.$$

Combining this with the constraints on the variables, we have $\lambda = 2$, $6x_1^{\mathbf{2}} + x_2 + (1/24) - 2 = 6x_1^{\mathbf{2}} + (2 - x_1) - 2 + (1/24) = 6x_1^{\mathbf{2}} - x_1 + (1/24) = 0$. This leads to the unique solution satisfying the first order necessary optimality conditions ($\overline{x} = (1/12, 23/12)^T$, $\overline{\lambda} = 2$). The tangent hyperplane at any feasible solution is $\{y : y_1 + y_2 = 0\}$. The Hessian of the Lagrangian is

$$H_x(L(\overline{x}, \overline{\lambda})) = \begin{pmatrix} 12\overline{x}_1 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

So, on the tangent hyperplane to the feasible region at $\overline{x}$ we have $y^T H_x(L(\overline{x}, \overline{\lambda}))y = (y_1 + y_2)^2 = 0$. Thus the second order necessary conditions for a local minimum are also satisfied. However, the point $\overline{x}$ does not satisfy the sufficient conditions for being a local minimum in this problem, (64), discussed above.

Optimality Conditions for the Inequality Constrained Minimization Problems

Consider the general NLP

$$\begin{array}{lll} \text{minimize} & \theta(x) & \\ \text{subject to} & h_i(x) = 0, \ i = 1 \text{ to } m & \\ & g_p(x) \geqq 0, \ p = 1 \text{ to } t & \end{array} \tag{66}$$

where $\theta(x)$, $h_i(x)$, $g_p(x)$ are all real valued continuously differentiable functions defined on $\mathbf{R}^n$. Let $h(x) = (h_1(x), \ldots, h_m(x))^T$ and $g(x) = (g_1(x), \ldots, g_t(x))^T$.

Let $\overline{x}$ be a feasible solution for (66). The active constraints at $\overline{x}$ are all the equality constraints in (66) and all the inequality constraints which hold as equations at $\overline{x}$ (i. e., $g_p(x)$ for p such that $g_p(\overline{x}) = 0$). Let $\mathbf{P}(\overline{x}) = \{p : p = 1 \text{ to } t, g_p(\overline{x}) = 0\}$. The feasible solution $\overline{x}$ is said to be a regular point for (66) if $\{\nabla h_i(\overline{x}) : i = 1 \text{ to } m\} \cup \{\nabla g_p(\overline{x}) : p \in \mathbf{P}(\overline{x})\}$ is linearly independent. This is a **constraint qualification** known as the **regularity condition** for (66). As mentioned earlier, this is a condition on the active

constraints at $\overline{x}$, and not on the set of feasible solutions. As an example, consider the system of constraints

$$\begin{array}{rcrcl} (x_1-1)^{\mathbf{2}} & + & (x_2-1)^{\mathbf{2}} & = & 0 \\ x_1^{\mathbf{4}} & + & x_2^{\mathbf{4}} & = & 2 \\ x_1 & & & \leqq & 1 \\ & & x_2 & \leqq & 1 \\ x_1 & + & x_2 & \geqq & 2\,. \end{array}$$

This system has the unique solution $(x_1, x_2)^T = (1, 1)^T$, all the constraints are active and it can be verified that the regularity condition does not hold at this point. On the other hand, if this singleton set is represented by the system of constraints

$$\begin{array}{rcl} x_1 & = & 1 \\ x_2 & = & 1 \end{array}$$

then the regularity condition holds at the point. Thus, whether regularity conditions hold or not could depend on the system of constraints chosen to represent the set of feasible solutions. This points out the importance of exercising great care in constructing the model for the problem.

Since the inequality constraints "$g_i(x) \geqq 0$" for $i \notin \mathbf{P}(\overline{x})$ are inactive at $\overline{x}$, the local feasible region around $\overline{x}$ remains unchanged if these inactive inequality constraints are ignored. See Figure 20.

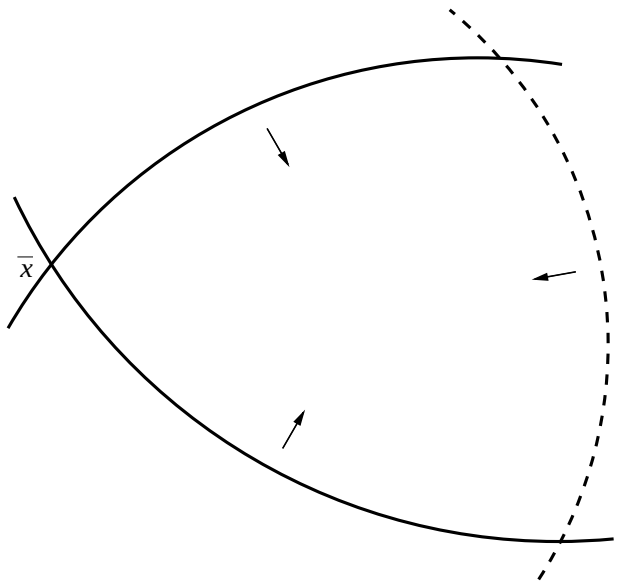


Figure 20 The region which lies on the side of the arrow of each nonlinear surface is the feasible region. The inequality constraint corresponding to the dashed surface is inactive at $\overline{x}$, and it can be ignored for the purpose of deriving optimality conditions for $\overline{x}$ to be a local minimum in the feasible region.

Thus for the purpose of deriving optimality conditions for $\overline{x}$ to be a local minimum for (66), we can ignore the inactive inequality constraints at $\overline{x}$. Also, when all the active

constraints at $\overline{x}$ are treated as equality constraints, the local feasible region around $\overline{x}$ becomes smaller, and hence, if $\overline{x}$ is a local minimum for (66), it must be a local minimum for the problem obtained by treating all active constraints at $\overline{x}$ as equality constraints.

Let $\overline{x}$, a feasible regular point for (66), be a local minimum for (66). By the above arguments, it must be a local minimum for the problem,

$$\begin{array}{ll} \text{minimize} & \theta(x) \\ \text{subject to} & h_i(x) = 0,\ i = 1 \text{ to } m \\ & g_p(x) = 0,\ p \in \mathbf{P}(\overline{x}). \end{array} \tag{67}$$

So by previous results, there exists $(\overline{\mu}_1, \ldots, \overline{\mu}_m)$ and $\overline{\pi}_p$ for $p \in \mathbf{P}(\overline{x})$ satisfying

$$\nabla\theta(\overline{x}) - \sum_{i=1}^{m} \overline{\mu}_i \nabla h_i(\overline{x}) - \sum_{p \in \mathbf{P}(\overline{x})} \overline{\pi}_p \nabla g_p(\overline{x}) = 0. \tag{68}$$

We will now prove that if $\overline{x}$ is a local minimum for (66), then $\overline{\pi}_p \geqq 0$ for all $p \in \mathbf{P}(\overline{x})$.

Suppose in (68), $\overline{\pi}_p < 0$ for some $p \in \mathbf{P}(\overline{x})$, say for $p = r$. By the regularity condition, the set $\{\nabla h_i(\overline{x}) : i = 1 \text{ to } m\} \cup \{\nabla g_p(\overline{x}) : p \in \mathbf{P}(\overline{x})\}$ is linearly independent, and by our assumption $r \in \mathbf{P}(\overline{x})$. So there exists a $y \in \mathbf{R}^n$ satisfying

$$\begin{array}{l} (\nabla h_i(\overline{x}))y = 0,\ i = 1 \text{ to } m \\ (\nabla g_p(\overline{x}))y = 0,\ p \in \mathbf{P}(\overline{x}), p \neq r \\ (\nabla g_r(\overline{x}))y = 1. \end{array} \tag{69}$$

By Theorem 20 there exists a differentiable curve $x(\alpha)$ with $x(0) = \overline{x}$, defined for values of α in an interval around $\alpha = 0$, lying on the set of feasible solutions of

$$\begin{array}{l} h_i(x) = 0,\ i = 1 \text{ to } m \\ g_p(x) = 0,\ p \in \mathbf{P}(\overline{x}), p \neq r \end{array} \tag{70}$$

with $\frac{dx(0)}{d\alpha} = y$. Since $(\frac{dg_r(x(\alpha))}{d\alpha})_{\alpha=0} = (\nabla g_r(\overline{x}))y = 1 > 0$, by Taylor's theorem we know that there exists a $\lambda > 0$ such that for all $0 \leqq \alpha \leqq \lambda$, points on the curve $x(\alpha)$ satisfy $g_r(x) \geqq 0$. Using this, it can be verified that when α is positive but sufficiently small, $x(\alpha)$ remains feasible to (66) and since $(\frac{d\theta(x(\alpha))}{d\alpha})_{\alpha=0} = \overline{\pi}_r(\nabla g_r(\overline{x}))y$ (by (68)) < 0, it is a better feasible solution for (66) than $\overline{x}$, contradicting the local minimum property of $\overline{x}$. Thus if $\overline{x}$ is a local minimum for (66) and is a regular point, there must exist $\overline{\mu} = (\overline{\mu}_1, \ldots, \overline{\mu}_m)$, and $\overline{\pi}_p$ for $p \in \mathbf{P}(\overline{x})$ satisfying (68), and $\overline{\pi}_p \geqq 0$ for all $p \in \mathbf{P}(\overline{x})$. Define $\overline{\pi}_p = 0$ for all $p = 1$ to t, $p \notin \mathbf{P}(\overline{x})$ and let $\overline{\pi} = (\overline{\pi}_1, \ldots, \overline{\pi}_t)$. Let $L(x, \mu, \pi) = \theta(x) - \mu h(x) - \pi g(x)$. $L(x, \mu, \pi)$ is the Lagrangian for (66) and $(\overline{\mu}, \overline{\pi})$ are the Lagrange multipliers. These facts imply that if $\overline{x}$ is a regular point local minimum for (66), there exist $\overline{\mu}$, $\overline{\pi}$ satisfying

$$\begin{array}{c} \nabla_x L(\overline{x}, \overline{\mu}, \overline{\pi}) = 0 \\ \overline{\pi} \geqq 0 \\ \overline{\pi}_p g_p(\overline{x}) = 0 \text{ for all } p = 1 \text{ to } t \end{array} \tag{71}$$

and the feasible conditions

$$h(\overline{x}) = 0,\ g(\overline{x}) \geqq 0.$$

(71) are known as the **first order necessary optimality conditions** for the regular feasible point $\overline{x}$ to be a local minimum for (66). They are also known as the Karush-Kuhn-Tucker (or KKT) necessary conditions for optimality.

Let $\mathbf{T} = \{y : (\nabla h_i(\overline{x}))y = 0,\ i = 1 \text{ to } m, \text{ and } (\nabla g_p(\overline{x}))y = 0, p \in \mathbf{P}(\overline{x})\}$. If all the functions $\theta(x)$, $h_i(x)$, $g_p(x)$ are twice continuously differentiable, and $\overline{x}$ is a regular feasible point for (66), using similar arguments as before, it can be shown that a necessary condition for $\overline{x}$ to be a local minimum for (66) is that there exist Lagrange multiplier vectors $\overline{\mu}$, $\overline{\pi}$ such that

$$\text{(71) holds, and } y^T H_x(L(\overline{x}, \overline{\mu}, \overline{\pi}))y \geqq 0 \text{ for all } y \in \mathbf{T}. \tag{72}$$

(72) are known as **second order necessary conditions** for $\overline{x}$ to be a local minimum for (66).

We now state a sufficient optimality condition for (66) in the form of a theorem.

Theorem 22. *Suppose $\theta(x)$, $h_i(x)$, $g_p(x)$ are all twice continuously differentiable functions, and $\overline{x}$ is a feasible point such that there exists Lagrange multiplier vectors $\overline{\mu} = (\overline{\mu}_1, \ldots, \overline{\mu}_m)$, $\overline{\pi} = (\overline{\pi}_1, \ldots, \overline{\pi}_t)$ which together satisfy*

$$\begin{aligned} &h(\overline{x}) = 0,\ g(\overline{x}) \geqq 0 \\ &\nabla_x L(\overline{x}, \overline{\mu}, \overline{\pi}) = 0 \\ &\overline{\pi} \geqq 0,\ \overline{\pi} g(\overline{x}) = 0 \end{aligned} \tag{73}$$

$$y^T H_x(L(\overline{x}, \overline{\mu}, \overline{\pi}))y > 0, \quad \text{for all } y \in \mathbf{T}_1,\ y \neq 0$$

where $\mathbf{T}_1 = \{y : (\nabla h_i(\overline{x}))y = 0,\ i = 1 \text{ to } m \text{ and } (\nabla g_p(\overline{x}))y = 0 \text{ for } p \in \mathbf{P}(\overline{x}) \cap \{p : \overline{\pi}_p > 0\},\ (\nabla g_p(\overline{x}))y \geqq 0 \text{ for } p \in \mathbf{P}(\overline{x}) \cap \{p : \overline{\pi}_p = 0\}\}$, then $\overline{x}$ is a local minimum for (66).

Proof. Suppose $\overline{x}$ is not a local minimum for (66). As in the proof of Theorem 21, there must exist a sequence of distinct feasible solutions $x^r = \overline{x} + \delta_r y^r$, $r = 1, 2, \ldots$ converging to $\overline{x}$ as $r \to 0^+$, where $\|y^r\| = 1$ for all r; $y^r \to \overline{y}$ and $\delta_r \to 0^+$; such that $\theta(x^r) < \theta(\overline{x})$ for all r. By feasibility, as in the proof of Theorem 21, we have

$$(\nabla h_i(\overline{x}))\overline{y} = 0,\ i = 1 \text{ to } m. \tag{74}$$

For each $p \in \mathbf{P}(\overline{x})$, we have $g_p(\overline{x}) = 0$, and $g_p(x^r) \geqq 0$ by feasibility. So

$$0 \leqq g_p(\overline{x} + \delta_r y^r) - g_p(\overline{x}) = \delta_r(\nabla g_p(\overline{x}))y^r + 0(\delta_r)$$

Dividing by $\delta_r > 0$, and taking the limit as $r \to \infty$, we conclude that

$$(\nabla g_p(\overline{x}))\overline{y} \geqq 0 \text{ for all } p \in \mathbf{P}(\overline{x}). \tag{75}$$

Also, $0 > \theta(\overline{x} + \delta_r y^r) - \theta(\overline{x}) = \delta_r(\nabla\theta(\overline{x}))y^r + 0(\delta_r)$, and again dividing by $\delta_r > 0$, and taking the limit as $r \to \infty$ we conclude that $(\nabla\theta(\overline{x}))\overline{y} \leqq 0$.

Suppose $(\nabla g_p(\overline{x}))\overline{y} > 0$ for some $p \in \mathbf{J} = \{p : \overline{\pi}_p > 0\}$. Then

$$\begin{aligned} 0 \geqq (\nabla\theta(\overline{x}))\overline{y} &= \overline{\mu}(\nabla h(\overline{x}))\overline{y} + \overline{\pi}(\nabla g(\overline{x}))\overline{y}, \text{ by (73)} \\ &= \sum_{p\in\mathbf{J}} \overline{\pi}_p(\nabla g_p(\overline{x}))\overline{y}, \text{ by (73), (74).} \\ &> 0, \text{ by (73), (75) and the assumption that} \\ &\quad (\nabla g_p(\overline{x}))\overline{y} > 0 \text{ for some } p \in \mathbf{J} \end{aligned}$$

a contradiction. So $\overline{y}$ satisfies

$$(\nabla g_p(\overline{x}))\overline{y} = 0 \text{ for all } p \in \mathbf{P}(\overline{x}) \cap \{p : \overline{\pi}_p > 0\}. \tag{76}$$

By (74), (75), (76), we see that $\overline{y} \in \mathbf{T}_1$. From (73) and feasibility we have

$$\begin{aligned} &\theta(x^r) - \theta(\overline{x}) = L(\overline{x} + \delta_r y^r, \overline{\mu}, \overline{\pi}) - L(\overline{x}, \overline{\mu}, \overline{\pi}) = \\ &(1/2)\delta_r^{\mathbf{2}}(y^r)^T H_x(L(\overline{x} + \alpha_r y^r, \overline{\mu}, \overline{\pi}))y^r \end{aligned} \tag{77}$$

where $0 \leqq \alpha_r \leqq \delta_r$, by using (73) on the expression given by Taylor's theorem. When r is sufficiently large, from the continuity, and the conditions satisfied by $\overline{y}$ proved above, and (73), we conclude that the right-hand side of (77) is $\geqq 0$, while $\theta(x^r) - \theta(\overline{x})$ is < 0, a contradiction. So, $\overline{x}$ must be local minimum for (66).

□

Thus (73) provides a **sufficient local minimality condition** for (66). See references [A8, A10, 10.2, 10.12, 10.13, 10.17, 10.27] for a complete discussion of optimality conditions for nonlinear programs.

In inequality constrained problems, we notice that the gap between known second order necessary optimality conditions and sufficient optimality conditions, is quite wide.

The NLP (66) is said to be a **convex programming problem** if $\theta(x)$ is convex, $h_i(x)$ is affine for all i, and $g_p(x)$ is concave for all p. In this case the set of feasible solutions is a convex set. For convex programming problems, we will now show that (71) are both **necessary and sufficient conditions for global optimality**.

Theorem 23. *Suppose (66) is a convex program. The feasible regular point $\overline{x}$ is a global minimum for (66) iff there exists a Lagrange multiplier vector $(\overline{\mu}, \overline{\pi})$ such that $\overline{x}$, $\overline{\mu}$, $\overline{\pi}$ together satisfy (71).*

Proof. The necessity of (71) for optimality has already been established above. We will now prove the sufficiency. Suppose $\overline{x}$ is a feasible solution of (66) satisfying (71).

Let x be any other feasible solution for (66). By Theorem 15

$$\begin{aligned}
\theta(x) - \theta(\overline{x}) &\geqq (\nabla\theta(\overline{x}))(x - \overline{x}) \\
&= \Big(\sum_{i=1}^{m} \overline{\mu}_i \nabla h_i(\overline{x}) + \sum_{p \in \mathbf{P}(\overline{x})} \overline{\pi}_p \nabla g_p(\overline{x})\Big)(x - \overline{x}) \text{ by (71)} \\
&= \sum_{p \in \mathbf{P}(\overline{x})} \overline{\pi}_p \nabla g_p(\overline{x})(x - \overline{x}), \text{ since } h(x) \text{ is affine} \\
&\geqq \sum_{p \in \mathbf{P}(\overline{x})} \overline{\pi}_p (g_p(x) - g_p(\overline{x})) \text{ by Theorem 16, since } g_p(x) \text{ is concave.} \\
&= \sum_{p \in \mathbf{P}(\overline{x})} \overline{\pi}_p g_p(x), \text{ since } g_p(\overline{x}) = 0 \text{ for } p \in \mathbf{P}(\overline{x}). \\
&\geqq 0, \text{ since } \overline{\pi} \geqq 0 \text{ and } g(x) \geqq 0 \text{ for feasibility.}
\end{aligned}$$

So $\overline{x}$ is a global minimum for (66).

□

Example 12

Consider the problem of determining the electrical current flows in the following electrical network.

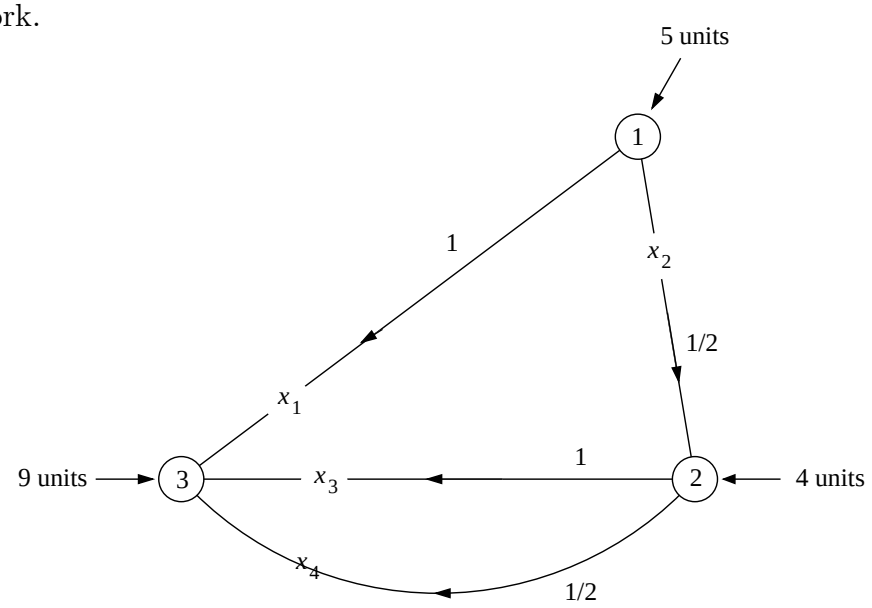


Figure 21

Assume that the current flows on each arc in the direction indicated. A total of 5, 4 units of current enters the system at nodes 1, 2 respectively per unit time. The numbers given on the arcs are the resistences of the arcs. Let x_1, x_2, x_3, x_4 denote the current flows on the arcs as indicated. If r_j denotes the resistence associated with x_j it is known that the power loss is $\sum_{j=1}^{4} r_j x_j^{\mathbf{2}}$. It is required to find out the current flows, under the assumption that the flows would occur so as to minimize the power

loss. Hence the x-vector is the optimum solution of the problem

$$\begin{array}{lllll} \text{minimize} & x_1^2 + (1/2)x_2^2 + x_3^2 + (1/2)x_4^2 & & & \\ \text{subject to} & x_1 & + x_2 & & = 5 \\ & & - x_2 + x_3 + & x_4 & = 4 \\ & & x_j \geqq 0, \ j = 1 \text{ to } 4. & & \end{array} \tag{78}$$

So, the Lagrangian is $L(x, \mu, \pi) = x_1^2 + (1/2)x_2^2 + x_3^2 + (1/2)x_4^2 - \mu_1(x_1 + x_2 - 5) - \mu_2(-x_2 + x_3 + x_4 - 4) - \sum_{j=1}^{4} \pi_j x_j$.

So, the first order necessary optimality conditions are

$$\begin{array}{lllll} \dfrac{\partial L}{\partial x_1} = & 2x_1 & & -\mu_1 & -\pi_1 = 0 \\ \dfrac{\partial L}{\partial x_2} = & & x_2 & -\mu_1 + \mu_2 & -\pi_2 = 0 \\ \dfrac{\partial L}{\partial x_3} = & & 2x_3 & -\mu_2 & -\pi_3 = 0 \\ \dfrac{\partial L}{\partial x_4} = & & & x_4 - \mu_2 & -\pi_4 = 0 \end{array} \tag{79}$$

$$\pi_1, \pi_2, \pi_3, \pi_4 \geqq 0 \tag{80}$$

$$\pi_1 x_1 = \pi_2 x_2 = \pi_3 x_3 = \pi_4 x_4 = 0 \tag{81}$$

and the constraints (78) on the x-variables for feasibility.

The complementary slackness conditions (81) imply that for each j, either the Lagrange multiplier π_j is zero, or the inequality constraint $x_j \geqq 0$ holds as an equality constraint (i. e., it is active) at the optimum. One technique to find a solution to the first order necessary optimality conditions here is to guess the subset of inequality constraints in (78) which will be active at the optimum, called the **active set**. Treat each of the inequality constraints in (78) in this active set as an equation, ignore the inequality constraints in (78) outside the active set (we are assuming that they will be inactive at the optimum). Set the Lagrange multiplier π_j corresponding to each inequality constraint in (78), not in the active set to zero. What remains among (78), (79) is a system of equations, which is solved. If the solution of this system satisfies (80) and the ignored inequality constraints in (78) not in the active set, we are done, this solution solves the first order necessary optimality conditions. If some of these conditions are violated, repeat this process with a different active set. This process, therefore, involves a combinatorial search, which may eventually involve solving 2^t systems where t is the number of inequality constraints in the original NLP ($t = 4$ here), not efficient if t is large. Efficient algorithms for solving NLP's involving inequality constraints either carry out this combinatorial search very efficiently; or do not use it at all, but operate with other efficient methods to find a solution to the first order necessary optimality conditions (see Chapters 2, 10).

We first try treating the inequality constraint $x_3 \geqq 0$ as active, and all the other inequality constraints $x_j \geqq 0$, $j = 1, 2, 4$ as inactive. Ignoring these inactive inequality constraints, and setting $\pi_j = 0$, $j = 1, 2, 4$ leads to the system of equations:

$$\begin{array}{rrrrrr} x_1 & +x_2 & & & & = 5 \\ & -x_2 & +x_4 & & & = 4 \\ 2x_1 & & & -\mu_1 & & = 0 \\ & x_2 & & -\mu_1 & -\mu_2 & = 0 \\ & & & & -\mu_2 & -\pi_3 = 0 \\ & & x_4 & & -\mu_2 & = 0. \end{array}$$

This system has the unique solution $(x_1, x_2, x_4) = (-2, 7, 11)$, $(\mu_1, \mu_2) = (-4, 11)$, $\pi_3 = -11$. This solution violates the constraints "$x_1 \geqq 0$, $\pi_3 \geqq 0$", so this choice of active set did not lead to a solution of the first order necessary optimality conditions in this problem.

Let us now try treating all the constraints "$x_j \geqq 0$, $j = 1$ to 4" as inactive. Ignoring all these inactive constraints, and setting $\pi_j = 0$, $j = 1$ to 4 leads to the system of equations

$$\begin{array}{rrrrrr} 2x_1 & & & & -\mu_1 & = 0 \\ & x_2 & & & -\mu_1 & +\mu_2 = 0 \\ & & 2x_3 & & & -\mu_2 = 0 \\ & & & x_4 & & -\mu_2 = 0 \\ x_1 & +x_2 & & & & = 5 \\ & -x_2 & +x_3 & +x_4 & & = 4. \end{array}$$

This system has the unique solution $\overline{x} = (3, 2, 2, 4)^T$, $\overline{\mu} = (6, 4)$. This solution also satisfies the inequality constraints, on the x_j which were ignored. So $(\overline{x}, \overline{\mu}, \overline{\pi} = 0)$ satisfies the first order necessary optimality conditions for this problem. It can be verified that $\overline{x}$ also satisfies the second order necessary optimality conditions, as well as the sufficient conditions for being a local minimum for this problem. Since $\theta(x)$ is convex here, $\overline{x}$ is in fact a global minimum for this problem.

Optimality Conditions for Linearly Constrained Optimization Problems

Consider the nonlinear program,

$$\begin{array}{ll} \text{minimize} & \theta(x) \\ \text{subject to} & Ax = b \\ & Dx \geqq d \end{array} \tag{82}$$

where A, b, D, d are given matrices of orders $m \times n$, $m \times 1$, $t \times n$ and $t \times 1$ respectively, and $\theta(x)$ is continuously differentiable. Since the constraints are linear, for this problem, we can establish first order necessary optimality conditions of the form in (71) without requiring a regularity type of constraint qualification.

Theorem 24. *If $\overline{x}$ is a local minimum for (82), there exist Lagrange multiplier vectors $\overline{\mu} = (\overline{\mu}_1, \ldots, \overline{\mu}_m)$, $\overline{\pi} = (\overline{\pi}_1, \ldots, \overline{\pi}_t)$ such that*

$$\begin{aligned} \nabla_x L(\overline{x}, \overline{\mu}, \overline{\pi}) = \nabla\theta(\overline{x}) - \overline{\mu}A - \overline{\pi}D = 0 \\ A\overline{x} = b,\ D\overline{x} \geqq d \\ \overline{\pi} \geqq 0,\ \overline{\pi}(D\overline{x} - d) = 0 \end{aligned} \tag{83}$$

where $L(x, \mu, \pi) = \theta(x) - \mu(Ax - b) - \pi(Dx - d)$ is the Lagrangian for (82).

Proof. Let $\mathbf{P}(\overline{x}) = \{p : 1 \leqq p \leqq t \text{ and } D_{p.}\overline{x} = d_p\}$, it is the index set of active inequality constraints in (82) at the feasible point $\overline{x}$. Since the constraints are linear, the tangent plane to the system determined by the active constraints in (82) at $\overline{x}$ is

$$\mathbf{T} = \{y : A_{i.}y = 0,\ i = 1 \text{ to } m, \text{ and } D_{p.}y = 0,\ p \in \mathbf{P}(\overline{x})\} \tag{84}$$

whether $\overline{x}$ satisfies the regularity condition for (82) or not. Let

$$\begin{aligned} \mathbf{\Gamma}_1 &= \{y : y \in \mathbf{R}^n,\ A_{i.}y = 0,\ i = 1 \text{ to } m \text{ and } D_{p.}y \geqq 0 \text{ for all } p \in \mathbf{P}(\overline{x})\} \\ \mathbf{\Gamma}_2 &= \{y : y \in \mathbf{R}^n,\ (\nabla\theta(\overline{x}))y < 0\}. \end{aligned}$$

We will now show that the fact that $\overline{x}$ is a local minimum for (82) implies that $\mathbf{\Gamma}_1 \cap \mathbf{\Gamma}_2 = \emptyset$. Suppose not. Let $\overline{y} \in \mathbf{\Gamma}_1 \cap \mathbf{\Gamma}_2$. Since $D_{p.}\overline{x} > d_p$ for $p \notin \mathbf{P}(\overline{x})$, and $\overline{y} \in \mathbf{\Gamma}_1$, it can be verified that $\overline{x} + \alpha\overline{y}$ is feasible to (82) when α is positive and sufficiently small, and since $\overline{y} \in \mathbf{\Gamma}_2$, we have $\theta(\overline{x} + \alpha\overline{y}) < \theta(\overline{x})$, contradicting the local minimality of $\overline{x}$ to (82). So $\mathbf{\Gamma}_1 \cap \mathbf{\Gamma}_2 = \emptyset$.

$\mathbf{\Gamma}_1 \cap \mathbf{\Gamma}_2 = \emptyset$ implies by Farkas' theorem (Theorem 3 of Appendix 1) that there exist $\overline{\mu} = (\overline{\mu}_1, \ldots, \overline{\mu}_m)$ and $\overline{\pi}_p$ for $p \in \mathbf{P}(\overline{x})$ satisfying

$$\begin{aligned} \nabla\theta(\overline{x}) = \sum_{i=1}^{m} \overline{\mu}_i A_{i.} + \sum_{p \in \mathbf{P}(\overline{x})} \overline{\pi}_p D_{p.} \\ \overline{\pi}_p \geqq 0 \text{ for all } p \in \mathbf{P}(\overline{x}). \end{aligned}$$

Now define $\overline{\pi}_p = 0$ for $p \notin \mathbf{P}(\overline{x})$, and let $\overline{\pi} = (\overline{\pi}_1, \ldots, \overline{\pi}_t)$. From the above, we verify that $\overline{x}$, $\overline{\mu}$, $\overline{\pi}$ together satisfy (83).

□

The conditions (83) are the **first order necessary optimality conditions** for the linearly constrained optimization problem (82).

If $\theta(x)$ is twice continuously differentiable in (82), since the constraints are linear in (82), it can be verified that the Hessian matrix of the Lagrangian is the same as the Hessian matrix of $\theta(x)$. Using the Taylor series approximation up to the second order, it can be shown that if $\overline{x}$ is a local minimum for (82), there must exist Lagrange multiplier vectors $\overline{\mu}$, $\overline{\pi}$ such that

$$(83) \text{ holds and } y^T H(\theta(\overline{x}))y \geqq 0 \text{ for all } y \in \mathbf{T} \text{ of (84).} \tag{85}$$

The conditions (85) correspond to (72), they are the **second order necessary optimality conditions** for (82).

5. Summary of Some Optimality Conditions

All the functions (objective and constraint function) are assumed to be continuously differentiable. They are assumed to be twice continuously differentiable, if the Hessian matrix, appears in the expressions.

Problem	necessary optimality conditions for point $\bar{x}$ to be a local minimum	sufficient optimality conditions for point $\bar{x}$ to be a local miminum
Unconstrained minimization. minimize $\theta(x)$ over $x \in \mathbf{R}^n$	<u>First order conditions</u> $\nabla\theta(\bar{x}) = 0$ <u>Second order conditions</u> $\nabla\theta(\bar{x}) = 0$ and $H(\theta(\bar{x}))$ is PSD.	 $\nabla\theta(\bar{x}) = 0$ and $H(\theta(\bar{x}))$ is PD.

Problem	necessary optimality conditions for point $\bar{x}$ to be a local minimum	sufficient optimality conditions for point $\bar{x}$ to be a local miminum
Equality constrained minimization. minimize $\theta(x)$ subject to $h_i(x) = 0$, $i = 1$ to m.	Denote $h(x)=(h_1(x), ..., h_m(x))^T$. $\mu=(\mu_1, ..., \mu_m)$ is the row vector of Lagrange multipliers. The Largrangian is $L(x,\mu)=\theta(x)-\mu h(x)=\theta(x)-\sum_{i=1}^{m} \mu_i h_i(x)$.	
	Conditions given here hold under the constraint qualifications that $\bar{x}$ is a regular point (i.e., $\{\nabla h_i(\bar{x})$: $i = 1$ to $m\}$ is a linearly independent set) or under weaker constraint qualifications, for which see references [A8, A10, A12, 10.12, 10.13, 10.26], or if the constraints are all linear.	Feasibility, $h(\bar{x})=0$ and there exists a Lagrange multiplier vector $\bar{\mu}$, which together with $\bar{x}$ satisfies $\nabla_x(L(\bar{x},\bar{\mu})) = 0$ and $y^T H_x(L(\bar{x},\bar{\mu}))y > 0$ for all $y \in \{y$: $(\nabla h_i(\bar{x}))y = 0$, $i = 1$ to m, $y \neq 0\}$.
	First order conditions Feasibility, $h(\bar{x})=0$ and there exists a Lagrange multiplier vector $\bar{\mu}$, which together with $\bar{x}$ satisfies $\nabla_x L(\bar{x},\bar{\mu}) = \nabla\theta(\bar{x})-\bar{\mu}\nabla h(\bar{x}) = 0$ (i.e., objective gradient is a linear combination of the constraint gradients).	
	Second order conditions $h(\bar{x})=0$ and there exists a Lagrange multiplier vector $\bar{\mu}$, which together with $\bar{x}$ satisfies $\nabla_x(L(\bar{x},\bar{\mu})) = 0$ and $y^T H_x(L(\bar{x},\bar{\mu}))y \geqq 0$ for all $y \in \{y$: $(\nabla_x h_i(\bar{x}))y = 0$, $i=1$ to $m\}$.	

Problem	necessary optimality conditions for point $\bar{x}$ to be a local minimum	sufficient optimality conditions for point $\bar{x}$ to be a local miminum
Constrained minimization. minimize $\theta(x)$ subject to $h_i(x)=0$, $i=1$ to m; $g_p(x) \geqq 0$, $p=1$ to t.	Let $h(x)=(h_1(x), ..., h_m(x))^T$. $g(x) = (g_1(x), ..., g_t(x))^T$. Let $\mu=(\mu_1, ..., \mu_m)$, $\pi=(\pi_1, ..., \pi_t)$ be row vectors of Lagrange multipliers. The Largrangian is $L(x,\mu,\pi)= \theta(x)-\mu h(x)-\pi g(x)=\theta(x) - \sum_{i=1}^{m} \mu_i h_i(x)-\sum_{p=1}^{t} \pi_p g_p(x)$.	
	Conditions given here hold under the constraint qualification that $\bar{x}$ is a regular point (i.e., if $P(\bar{x})=\{p: g_p(\bar{x})=0\}$, then $\{\nabla h_i(\bar{x}): i=1 \text{ to } m\} \cup \{\nabla g_p(\bar{x}): p \in P(\bar{x})\}$ is a linearly independent set), or under weaker constraint qualifications, for which see references cited above, or if the constraints are all linear.	Feasibility, $h(\bar{x}) = 0$, $g(\bar{x}) \geqq 0$ and there exists Lagrange multiplier vectors $\bar{\mu}$, $\bar{\pi}$ which together with $\bar{x}$ satisfy $\nabla_x L(\bar{x},\bar{\mu},\bar{\pi}) = 0$, $\bar{\pi} \geqq 0$, $\bar{\pi} g(\bar{x}) = 0$
	<u>First order conditions (KKT conditions)</u> Feasibility, $h(\bar{x})=0$, $g(\bar{x}) \geqq 0$ and there exists Lagrange multiplier vectors $\bar{\mu}$, $\bar{\pi}$ which together with $\bar{x}$ satisfy $\nabla_x L(\bar{x},\bar{\mu},\bar{\pi})=\nabla\theta(\bar{x})-\sum_{i=1}^{m} \bar{\mu}_i \nabla h_i(\bar{x}) - \sum_{p=1}^{t} \bar{\pi}_p \nabla g_p(\bar{x}) = 0$, $\bar{\pi} \geqq 0$, $\bar{\pi}_p g_p(\bar{x}) = 0$ for all p. The last set of conditions here are the **complementary slackness conditions.** It is equivalent to $\bar{\pi} g(\bar{x})=0$.	and $y^T H_x(L(\bar{x},\bar{\mu},\bar{\pi}))y>0$ for all $y \in T_1=\{y: (\nabla h_i(\bar{x}))y=0, i=1 \text{ to } m, \text{ and } (\nabla g_p(\bar{x}))y=0 \text{ for } p \in P(\bar{x}) \cap \{p: \bar{\pi}_p>0\}$ and $(\nabla g_p(\bar{x}))y \geqq 0$, for $p \in P(\bar{x}) \cap \{p: \bar{\pi}_p=0\}$, $y \neq 0$.
	<u>Second order conditions</u> $h(\bar{x})=0$, $g(\bar{x}) \geqq 0$ and there exists Lagrange multiplier vectors $\bar{\mu}$, $\bar{\pi}$, which together with $\bar{x}$ satisfy $\nabla_x L(\bar{x},\bar{\mu},\bar{\pi}) = 0$, $\bar{\pi} \geqq 0$, $\bar{\pi} g(\bar{x}) = 0$ and $y^T H_x(L(\bar{x},\bar{\mu},\bar{\pi}))y \geqq 0$ for all $y \in \{y: (\nabla h_i(\bar{x}))y=0, i=1 \text{ to } m \text{ and } (\nabla g_p(\bar{x}))y=0, p \in P(\bar{x})\}$.	For convex programming problem, that is, when $\theta(x)$ is convex, $h(x)$ is affine and $g(x)$ is concave, the First order conditions (KKT conditions) are both necessary and sufficient for $\bar{x}$ to be a global minimum.

6. Exercises

---

13. Consider the quadratic program

$$\begin{array}{ll} \text{minimize} & cx + (1/2)x^T Dx \\ \text{subject to} & Ax \geqq b \\ & x \geqq 0 \end{array}$$

where D is a general symmetric matrix of order n. Prove that the necessary and sufficient conditions for x^* to be a local minimum to this general quadratic program is that there exist vectors y^*, u^*, v^*, such that

$$\begin{pmatrix} u^* \\ v^* \end{pmatrix} = \begin{pmatrix} D & -A^T \\ A & 0 \end{pmatrix} \begin{pmatrix} x^* \\ y^* \end{pmatrix} + \begin{pmatrix} c^T \\ -b \end{pmatrix}$$

$$\begin{pmatrix} u^* \\ v^* \end{pmatrix} \geqq 0, \ \begin{pmatrix} x^* \\ y^* \end{pmatrix} \geqq 0, \ \begin{pmatrix} u^* \\ v^* \end{pmatrix}^T \begin{pmatrix} x^* \\ y^* \end{pmatrix} = 0$$

hold, and for every vector $\xi \in \mathbf{R}^n$ satisfying

$$\begin{aligned} A_i.\xi &= 0 \text{ if } y_i^* > 0 \\ A_i.\xi &\geqq 0 \text{ if } v_i^* = y_i^* = 0 \\ \xi_j &= 0 \text{ if } u_j^* > 0 \\ \xi_j &\geqq 0 \text{ if } x_j^* = u_j^* = 0 \end{aligned}$$

we have $\xi^T D\xi \geqq 0$. (A. Majthay [A9])

14. Consider the quadratic programming problem

$$\begin{array}{ll} \text{minimize} & cx + (1/2)x^T Dx \\ \text{subject to} & 0 \leqq x \leqq u \end{array}$$

where

$$D = \begin{pmatrix} -2 & -3 & -3 \\ -3 & -5 & -1 \\ -3 & -1 & -4 \end{pmatrix}, \ c = \begin{pmatrix} 4 \\ 3 \\ 5 \end{pmatrix}^T, \ u = \begin{pmatrix} 10 \\ 10 \\ 10 \end{pmatrix}$$

and identify the global optimum solution of this problem. (W. P. Hallman and I. Kaneko [2.15])

15. Let $f(x)$ be a real valued differentiable function defined on $\mathbf{R}^1$. Let $x^0 \in \mathbf{R}^1$. Is the following statement true? "For x^0 to be a local minimum for $f(x)$ in $\mathbf{R}^1$, it is

necessary that the derivative $f'(x^0) = 0$; and there must exist an open interval (a, b) around x^0 such that $f'(x) < 0$ for all x in the open interval (a, x^0), and $f'(x) > 0$ for all x in the open interval (x^0, b)". Is this condition sufficient for x^0 to be a local minimum of $f(x)$? Use the function defined by

$$f(x) = x^{\mathbf{2}}(2 + \sin(1/x)), \text{ when } x \neq 0$$
$$f(0) = 0$$

and $x^0 = 0$, as an example. (K. Sydsaeter [A14])

16. Let $f(x)$ be a real valued function defined on $\mathbf{R}^1$. Let $x^0 \in \mathbf{R}^1$, and suppose $f(x)$ has continuous nth derivative. A sufficient condition for x^0 to be a strict local minimum for $f(x)$ in $\mathbf{R}^1$, is that $f^{(1)}(x^0) = f^{(2)}(x^0) = \ldots = f^{(n-1)}(x^0) = 0$, and $f^{(n)}(x^0) > 0$ for n even, where, $f^{(r)}(x^0)$ is the rth derivative of $f(x)$ at x^0. Is this condition necessary for x^0 to be a local minimum for $f(x)$? Use the function defined by

$$f(x) = e^{-(\mathbf{1}/\mathbf{x}^{\mathbf{2}})}, x \neq 0$$
$$f(0) = 0$$

and $x^0 = 0$, as an example. (K. Sydsaeter [A14])

17. It is sometimes stated that minimizing a function subject to constraints is equivalent to finding the unconstrained minimum of the Lagrangian function. Examine whether this statement is true, using the example

$$\begin{array}{ll} \text{minimize} & -x_1 x_2 \\ \text{subject to} & x_1 + x_2 = 2 \end{array}$$

and the point $\overline{x} = (1, 1)^T$ which is optimal for it. (K. Sydsaeter [A14])

18. Consider the equality constrained optimization problem (56) and the Lagrangian $L(x, \mu) = \theta(x) - \mu h(x)$ for it. If $(\overline{x}, \overline{\mu})$ is an unconstrained local minimum for $L(x, \mu)$ over $x \in \mathbf{R}^n$, $\mu \in \mathbf{R}^m$, prove that the point $\overline{x}$ must be feasible to (56) and in fact it must be a local minimum for (56). However, show that the converse may not be true, that is, even if $\hat{x}$ is a local minimum for (56), there may not exist a Lagrange multiplier vector $\hat{\mu}$ such that $(\hat{x}, \hat{\mu})$ is an unconstrained local minimum for $L(x, \mu)$. See Exercise 17 above. Develop general conditions on the NLP (56) and the point $\overline{x}$ which can guarantee that if $\overline{x}$ is a local minimum for (56), there exists a Lagrange multiplier vector $\overline{\mu}$ such that $(\overline{x}, \overline{\mu})$ is a local minimum for $L(x, \mu)$ over $x \in \mathbf{R}^n$, $\mu \in \mathbf{R}^m$.

19. Consider the NLP (66) and the Lagrangian $L(x, \mu, \pi) = \theta(x) - \mu h(x) - \pi g(x)$. If $(\overline{x}, \overline{\mu}, \overline{\pi})$ is a local minimum for the problem

$$\begin{array}{ll} \text{minimize} & L(x, \mu, \pi) \\ \text{subject to} & x \in \mathbf{R}^n, \mu \in \mathbf{R}^m \\ \text{and} & \pi \geqq 0 \end{array} \tag{86}$$

prove that $\overline{x}$ must be feasible to (66) and in fact must be a local minimum for (66). However, show that the converse may not be true, that is even if $\hat{x}$ is a local minimum for (66), there may not exist a $\hat{\mu} \in \mathbf{R}^m$ and $\hat{\pi} \in \mathbf{R}^t$, $\hat{\pi} \geqq 0$, such that $(\hat{x}, \hat{\mu}, \hat{\pi})$ is a local minimum for (86).

Develop general conditions on the NLP (66) and the point $\overline{x}$, which can guarantee that if $\overline{x}$ is a local minimum for (66), there exist Lagrange multiplier vectors $\overline{\mu}$, $\overline{\pi}$ such that $(\overline{x}, \overline{\mu}, \overline{\pi})$ is a local minimum for (86).

20. Let $\theta(x)$ be a real valued function defined on $\mathbf{R}^n$ and let $\overline{x} \in \mathbf{R}^n$. Examine the following statement "If $\overline{x}$ is a local minimum along each straight line through $\overline{x}$ in $\mathbf{R}^n$, then $\overline{x}$ is a local minimum for $\theta(x)$ in $\mathbf{R}^n$", and mark whether it is true or false. Use $\theta(x_1, x_2) = (x_2 - x_1^{\mathbf{2}})(x_2 - 2x_1^{\mathbf{2}})$ defined on $\mathbf{R}^2$ and $\overline{x} = (0, 0)^T$ as an example. (K. Sydsaeter [A14])

21. Let A, D be given PD matrices of order n. Solve the following two optimization problems.

$$\begin{array}{lll} \text{(i)} & \text{minimize} & cx \\ & \text{subject to} & (x - \overline{x})^T A(x - \overline{x}) \leqq 1 \\ \text{(ii)} & \text{minimize} & cx + (1/2)x^T Dx \\ & \text{subject to} & (x - \overline{x})^T A(x - \overline{x}) \leqq 1. \end{array}$$

Discuss what happens if A is PD but D is either PSD or not even PSD.

22. Consider the following quadratic programming problem

$$\begin{aligned} f(b) = {} & \text{minimum value of } Q(x) = cx + (1/2)x^T Dx \\ & \text{subject to } Ax \geqq b \\ & x \geqq 0 \end{aligned}$$

where D is a symmetric PSD matrix of order n, $f(b)$ denotes the optimum objective value in this problem as a function of the vector b, and A, b are given matrices of orders $m \times n$ and $m \times 1$ respectively. In this problem, assume that A, c, d remain fixed, but b may vary.

(i) If $f(b)$ is finite for some b, prove that $f(b)$ is finite for all b for which the problem is feasible.

(ii) If $f(b)$ is finite for some b, prove that $f(b)$ is convex over $b \in \text{Pos}(A, -I_m)$.

(iii) What is $\partial f(b)$?

Note: The result in (i) above could be false if D is not PSD. Consider the following

example from B. C. Eaves [2.9]

$$\begin{array}{lrcl} \text{minimize } Q(x) = & -4x_1 + x_1^{\mathbf{2}} - x_2^{\mathbf{2}} & & \\ \text{subject to} & -x_1 + x_2 & \geqq & b_1 \\ & -x_1 + x_2 & \geqq & b_2 \\ & x_1, x_2 & \geqq & 0. \end{array}$$

Let $b = (b_1, b_2)^T$. If $b = b^1 = (-2, -4)$, or if $b = b^2 = (-4, -2)$, verify that the problem is feasible and that $Q(x)$ is bounded below on the set of feasible solution. If $b = (b^1 + b^2)/2 = (-3, -3)^T$, verify that $Q(x)$ becomes unbounded below on the set of feasible solutions.

23. Let $\mathbf{K} \subset \mathbf{R}^n$ be a closed convex set. For $x \in \mathbf{R}^n$, define

$$f(x) = \text{ Minimum } \{\|y - x\| : y \in \mathbf{K}\}.$$

Prove that $f(x)$ is convex.

24. Let $\theta(x) = (2x_2 - x_1^{\mathbf{2}})^{\mathbf{2}}$. Check whether $\theta(x)$ is convex, or concave, or neither, on $-1 \leqq x_1 \leqq 1$, $-1 \leqq x_2 \leqq 1$.

25. Consider the linear program in standard form

$$\begin{array}{lrcl} \text{minimize} & cx & & \\ \text{subject to} & Ax & = & b \\ & x & \geqq & 0. \end{array}$$

This problem can be written as the following NLP in which the constraints are all equalities, but there are new variables u_j.

$$\begin{array}{ll} \text{minimize} & cx \\ \text{subject to} & Ax = b \\ & u_j^{\mathbf{2}} - x_j = 0, \text{ for all } j. \end{array}$$

Write down the necessary optimality conditions for this equality constrained NLP, and show that they are equivalent to the duality-complementary slackness conditions for optimality in the above LP.

26. Consider the NLP

$$\begin{array}{ll} \text{minimize} & \theta(x) \\ \text{subject to} & h_i(x) = 0, \ i = 1 \text{ to } m \\ & g_p(x) \geqq 0, \ p = 1 \text{ to } t \end{array}$$

where all the functions $\theta(x)$, $h_i(x)$, $g_p(x)$ are continuously differentiable. If $\overline{x}$ is a local minimum for this problem, and

(a) $\{\nabla h_i(\overline{x}) : i = 1 \text{ to } m\}$ is linearly independent,

(b) there exists a $y \in \mathbf{R}^n$ satisfying

$$\nabla h_i(\overline{x})y = 0,\ i = 1 \text{ to } m$$
$$\nabla g_p(\overline{x})y > 0,\ p \in \mathbf{P}(\overline{x})$$

where $\mathbf{P}(\overline{x}) = \{p : g_p(\overline{x}) = 0\}$.

Prove that there must exist $\overline{\mu} = (\overline{\mu}_1, \ldots, \overline{\mu}_m)$, $\overline{\pi} = (\overline{\pi}_1, \ldots, \overline{\pi}_t)$ such that

$$\nabla\theta(\overline{x}) - \overline{\mu}\nabla h(\overline{x}) - \overline{\pi}\nabla g(\overline{x}) = 0$$
$$\overline{\pi} \geqq 0 \text{ and } \overline{\pi}_p g_p(\overline{x}) = 0 \text{ for all } p = 1 \text{ to } t.$$

27. Consider the NLP

$$\begin{array}{ll} \text{minimize} & \theta(x) = x_1^{\mathbf{2}} + x_2^{\mathbf{2}} \\ \text{subject to} & (x_1 - 1)^{\mathbf{3}} - x_2^{\mathbf{2}} = 0. \end{array}$$

i) If $\overline{x} = (\overline{x}_1, \overline{x}_2)^T$ is a feasible solution to this problem, prove that $\overline{x}_1$ must be $\geqq 1$. Using this information, prove that $\hat{x} = (1, 0)^T$ is the global minimum for this problem.

ii) Write down the first order necessary optimality conditions for this problem. Does $\hat{x}$ satisfy these conditions? Why? Explain clearly. (R. Fletcher [10.13])

28. Consider the NLP

$$\begin{array}{lll} \text{minimize} & \theta(x) = x_2 & \\ \text{subject to} & (1 - x_1)^{\mathbf{3}} - x_2 & \geqq 0 \\ & x_1 & \geqq 0 \\ & x_2 & \geqq 0. \end{array}$$

Verify that $\overline{x} = (1, 0)^T$ is a global optimum solution to this problem. Is $\overline{x}$ a regular point? Do the first order necessary optimality conditions hold at $\overline{x}$?

If the problem is to minimize: $-x_1$, subject to the constraints given above, verify that $\overline{x}$ is again the global optimum. Do the first order necessary optimality conditions hold at $\overline{x}$ for this problem? Why?

29. In each of the following NLPs, find out the global optimum and check whether the first order necessary optimality conditions hold at it. Explain the reasons for it.

(87)

$$\begin{array}{ll} \text{minimize} & -x_1 \\ \text{subject to} & -x_1^{\mathbf{2}} \geqq 0 \\ & x_1 \geqq 0 \end{array}$$

(88)

$$\begin{array}{ll} \text{minimize} & -x_1 \\ \text{subject to} & -x_1^{\mathbf{2}} + x_2 \geqq 0 \\ & -x_2 \geqq 0. \end{array}$$

30. Find an optimum solution to the following NLP, using a combinatorial search for the set of active constraints at the optimum

$$\begin{array}{ll} \text{minimize} & \theta(x) = 2x_1^{\mathbf{2}} + 2x_1x_2 + x_2^{\mathbf{2}} - 10x_1 - 10x_2 \\ \text{subject to} & x_1^{\mathbf{2}} + x_2^{\mathbf{2}} \leqq 5 \\ & 3x_1 + x_2 \leqq 6. \end{array}$$

31. Consider the following NLP

$$\begin{array}{ll} \text{minimize} & -x_1^{\mathbf{4}} - x_2^{\mathbf{4}} \\ \text{subject to} & x_1^2 + (x_2 - 1)^2 - 1 = 0. \end{array}$$

Verify that $\overline{x} = (0, 2)^T$ is a global minimum for this problem. Do the first order necessary optimality conditions hold at $\overline{x}$? Is there a $\overline{\mu}$ such that $(\overline{x}, \overline{\mu})$ is a local minimum for the Lagrangian in this problem?

32. Consider the general NLP

$$\begin{array}{ll} \text{minimize} & \theta(x) \\ \text{subject to} & h_i(x) = 0, \ i = 1 \text{ to } m \\ & g_p(x) \geqq 0, \ p = 1 \text{ to } t \end{array}$$

where $\theta(x)$, $h_i(x)$, $g_p(x)$ are all continuously differentiable functions defined on $\mathbf{R}^n$.

One of the hard unsolved problems in NLP is to develop a computationally viable method or characterization to determine whether $\theta(x)$ is bounded below on the feasible solution set for this problem, or diverges $-\infty$ on this set; and when $\theta(x)$ is bounded on the solution set, to determine whether $\theta(x)$ attains its minimum at some finite feasible solution ($\theta(x)$ may only have an infimum in this problem, it may not be an attained minimum).

Another hard problem is to develop optimality conditions for a feasible solution $\overline{x}$ of this problem to be a global minimum for it. In the absence of convexity of $\theta(x)$, concavity of $g(x)$ and affineness of $h(x)$, at present we do not have any conditions for distinguishing the global minimum for this problem, from other local minima that may exist (the only known condition for the global minimum is its definition, that is, $\overline{x}$ is a global minimum iff $\theta(x) \geqq \theta(\overline{x})$ for all feasible solutions x, this condition is not computationally useful, since checking it directly may require computing the function value at uncountably many points).

33. Let A be a given matrix of order $m \times n$. Prove that the following three conditions are equivalent

(i) there exists no $x \in \mathbf{R}^n$ satisfying $Ax \leqq 0$, $x \geq 0$,

(ii) for every $b \in \mathbf{R}^m$, the set $\{x : Ax \leqq b, x \geqq 0\}$ is bounded,

(iii) there exists a $\pi \geqq 0$ satisfying $\pi A > 0$.

34. If $\theta(x)$ is a continuous real valued function defined over $\mathbf{R}^n$ with the monotonicity property (that is for every $0 \leqq x \leqq y$ we have $\theta(x) \leqq \theta(y)$), then prove that the problem

$$\begin{array}{ll} \text{minimize} & \theta(x) \\ \text{subject to} & Ax \geqq b \\ & x \geqq 0 \end{array}$$

has an optimum solution, if it is feasible. (B. C. Eaves [2.8])

35. Let $Q(x) = cx + (\frac{1}{2})x^T Dx$ where D is a symmetric matrix. Let $\alpha > 0$ be given. Prove that the point x^* solves the problem

$$\begin{array}{ll} \text{minimize} & Q(x) \\ \text{subject to} & \|x\| \leqq \alpha \end{array}$$

iff it is feasible and there exists a $\lambda \geqq 0$ satisfying

$$\begin{aligned} &(x^*)^T(D + \lambda I) = -c \\ &\lambda(\alpha - \|x^*\|) = 0 \\ &(D + \lambda I) \text{ is a PSD matrix.} \end{aligned}$$

36. Consider the following NLPs in each of which the variables are $x \in \mathbf{R}^n$.

(89)		(90)	
minimize	cx	minimize	$x^T x$
subject to	$x^T x \leqq 1$	subject to	$-cx \geqq 1$
	$Ax \geqq 0$		$Ax \geqq 0.$

The data in both the problems, the matrices A, c of order $m \times n$ and $1 \times n$ respectively, are the same. Prove that these two problems are equivalent.

37. Let $f(\lambda) : \mathbf{R}^1 \to \mathbf{R}^1$ be a real valued convex function defined on $\mathbf{R}^1$. For any λ, the limit of $\frac{f(\lambda+\varepsilon)-f(\lambda)}{\varepsilon}$ as $\varepsilon \to 0$ through positive values is called the right derivative of $f(\lambda)$ at λ, and denoted by $f'_+(\lambda)$, the limit of the same as $\varepsilon \to 0$ through negative values is called the left derivative of $f(\lambda)$ at λ and denoted by $f'_-(\lambda)$. Prove the following

i) If $\lambda < \gamma$, then $f'_-(\lambda) \leqq f'_+(\lambda) \leqq f'_-(\gamma) \leqq f'_+(\gamma)$.
ii) A necessary and sufficient condition for λ_* to minimize $f(\lambda)$ over $\lambda \in \mathbf{R}^1$ is: $f'_-(\lambda_*) \leqq 0 \leqq f'_+(\lambda_*)$.
iii) The subdifferential $\partial f(\lambda)$ is the line segment $[f'_-(\lambda), f'_+(\lambda)]$.
iv) For each λ, let $g(\lambda) \in \partial f(\lambda)$. Prove that

(a) $P(\lambda, \gamma) = f(\lambda) - [f(\gamma) + g(\gamma)(\lambda - \gamma)] \geqq 0$ for all λ, γ.
(b) If $f(\lambda) \leqq f(\gamma)$, then $P(\lambda, \gamma) \leqq |g(\gamma)| \cdot |\gamma - \lambda|$.
(c) If $g(\lambda)g(\gamma) < 0$, then

$$P(\lambda, \gamma) \geqq |g(\gamma)| \cdot |\lambda - \lambda_*| \text{ where } \lambda_* \text{ is the minimizer of } f(\lambda).$$

(C. Lemarechal and R. Mifflin [10.23])

38. Consider the problem

$$\begin{array}{ll} \text{minimize} & cx \\ \text{subject to} & Ax \leqq 0 \\ & \|x\| = 1 \end{array}$$

where A is of order $m \times n$.

Let $\mathbf{K} = \{y : y = \sum_{i=1}^{m} \mu_i A_{i.}, \mu_i \geqq 0 \text{ for all } i\}$. Prove the following

i) If $c \in \mathbf{K}$, the maximum objective value in this problem, is $\leqq 0$.
ii) If $c \notin \mathbf{K}$, let $b \in \mathbf{K}$ be the point in $\mathbf{K}$ that is closest to c. Then $(c - b)/\|b - c\|$ is the optimum solution of this problem, and the optimum objective value in the problem is $\|c - b\|$.

39. Let $\mathbf{K} \subset \mathbf{R}^n$ be a closed convex polyhedral set partitioned into closed convex polyhedral regions as $\bigcup_{t=1 \text{to } r} \mathbf{K}_t$. So if $u \neq v$, the interiors of $\mathbf{K}_u$ and $\mathbf{K}_v$ have an empty intersection, and $\mathbf{K}_u \cap \mathbf{K}_v$ is itself either empty or is either a face of lower dimension or a subset of a face of lower dimension of each of $\mathbf{K}_u$ and $\mathbf{K}_v$. Assume that each $\mathbf{K}_t$ has a nonempty interior. Suppose the real-valued function $f(x)$ is defined on $\mathbf{K}$ by the following

$$f(x) = f_t(x) = c_0^t + \sum_{j=1}^{n} c_j^t x_j, \text{ if } x \in \mathbf{K}_t$$

where c_0^t and c_j^t are all given constants. The definition assumes that if $\mathbf{K}_u \cap \mathbf{K}_v \neq \emptyset$, then $f_u(x) = f_v(x)$ for all $x \in \mathbf{K}_u \cap \mathbf{K}_v$. So $f(x)$ is a continuous piecewise linear function defined on $\mathbf{K}$.

Derive necessary and sufficient conditions for the continuous piecewise linear function $f(x)$ to be convex on $\mathbf{K}$, and develop an efficient algorithm to check whether these conditions hold.

As a numerical example, let $\mathbf{K} = \{x = (x_1, x_2)^T : -1 \leqq x_1 \leqq 1, -1 \leqq x_2 \leqq 1\}$. Consider the partition of $\mathbf{K}$ given in Figure 22. Two piecewise linear functions $f(x)$, $g(x)$ defined on $\mathbf{K}$ are provided in Figure 22. Check whether they are convex on $\mathbf{K}$. (See Section 8.14 in K. G. Murty [2.26].)

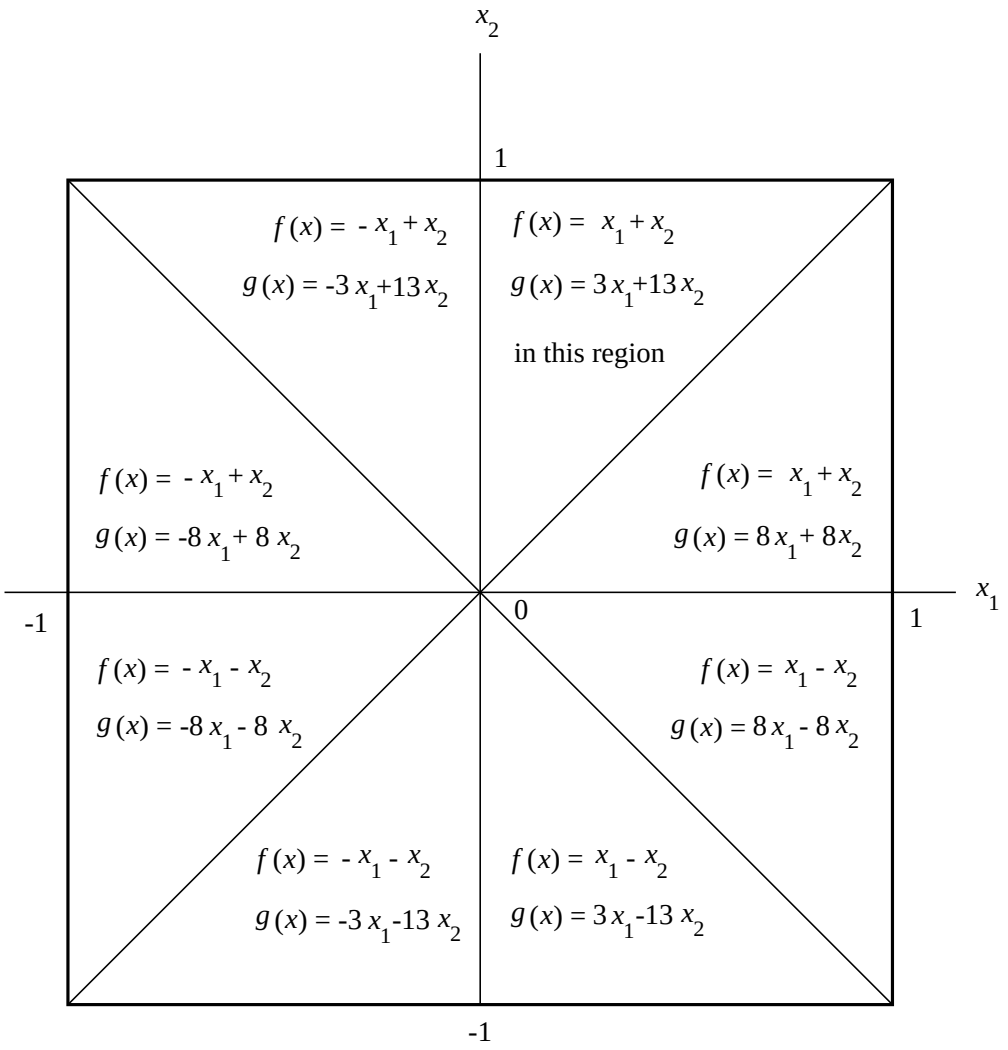


Figure 22

40. (Research Problem) For $i = 1$ to m, $g_i(x)$ is a real valued continuously differentiable function defined on $\mathbf{R}^n$, but $g(x) = (g_1(x), \ldots, g_m(x))^T$ is not convex. Let $\alpha = (\alpha_1, \ldots, \alpha_m)^T \in \mathbf{R}^m$. Let $\mathbf{K}(\alpha) = \{x : g_i(x) \leqq \alpha_i,\ i = 1 \text{ to } m\}$. Let $b = (b_1, \ldots, b_n) \in \mathbf{R}^n$ be a given point.

(i) Assuming that $\mathbf{K}(0) \neq \emptyset$ is a convex set, develop an efficient algorithm to check whether $\mathbf{K}(\alpha)$ is convex for given α. Is this problem easier to solve if either $\alpha \geqq 0$ or $\alpha \leqq 0$?

(ii) Assuming that $\mathbf{K}(0) \neq \emptyset$ is a convex set, and that b is a boundary point of $\mathbf{K}(0)$, (that is, there exists an i such that $g_i(b) = 0$), develop an efficient algorithm to find a $c = (c_1, \ldots, c_n) \neq 0$ satisfying $c(x - b) \geqq 0$ for all $x \in \mathbf{K}(0)$ (then $\mathbf{H} = \{x : c(x - b) = 0\}$ is a supporting hyperplane for the convex set $\mathbf{K}(0)$ at its boundary point b).

(iii) Assuming that $\mathbf{K}(0) \neq \emptyset$ is a convex set and that $b \notin \mathbf{K}(0)$, develop an efficient algorithm to determine a hyperplane separating b from $\mathbf{K}(0)$.

(iv) Consider the special cases of the above problems when all $g_i(x)$ are affine functions, excepting one which is quadratic and nonconvex.

41. Let $\theta(x)$ be a continuously differentiable real valued function defined on $\mathbf{R}^n$. Let $\mathbf{K}$ be a subspace of $\mathbf{R}^n$. If $\overline{x} \in \mathbf{K}$ minimizes $\theta(x)$ over $\mathbf{K}$, prove that $\nabla\theta(\overline{x})$ is orthogonal to every vector in $\mathbf{K}$.

42. Let $\theta(x)$; $g_i(x)$, $i = 1$ to m, be continuously differentiable convex functions defined on $\mathbf{R}^n$. Let $\overline{\theta}$ be the optimum objective value; and $\overline{\pi}$, an optimum Lagrange multiplier vector associated with the NLP

$$\begin{array}{ll} \text{minimize} & \theta(x) \\ \text{subject to} & -g_i(x) \geqq 0,\ i = 1 \text{ to } m. \end{array}$$

Then prove that $\overline{\theta} = \text{Infimum } \{\theta(x) + \overline{\pi}g(x) : x \in \mathbf{R}^n\}$.

43. Arithmetic Mean — Geometric Mean Inequality:

Let $x_1, \ldots, x_n$ be positive real numbers. Let $\delta_1, \ldots, \delta_n$ be positive real numbers satisfying $\delta_1 + \ldots + \delta_n = 1$. Prove that

$$\prod_{i=1}^{n} (x_i)^{\delta_i} \leqq \sum_{i=1}^{n} \delta_i x_i$$

with equality holding iff $x_1 = x_2 = \ldots = x_n$, where "$\prod$" indicates the product sign.

44. Young's Inequality:

Let x, y, p, q be all positive real numbers, and $p > 1$, $q > 1$ satisfying $\frac{1}{p} + \frac{1}{q} = 1$. Prove that

$$xy \leqq \frac{x^{\mathbf{p}}}{p} + \frac{y^{\mathbf{q}}}{q}$$

with equality holding only when $x^{\mathbf{p}} = y^{\mathbf{q}}$.

45. Holder's Inequality:

Let p, q, be positive real numbers > 1 satisfying $\frac{1}{p} + \frac{1}{q} = 1$. Let $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n)$ be real vectors. Prove that

$$\sum_{i=1}^{n} x_i y_i \leqq \left(\sum_{i=1}^{n} |x_i|^{\mathbf{p}} \right)^{\mathbf{1/p}} \left(\sum_{i=1}^{n} |y_i|^{\mathbf{q}} \right)^{\mathbf{1/q}}.$$

46. Minkowski's Inequality:

Let $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n)$ be real vectors and $p \geqq 1$. Prove that

$$\left(\sum_{i=1}^{n} |x_i + y_i|^p \right)^{\mathbf{1/p}} \leqq \left(\sum_{i=1}^{n} |x_i|^{\mathbf{p}} \right)^{\mathbf{1/p}} + \left(\sum_{i=1}^{n} |y_i|^{\mathbf{p}} \right)^{\mathbf{1/p}}.$$

47. Let $f(\lambda)$ denote a smooth real valued function defined over $\mathbf{R}^1$. A classical sufficient condition for $\overline{\lambda} \in \mathbf{R}^1$ to be a local minimum for $f(\lambda)$ over $\mathbf{R}^1$ states "$\overline{\lambda}$ is a local minimum for $f(\lambda)$ over $\mathbf{R}^1$ if the first nonzero derivative of $f(\lambda)$ at $\overline{\lambda}$ is of even order, and this derivative is > 0". Develop a generalization of this result to $\mathbf{R}^n$, $n > 1$.

48. Given a vector $y = (y_j) \in \mathbf{R}^n$, define

$$\begin{aligned}
\|y\|_1 &= \sum_{j=1}^{n} |y_j| \\
\|y\|_\infty &= \text{ maximum } \{|y_j| : j = 1 \text{ to } n\} \\
\|y\|_2 &= \sqrt{\sum_{j=1}^{n} y_j^{\mathbf{2}}} \\
y^+ &= (y_j^+) \text{ where } y_j^+ = \text{ maximum } \{0, y_j\}
\end{aligned}$$

$\|y\|_1$, $\|y\|_\infty$, $\|y\|_2$ are called the 1-norm, ∞-norm, 2-norm, respectively, of the vector y.

Consider the system

$$\begin{aligned}
Ax &\leqq b \\
Bx &= d
\end{aligned} \tag{91}$$

where A, B are fixed matrices of orders $m \times n$, $p \times n$ respectively; and b, d are column vectors of appropriate dimensions. Assume that each row vector of A contains at least one nonzero entry, and if equality constraints do exist, then B is of full row rank (B

could be vacuous, that is, there may be no equality constraints in (91)). Let $\mathbf{K}(b,d)$ denote the set of feasible solutions of (91). Define

$$\begin{aligned} \mu(A,B) = &\text{ supremum } \|u,v\|_2 \\ &\text{subject to, } u,\ v \text{ are row vectors in } \mathbf{R}^m,\ \mathbf{R}^p, \\ \|uA+vB\|_1 = &\ 1 \\ u \geqq &\ 0 \\ &\text{and the set of rows of } \begin{pmatrix} A \\ B \end{pmatrix} \text{ corresponding to} \\ &\text{nonzero elements of } (u,v) \text{ is linearly independent.} \end{aligned} \tag{92}$$

(i) Prove that $\mu(A,B)$ is finite.

(ii) If $\begin{pmatrix} b^1 \\ d^1 \end{pmatrix}$, $\begin{pmatrix} b^2 \\ d^2 \end{pmatrix}$ are such, that $\mathbf{K}(b^1,d^1)$ and $\mathbf{K}(b^2,d^2)$ are both nonempty; for each $x^1 \in \mathbf{K}(b^1,d^1)$, prove that there exists an $x^2 \in \mathbf{K}(b^2,d^2)$ satisfying

$$\|x^1 - x^2\|_\infty \leqq \mu(A,B) \left\| \begin{pmatrix} b^1 \\ d^1 \end{pmatrix} - \begin{pmatrix} b^2 \\ d^2 \end{pmatrix} \right\|_2 .$$

This result can be interpreted as implying that feasible solutions of (91) are Lipschitz continuous with respect to right hand side constants vector perturbations, with Lipschitz constant $\mu(A,B)$ depending only on the coefficient matrix $\begin{pmatrix} A \\ B \end{pmatrix}$.

(iii) In (91), if B is of full row rank and the system "$Ay < 0$, $By = 0$" has a solution y, prove that $\mathbf{K}(b,d) \neq \emptyset$ for all $\begin{pmatrix} b \\ d \end{pmatrix} \in \mathbf{R}^{m+p}$, and that for any $\begin{pmatrix} b^1 \\ d^1 \end{pmatrix}$, $\begin{pmatrix} b^2 \\ d^2 \end{pmatrix} \in \mathbf{R}^{m+p}$, and $x^1 \in \mathbf{K}(b^1,d^1)$, there exists an $x^2 \in \mathbf{K}(b^2,d^2)$ satisfying

$$\|x^1 - x^2\|_\infty \leqq \overline{\mu}(A,B) \left\| \begin{pmatrix} b^1 \\ d^1 \end{pmatrix} - \begin{pmatrix} b^2 \\ d^2 \end{pmatrix} \right\|_2$$

where

$$\begin{aligned} \overline{\mu}(A,B) = &\text{ maximum } \|u,v\|_2 \\ &\text{subject to, } u,\ v \text{ are row vectors in } \mathbf{R}^m,\ \mathbf{R}^p, \\ \|uA+vB\|_1 = &\ 1 \\ u \geqq &\ 0. \end{aligned} \tag{93}$$

(iv) Suppose $\begin{pmatrix} b^1 \\ d^1 \end{pmatrix}$ is such that $\mathbf{K}(b^1,d^1) \neq \emptyset$. For any $x \in \mathbf{R}^n$, prove that there exists an $x^1 \in \mathbf{K}(b^1,d^1)$ satisfying

$$\|x - x^1\|_\infty \leqq \mu(A,B) \left\| \begin{matrix} (Ax - b^1)^+ \\ (Bx - d^1) \end{matrix} \right\|_2 .$$

If the Lipschitz constant $\mu(A, B)$ is available, this inequality provides an error bound on how far x is from a feasible solution of (91).

(v) Consider the LP

$$\begin{array}{ll} \text{minimize} & cx \\ \text{subject to} & Ax \leqq b \\ & Bx = d \end{array} \tag{94}$$

with A, B fixed, let $\mathbf{\Gamma}(b, d)$ denote the set of optimum solutions of (94). If it is known that $\mathbf{\Gamma}(b^1, d^1)$ and $\mathbf{\Gamma}(b^2, d^2)$ are both nonempty; for any $x^1 \in \mathbf{\Gamma}(b^1, d^1)$ prove that there exists an $x^2 \in \mathbf{\Gamma}(b^2, d^2)$ satisfying

$$\|x^1 - x^2\|_\infty \leqq \mu(A, B) \left\| \begin{pmatrix} b^1 \\ d^1 \end{pmatrix} - \begin{pmatrix} b^2 \\ d^2 \end{pmatrix} \right\|_2$$

where $\mu(A, B)$ is the Lipschitz constant defined in (92). This result can be interpreted as implying that optimum solutions of LPs are Lipschitz continuous with respect to right hand side constants vector perturbations.

(vi) Consider the LP

$$\begin{array}{lll} \text{minimize} & -(1+\delta)x_1 - x_2 & \\ \text{subject to} & x_1 + x_2 \leqq 1 & \\ & x_1, x_2 \geqq 0 & \end{array} \tag{95}$$

where δ is a real parameter. Show that when $\delta > -1$ and $\delta \neq 0$, this problem has a unique optimum solution $x(\delta)$ given by

$$x(\delta) = \begin{cases} (1, 0)^T, & \text{if } \delta > 0 \\ (0, 1)^T, & \text{if } -1 < \delta < 0. \end{cases}$$

By showing that

$$\lim_{\delta \to 0^+} \frac{\|x(\delta) - x(-\delta)\|}{2\delta} = +\infty$$

prove that $x(\delta)$ is not Lipschitzian with respect to δ.

This shows that in general, optimum solutions of linear programs are not Lipschitzian with respect to perturbations in the objective function coefficients.

(vii) Consider the LCP (q, M) of order n. Let $\mathbf{J} \subset \{1, \ldots, n\}$. Consider the system

$$\begin{array}{ll} M_{i.}z + q_i \geqq 0,\ z_i = 0, & \text{for all } i \in \mathbf{J} \\ M_{i.}z + q_i = 0,\ z_i \geqq 0, & \text{for all } i \notin \mathbf{J}. \end{array} \tag{96}$$

If z is any solution of (96) then z leads to a solution of the LCP (q, M) (that is, $(w = Mz + q, z)$ is a solution of the LCP (q, M)). Using this fact, Lipschitz continuity of solutions with respect to the right hand side constants vector perturbations, can be established for certain classes of LCPs.

For any $\mathbf{J} \subset \{1, \ldots, n\}$, define $A(\mathbf{J})$ to be the square matrix of order n such that

$$(A(\mathbf{J}))_{i.} = \begin{cases} -M_{i.} & \text{for } i \in \mathbf{J} \\ I_{i.} & \text{for } i \notin \mathbf{J}. \end{cases}$$

Similarly, define the square matrix $B(\mathbf{J})$ of order n by

$$(B(\mathbf{J}))_{i.} = \begin{cases} I_{i.} & \text{for } i \in \mathbf{J}. \\ -M_{i.} & \text{for } i \notin \mathbf{J}. \end{cases}$$

Now define

$$\sigma(M) = \text{ maximum } \{\mu(A(\mathbf{J}), B(\mathbf{J})) : \mathbf{J} \subset \{1, \ldots, n\}\}$$

where $\mu(A(\mathbf{J}), B(\mathbf{J}))$ is $\mu(A, B)$ of (92) with $A = A(\mathbf{J})$, $B = B(\mathbf{J})$.

Suppose M is a P-matrix and (w^r, z^r) is the unique solution of the LCP (q, M) when $q = q^r$, $r = 1, 2$. Prove that

$$\|z^1 - z^2\|_\infty \leqq \sigma(M)\|q^1 - q^2\|_2.$$

This establishes that when M is a P-matrix and fixed, the solution of the LCP (q, M) is Lipschitz continuous in q with Lipschitz constant $\sigma(M)$ defined above.

(viii) Let $M = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, $q^1 = \begin{pmatrix} -\varepsilon \\ 1 \end{pmatrix}$, $q^2 = \begin{pmatrix} \varepsilon \\ 1 \end{pmatrix}$ where $\varepsilon > 0$. Show that if $z^1 = (1, \varepsilon)^T$, $z^2 = (0, 0)^T$, the solution of the LCP (q^r, M) is $(w^r = Mz^r + q, z^r)$, $r = 1, 2$. Verify that

$$\lim_{\varepsilon \to 0^+} \frac{\|z^1 - z^2\|_\infty}{2\varepsilon} = +\infty.$$

This shows that the solution of the LCP (q, M) may not be Lipschitzian in q for fixed M, when M is positive semidefinite but not a P-matrix. (O. L. Mangasarian and T. H. Shiau [A11])

49. Let A, b be given real matrices of orders $m \times n$ and $m \times 1$ respectively. Consider the system of equations

$$Ax = b. \tag{97}$$

This system may or may not have a solution. It is required to find a vector x that satisfies (97) as closely as possible using the least squares measure of deviation. Formulate this as a nonlinear program and write down the optimality conditions for it. Prove that this system of optimality conditions always has a solution.

Now consider the problem of finding a vector x satisfying (97) as closely as possible, subject to the additional constraints $\|x\| = 1$, which is required to be satisfied. This leads to the nonlinear program

$$\begin{array}{ll} \text{minimize} & \|Ax - b\|^{\mathbf{2}} \\ \text{subject to} & \|x\|^2 = 1. \end{array} \tag{98}$$

Discuss how (98) can be solved to optimality efficiently.

50. Let $f(x)$ be a real valued function defined on $\mathbf{R}^n$ which is thrice continuously differentiable. Consider the NLP

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & x \geqq 0. \end{array}$$

i) Prove that the first order necessary optimality conditions for this NLP can be posed as a nonlinear complementarity problem.

ii) Let $f_i(x) = \frac{\partial f(x)}{\partial x_i}$, $i = 1$ to n. Define $g(x) = (g_i(x) : i = 1 \text{ to } n)^T$ where

$$g_i(x) = -|x_i - f_i(x)|^{\mathbf{3}} + x_i^{\mathbf{3}} + (f_i(x))^{\mathbf{3}}.$$

Prove that solving he NLCP described in (i) above, is equivalent to solving the system of n equations in n unknowns

$$g(x) = 0.$$

Show that $g(x)$ is twice continuously differentiable. (L. Watson [A15])

51. We have received a large shipment of an engineering item. A random sample of 10 items selected from this lot had the following lifetimes in time units.

1.600	1.506	0.501	1.118
0.295	0.070	1.821	
0.055	0.499	3.102	

Assume that the lifetime, x, of items from the lot follows a Weibull distribution with the following probability density function

$$f(x) = \alpha\theta x^{\theta-1} e^{(-\alpha x^{\theta})},\ x \geqq 0.$$

Formulate the problem of obtaining the maximum liklihood estimators for the parameters α, θ as an NLP. Write down the optimality conditions for this NLP, and solve this NLP using them.

52. Consider the convex polyhedra $\mathbf{K}_1$, $\mathbf{K}_2$, which are the sets of feasible solutions of the systems given below

$\mathbf{K}_1$	$\mathbf{K}_2$
$Ax = b$	$Dy = d$
$x \geqq 0$	$y \geqq 0.$

It is required to find a pair of points $(x; y)$, $x \in \mathbf{K}_1$, $y \in \mathbf{K}_2$, which are closest in terms of the Euclidean distance, among all such pairs. Does this problem have a unique optimum solution? Why?

Formulate this problem as an NLP and write down the necessary optimality conditions for it. Are these conditions also sufficient for optimality for this problem?

53. Write down the necessary optimality conditions for Sylvester's problem, Exercise 1.25, and determine whether these conditions are also sufficient for optimality.

54. We are given smooth real valued functions $\theta_1(x), \ldots, \theta_r(x), g_1(x), \ldots, g_m(x)$, all defined over $\mathbf{R}^n$. Consider the following optimization problem:

$$\begin{array}{ll} \text{minimize} & v(x) \\ \text{subject to} & g_i(x) \geqq 0,\ i = 1 \text{ to } m \end{array}$$

where for each $x \in \mathbf{R}^n$, $v(x) =$ maximum $\{\theta_1(x), \ldots, \theta_r(x)\}$. Transform this problem into a smooth NLP with a linear objective function, but with additional constraints than those in this problem. Write down the necessary optimality conditions for the transformed problem and simplify them. State some general conditions on the data in the problem under which these conditions are also sufficient for optimality. Show that this technique can be used to transform any NLP into an NLP in which the objective function is linear.

55. The army has n types of weapons available. Using them, they want to destroy m targets. The following data is given:

$p_{ij} =$ probability that a weapon of type j shot at target type i will destroy it,
$v_i =$ value in \$ of target i,
$b_j =$ number of weapons of type j available.

Assume that a weapon shot at a target either destroys it, or leaves it absolutely unaffected.

Formulate the problem of determining the number of weapons of each type to be shot at each of the targets, so as to maximize the expected value destroyed, as an NLP. Neglecting the integer requirements on the decision variables in this problem, write down the necessary optimality conditions for it. Specialize these for the numerical example with the following data.

$$n = 2,\ m = 3$$

$$p = (p_{ij}) = \begin{pmatrix} .25 & .05 \\ .35 & .08 \\ .15 & .17 \end{pmatrix},\ v = (v_i) = \begin{pmatrix} 150 \\ 95 \\ 375 \end{pmatrix},\ b = (b_i) = \begin{pmatrix} 6 \\ 10 \end{pmatrix}.$$

56. Let B, A be matrices of order $n \times n$ and $m \times n$ respectively. Suppose rank$(A) = m$ and B is symmetric and PD on the subspace $\{x : Ax = 0\}$. Then prove that the matrix $\begin{pmatrix} B & A^T \\ A & O \end{pmatrix}$ is nonsigular.

57. Let $f(x)$ be a real valued convex function defined on $\mathbf{R}^n$. Assume that $f(x)$ is twice continuously differentiable at a given point $\overline{x} \in \mathbf{R}^n$. Define

$$l(x) = f(\overline{x}) + \nabla f(\overline{x})(x - \overline{x})$$
$$Q(x) = f(\overline{x}) + \nabla f(\overline{x})(x - \overline{x}) + \tfrac{1}{2}(x - \overline{x})^T H(f(\overline{x}))(x - \overline{x}).$$

The functions $l(x)$, $Q(x)$ are respectively the first and second order Taylor approximations for $f(x)$ around $\overline{x}$. In Theorem 15 we established that $f(x) - l(x)$ always has the same sign ($\geqq 0$) for all $x \in \mathbf{R}^n$. Discuss whether $f(x) - Q(x)$ always has the same sign for all $x \in \mathbf{R}^n$. If so, what is that sign? Why? (Richard Hughes)

58. Let $\mathbf{K}$ denote the set of feasible solutions of

$$Ax \geqq b \tag{99}$$

where A is an $m \times n$ matrix. We know that $\mathbf{K} \neq \emptyset$ and dimension of $\mathbf{K}$ is n. $\theta(x)$ is a strictly convex function defined on $\mathbf{R}^n$, with a unique unconstrained minimum in $\mathbf{R}^n$, $\overline{x}$. We know that $\overline{x}$ satisfies all but one constraint in (99). Suppose $A_i.\overline{x} \geqq b_i$ for $i = 2$ to m, but $A_1.\overline{x} < b_1$. Prove that if the problem

$$\begin{array}{ll} \text{minimize} & \theta(x) \\ & x \in \mathbf{K} \end{array}$$

has an optimum solution, the first constraint in (99) must be active at it. What is the appropriate generalization of this result when $\overline{x}$ violates more than one constraint in (99)? (M. Q. Zaman, S. U. Khan, and A. Bari, private communication).

59. Let $f(\lambda) : \mathbf{R}^1 \to \mathbf{R}^1$ be a continuously differentiable strictly increasing function of the real parameter λ.

Let $\theta(x) : \mathbf{R}^n \to \mathbf{R}^1$, $g(x) : \mathbf{R}^n \to \mathbf{R}^m$, $h(x) : \mathbf{R}^n \to \mathbf{R}^t$ be continuously differentiable functions.

Consider the constraint system

$$\begin{aligned} g(x) &\geqq 0 \\ h(x) &= 0 \end{aligned} \tag{100}$$

and the two optimization problems

Problem 1: Minimize $\theta(x)$, subject to (100)

Problem 2: Minimize $f(\theta(x))$, subject to (100).

Rigorously prove that both the problems have the same set of stationary points. (H. L. Li, private communication.)

60. Consider the following separable NLP

$$\begin{aligned} \text{minimize} \quad & \sum_{j=1}^{n} f_j(x_j) \\ \text{subject to} \quad & \sum_{j=1}^{n} x_j \geqq 1 \\ & x_j \geqq 0,\ j = 1 \text{ to } n \end{aligned}$$

where $f_j(x_j)$ is a differentiable function for all j. If $\overline{x} = (\overline{x}_j)$ solves this problem, prove that there must exist a nonnegative scalar k such that

$$\frac{df_j(\overline{x}_j)}{dx_j} \geqq k \text{ for all } j,$$

$$\text{and for all } j \text{ such that } \overline{x}_j > 0\,,\ \frac{df_j(\overline{x}_j)}{dx_j} = k.$$

61. Let A be a given matrix of order $m \times n$. Prove that at least one of the following systems

(I)	(II)
$Ax \geqq 0$	$\pi A \leqq 0$
$x \geqq 0$	$\pi \geqq 0$

has a nonzero feasible solution.

62. Consider the following LP

$$\begin{aligned} \text{minimize} \quad & cx \\ \text{subject to} \quad & Ax \geqq b \\ & x \geqq 0. \end{aligned}$$

Let $\mathbf{K}$, $\mathbf{\Gamma}$ denote the set of feasible solutions of the LP, and its dual respectivley. Prove that either both $\mathbf{K}$ and $\mathbf{\Gamma}$ are empty, or at least one of $\mathbf{K}$, $\mathbf{\Gamma}$ is an unbounded set.

63. Let D be the diagonal matrix $\text{diag}(\lambda_1, \ldots, \lambda_n)$, where $\lambda_1 \geqq \lambda_2 \geqq \ldots \geqq \lambda_n > 0$. Consider the following NLP

$$\begin{aligned} \text{minimize} \quad & (y^T D y)(y^T D^{-1} y) \\ \text{subject to} \quad & y^T y = 1. \end{aligned}$$

(i) Transform this NLP into another problem in new variables $x_1, \ldots, x_n$ in which the objective function to be optimized is a product of two linear functions, $g(x)$ and $h(x)$, say, and the constraints are all linear. Call this transformed problem (P).

(ii) Show that (P) must have a global optimum solution.

Assume that x^* is an optimum solution of (P). Let $h(x^*) = \delta$. Show that x^* must also be an optimum solution of the LP in the variables $x = (x_1, \ldots, x_n)$ in which the objective function to be optimized is $g(x)$, and the constraints are the same as those in (P) plus the additional constraint $h(x) = \delta$. Conversely, show that every optimum solution of this LP must also be optimal to (P). Using this, show that (P) has an optimum solution in which two variables among $x_1, \ldots, x_n$ are positive, and the others are all zero.

(iii) Consider the problem (P) again. In this problem, substitute $x_i = 0$ for all $i \neq p, q$, for some selected p, q between 1 to n. Show that in the optimum solution of this reduced problem, both x_p and x_q are equal.

(iv) Use the above results to prove Kantorovich's inequality which states the following: Let A be a symmetric PD matrix of order n with eigen values $\lambda_1 \geqq \lambda_2 \geqq \ldots \geqq \lambda_n > 0$. Then

$$(y^T A y)(y^T A^{-1} y) \leqq \frac{(\lambda_1 + \lambda_n)^{\mathbf{2}}}{4\lambda_1 \lambda_n} \text{ for all } y \text{ such that } \|y\| = 1.$$

(M. Raghavachari, "A linear programming proof of Kantorovich's inequality", The American Statistician, 40 (1986) 136–137.)

64. $f(x)$ is a real valued differentiable function defined on $\mathbf{R}^n$. Prove that $f(x)$ is a convex function iff

$$(\nabla f(x^2) - \nabla f(x^1))(x^2 - x^1) \geqq 0,$$

for all $x^1, x^2 \in \mathbf{R}^n$. Similarly, prove that a real valued differentiable function $g(x)$ defined on $\mathbf{R}^n$ is concave iff

$$(\nabla g(x^2) - \nabla g(x^1))(x^2 - x^1) \leqq 0,$$

for all $x^1, x^2 \in \mathbf{R}^n$.

65. Let $f(x)$ be a real valued convex function defined on $\mathbf{R}^n$. For each $x \in \mathbf{R}^n$ let $f^+(x)$ and $f^-(x)$ denote the positive and negative parts of $f(x)$, that is, $f^+(x)$ and $f^-(x)$ satisfy for all $x \in \mathbf{R}^n$, $f^+(x) \geqq 0$, $f^-(x) \geqq 0$, $f(x) = f^+(x) - f^-(x)$, $(f^+(x))(f^-(x)) = 0$. Are $f^+(x)$ and $f^-(x)$ both convex functions over $\mathbf{R}^n$? Why?

66. Consider the linearly constrained NLP

$$\begin{array}{ll} \text{minimize} & \theta(x) \\ \text{subject to} & Ax = b \end{array}$$

where $\theta(x)$ is a real valued continuously differentiable convex function defined on $\mathbf{R}^n$, and A is an $m \times n$ matrix of rank m. If x^* is a feasible solution for this problem satisfying

$$\nabla\theta(x^*)(I - A^T(AA^T)^{-1}A) = 0$$

prove that x^* is an optimum solution of the NLP.

67. Consider the NLP

$$\begin{array}{ll} \text{minimize} & cx \\ \text{subject to} & g_i(x) \geqq 0,\ i = 1 \text{ to } m \end{array}$$

where $c = (c_1, \ldots, c_n) \neq 0$, and $g_i(x)$ is a continuously differentiable real valued function defined over $\mathbf{R}^n$ for each $i = 1$ to m. Suppose x^* is a local minimum for this problem, and is a regular point. Prove that there exists at least one i such that $g_i(x^*) = 0$.

68. i) On the x_1, x_2-Cartesian plane, find the nearest point on the parabola $\{x = (x_1, x_2)^T : x_2^2 = 4x_1\}$ to $(1, 0)^T$ in terms of the Euclidean distance.
ii) For the following NLP, check whether either of $x^1 = (1, \frac{1}{2})^T$ or $x^2 = (\frac{1}{3}, -\frac{1}{6})^T$ are optimum solutions

$$\begin{array}{ll} \text{minimize} & x_1^2 + 2x_2^2 \\ \text{subject to} & x_1^2 + \ x_2^2 \leqq 5 \\ & 2x_1 - 2x_2 = 1. \end{array}$$

69. Let $f(x)$ be a real valued continuously differentiable convex function defined over $\mathbf{R}^n$ and let $\mathbf{K}$ be a closed convex subset of $\mathbf{R}^n$. Suppose $\overline{x} \in \mathbf{K}$ is such that it is the nearest point (in terms of the Euclidean distance) in $\mathbf{K}$ to $\overline{x} - \lambda\nabla f(\overline{x})$ for some $\lambda > 0$. Prove that $\overline{x}$ minimizes $f(x)$ over $x \in \mathbf{K}$. Construct the converse of this statement and prove it too.

70. Consider the following NLP

$$\begin{array}{ll} \text{minimize} & \theta(x) \\ \text{subject to} & l \leqq x \leqq k \end{array}$$

where $\theta(x)$ is a real valued twice continuously differentiable function defined on $\mathbf{R}^n$, and l, k are two bound vectors in $\mathbf{R}^n$ satisfying $l < k$. Develop an algorithm for solving this problem, which takes advantage of the special structure of the problem. Write down the termination criteria that you would use, and provide a justification for them. Also, mention what type of a solution the algorithm is guaranteed to obtain at termination.

71. Consider the following NLP

$$\begin{array}{ll} \text{minimize} & \theta(x) \\ \text{subject to} & a_i \leqq g_i(x) \leqq b_i(x),\ i = 1 \text{ to } m \\ \text{and} & l \leqq x \leqq k \end{array}$$

where $\theta(x)$, $g_i(x), \ldots, g_n(x)$ are all real valued twice continuously differentiable functions defined over $\mathbf{R}^n$, and $a = (a_i)$, $b = (b_i)$, l, k satisfy $a \leqq b$, $l < k$. Discuss how you can solve this problem using the algorithm developed in Exercise 70.

72. $\theta(x)$ is a real valued continuously differentiable convex function defined over $\mathbf{R}^n$. $\mathbf{K}$ is a closed convex subset of $\mathbf{R}^n$. If $\overline{x} \in \mathbf{K}$ is the global maximum for $\theta(x)$ over $x \in \mathbf{K}$, prove that

$$\nabla\theta(\overline{x})x \leqq \nabla\theta(\overline{x})\overline{x}, \quad \text{for all } x \in \mathbf{K}.$$

Is the converse of this statement also true? Why?

Would the above inequality hold for all $x \in \mathbf{K}$ if x is only a local maximum for $\theta(x)$ over $\mathbf{K}$ and not a global maximum? Why?

73. If M is a P-matrix of order n (not necessarily PD) prove that the system

$$\begin{aligned} \pi M &> 0 \\ \pi &\geqq 0 \end{aligned}$$

has a solution π.

74. Write down the first order necessary optimality conditions for the following NLP, and find an optimum solution for it.

$$\begin{aligned} \text{minimize} \quad & (x_1 - 4)^{\mathbf{2}} + (x_2 + 1)^{\mathbf{2}} \\ \text{subject to} \quad & 7 \leqq x_1 \leqq 14 \\ & 10 \leqq x_2 \leqq 22. \end{aligned}$$

75. Consider the following linear program

$$\begin{aligned} \text{minimize } z(x) = \quad & = cx \\ \text{subject to} \quad & Ax = b \\ & Dx \geqq d. \end{aligned}$$

Let $\mathbf{K}$ denote the set of feasible solutions for this problem. Show that the primal simplex algorithm for this problem, is exactly the gradient projection method (Section 10.10.5) applied on this problem, beginning with a feasible point x^0 which is an extreme point of $\mathbf{K}$.

76. $\theta(x)$; $h_i(x)$, $i = 1$ to m; $g_p(x)$, $p = 1$ to t are all real valued twice continuously differentiable functions defined over $\mathbf{R}^n$.

i) Consider the NLP

$$\begin{aligned} \text{minimize} \quad & \theta(x) \\ \text{subject to} \quad & h_i(x) = 0, \ i = 1 \text{ to } m, \\ & g_p(x) \geqq 0, \ p = 1 \text{ to } t. \end{aligned}$$

Let $L(x, \mu, \pi) = \theta(x) - \sum_{i=1}^{m} \mu_i h_i(x) - \sum_{p=1}^{t} \pi_p g_p(x)$ be the Lagrangian. Suppose we have a feasible solution $\overline{x}$ to this NLP and Lagrange multiplier vectors $\overline{\mu}$, $\overline{\pi}$ such that $(\overline{x}, \overline{\mu}, \overline{\pi})$ satisfy the first order necessary optimality conditions for this NLP, and the additional condition that $L(x, \overline{\mu}, \overline{\pi})$ is a convex function in x over $\mathbf{R}^n$ (notice that $L(x, \overline{\mu}, \overline{\pi})$ could be a convex function, even though $\theta(x)$, $-g_p(x)$, $h_i(x)$ and $-h_i(x)$ are not all convex functions). Then prove that $\overline{x}$ must be a global minimum for this NLP.

ii) Consider the numerical example

$$\begin{array}{ll} \text{minimize} & (x_1 - \alpha)^{\mathbf{2}} + (x_2 - \alpha)^{\mathbf{2}} \\ \text{subject to} & x_1^2 - 1 = 0 \\ & 1 - x_2^2 \leqq 0 \end{array}$$

where α is any real number satisfying $\|\alpha\| \leqq 1$. Let $\overline{x} = (\overline{x}_1, \overline{x}_2)^T = (1, 1)^T$, $\overline{\mu} = (\alpha - 1)$, $\overline{\pi} = (1 - \alpha)$. Verify that $\overline{x}$ is a global minimum for this problem using the result in (i). (P. Mereau and J. A. Paquet, "A sufficient condition for global constrained extrema", *Int. J. Control*, 17 (1973) 1065–1071).

77. $\theta(x)$; $g_i(x)$, $i = 1$ to m are all real valued convex functions defined over $\mathbf{R}^n$. Consider the NLP

$$\begin{array}{ll} \text{minimize} & \theta(x) \\ \text{subject to} & g_i(x) \leqq 0, \ i = 1 \text{ to } m. \end{array}$$

i) Prove that the set of all optimum solutions of this problem is a convex set.

ii) A real valued function defined on $\mathbf{R}^n$ is said to be a symmetric function if $f(x) = f(Px)$, for all $x \in \mathbf{R}^n$, and P any permutation matrix of order n. If all the functions $\theta(x)$, $g_i(x)$, $i = 1$ to m, are symmetric functions, and the above problem has an optimum solution, prove that it has one in which all the variables are equal.

Exercises 78 to 98 have been suggested to me by Vasant A. Ubhaya.

---

78. Let $\mathbf{J}$ be an interval of the real line. $f(x)$ is a real valued function defined on $\mathbf{J}$. Prove that $f(x)$ is convex iff for any three point x, y, z in $\mathbf{J}$ with $x < y < z$,

$$\text{determinant} \begin{vmatrix} x & f(x) & 1 \\ y & f(y) & 1 \\ z & f(z) & 1 \end{vmatrix} \geqq 0.$$

79. Let $a_1 \geqq a_2 \geqq \ldots \geqq a_n \geqq 0$ and let $f(x)$ be a real valued convex function defined on the interval $[0, a_1]$ with $f(0) = 0$. Show that

$$\sum_{k=1}^{n} (-1)^{\mathbf{k-1}} f(a_k) \geqq f\Big(\sum_{k=1}^{n} (-1)^{\mathbf{k-1}} a_k\Big). \qquad \text{(contd.)}$$

(E. F. Beckenbach and R. Bellman, *"Inequalities"*, Springer-Verlag, New York, 1983, and E. M. Wright, "An inequality for convex functions", *American Mathematical Monthly*, 61 (1984) 620-622.)

80. Let $\mathbf{J}$ be a closed interval of the real line. A real valued function $f(x)$ defined on $\mathbf{J}$ is said to be midconvex or Jensen-convex if

$$f\Big(\frac{x+y}{2}\Big) \leqq \frac{1}{2}(f(x)+f(y))$$

for all $x, y \in \mathbf{J}$. Prove that if $f(x)$ is midconvex, then

$$f(\lambda x + (1-\lambda)y) \leqq \lambda f(x) + (1-\lambda)f(y)$$

for all $x, y \in \mathbf{J}$ and all rational numbers λ between 0 and 1. Hence conclude that a continuous function is midconvex iff it is convex. (A. W. Roberts and D. E. Varberg, *"Convex Functions"*, Academic Press, New York, 1973.)

81. Let $\mathbf{\Gamma} \subset \mathbf{R}^n$ be a convex set, and let $f(x)$ be a real valued convex function defined on $\mathbf{\Gamma}$. Let $g(\lambda)$ be a nondecreasing convex function defined on a real interval $\mathbf{J}$ where the range of $f(x)$ is contained in $\mathbf{J}$. Prove that $h(x) = g(f(x))$ is convex. Use this to show the following:

a) If $f(x)$ is a positive concave function defined on $\mathbf{\Gamma}$, then $1/f(x)$ is convex.

b) If $f(x)$ is a nonnegative convex function defined on $\mathbf{\Gamma}$, then $(f(x))^{\mathbf{r}}$ is convex for $r \geqq 1$.

c) If $f(x)$ is a convex function defined on $\mathbf{\Gamma}$, then $\exp(f(x))$ is convex.

82. Let $y = (y_1, \ldots, y_{n-1})^T$, and $\mathbf{\Gamma} = \{y : 0 = y_0 < y_1 < \ldots < y_{n-1} < y_n = 1\}$. Define, for $j = 0, 1, \ldots, n$

$$\begin{aligned} D_j(y) &= \prod(|y_i - y_j| : \text{ over } 0 \leqq i \leqq n,\ i \neq j) \\ &= (-1)^j \prod((y_i - y_j) : \text{ over } 0 \leqq i \leqq n, i \neq j) \end{aligned}$$

where $\prod$ denotes the product sign. Define

$$F(y) = \sum((D_{n-1-2j}(y))^{-1} : \text{ over } 0 \leqq j \leqq \lfloor (n-1)/2 \rfloor).$$

Show that $F(y)$ is a strictly convex function of y over $\mathbf{\Gamma}$. Prove that $y^* = ((\sin(\pi/2n)^{\mathbf{2}}, (\sin(2\pi/2n))^{\mathbf{2}}, \ldots, (\sin((n-1)\pi/2n))^{\mathbf{2}})^T$ is the unique optimum solution for the problem of minimizing $F(y)$ over $\mathbf{\Gamma}$, and that $F(y^*) = 2^{\mathbf{2n-2}}$. Prove the following inequalities for all $y \in \mathbf{\Gamma}$.

(i) $\sum((D_j(y))^{-1} :$ over j odd, $1 \leqq j \leqq n-1) \geqq 2^{\mathbf{2n-2}}$, if n is even.

(ii) $\sum((D_j(y))^{-1} :$ over j even, $0 \leqq j \leqq n-1) \geqq 2^{\mathbf{2n-2}}$, if n is odd.

(iii) $\sum((D_j(y))^{-1} :$ over j odd, $1 \leqq j \leqq n\ \ ) \geqq 2^{\mathbf{2n-2}}$, if n is odd.

(iv) $\sum((D_j(y))^{-1} :$ over $0 \leqq j \leqq n\ \ ) \geqq 2^{\mathbf{2n-1}}$, if n is odd.

Furthermore prove that each of the above inequalities holds as an equation iff $y = y^*$ defined above. (V. A. Ubhaya "Nonlinear programming, approximation and optimization on infinitely differentiable functions", *Journal of Optimization Theory and Applications*, 29 (1979), 199-213.)

83. Let $\mathbf{S} \subset \mathbf{R}^{n+1}$ be a convex set. Define a set $\mathbf{\Gamma} \subset \mathbf{R}^n$ and a real valued function $f(x)$ on $\mathbf{\Gamma}$ as follows.

$$\mathbf{\Gamma} = \{x \in \mathbf{R}^n : u \in \mathbf{R}^1, (x, u) \in \mathbf{S}\}.$$
$$f(x) = \inf\{u : x \in \mathbf{\Gamma}, (x, u) \in \mathbf{S}\}.$$

Show that $\mathbf{\Gamma}$ is convex and $f(x)$ is a convex function on $\mathbf{\Gamma}$.

84. Let $\mathbf{\Gamma} \subset \mathbf{R}^n$ and $f(x)$ be any real valued function defined on $\mathbf{\Gamma}$. The epigraph $\mathbf{E}(f)$ of $f(x)$ is a subset of $\mathbf{R}^{n+1}$ defined as in Appendix 3. Assume that $\mathbf{\Gamma}$ is closed, and show that $\mathbf{E}(f)$ is closed iff $f(x)$ is lower semi-continuous. In particular, $\mathbf{E}(f)$ is closed if $\mathbf{\Gamma}$ is closed and $f(x)$ is continuous.

85. Let $\mathbf{\Gamma} \subset \mathbf{R}^n$ be a convex set and $f(x)$ be a real valued bounded function defined on $\mathbf{\Gamma}$. The greatest convex minorant $\overline{f}(x)$ of $f(x)$ is the largest convex function which does not exceed $f(x)$ at any point in $\mathbf{\Gamma}$, viz.,

$$\overline{f}(x) = \sup\{h(x) : h(y) \text{ is convex and } h(y) \leqq f(y) \text{ for all } y \text{ in } \mathbf{\Gamma}\}, x \in \mathbf{\Gamma}.$$

Show that $\overline{f}(x)$ defined in this way is, indeed, convex. If $\mathbf{E}(f)$ is the epigraph of $f(x)$ then show that

$$\overline{f}(x) = \inf\{u : x \in \mathbf{\Gamma}, (x, u) \in \text{co}(\mathbf{E}(f))\},$$

where $\text{co}(\mathbf{E}(f))$ is the convex hull of $\mathbf{E}(f)$, i. e., the smallest convex subset of $\mathbf{R}^{n+1}$ containing $\mathbf{E}(f)$.

86. Let $\mathbf{\Gamma} \subset \mathbf{R}^n$ be convex and $f(x)$ be a real valued convex function defined on $\mathbf{\Gamma}$. Assume $0 \leqq f(x) < 1$. Show that $(1+f(x))^{\mathbf{1/2}}$ and $(1-f(x))^{\mathbf{-1/2}}$ are convex functions on $\mathbf{\Gamma}$. Is $((1 + f(x))/(1 - f(x)))^{\mathbf{1/2}}$ convex?

87. Let $f(x)$ be a real homogeneous polynomial of degree 2 defined on $\mathbf{R}^n$, i. e.,

$$f(x) = \sum_i a_i x_i^{\mathbf{2}} + \sum_{i<j} b_{ij} x_i x_j,$$

where $x = (x_1, x_2, \ldots, x_n)$; and a_i, b_{ij} are given numbers. Show that $f(x)$ is convex iff $f(x)$ is nonnegative on $\mathbf{R}^n$.

88. Let $f(x)$ be a real valued function defined on the interval $\mathbf{J} = [0, 1]$. The nth ($n \geqq 1$) Bernstein polynomial for $f(x)$ is defined by

$$B_n(f, x) = \sum_{k=0}^{n} f\Big(\frac{k}{n}\Big)\binom{n}{k} x^{\mathbf{k}}(1-x)^{\mathbf{n-k}}.$$

Note that $B_n(f, 0) = f(0)$ and $B_n(f, 1) = f(1)$. Show the following:

(a) If $f(x)$ is nondecreasing on $\mathbf{J}$, then $B_n(f, x)$ is nondecreasing on $\mathbf{J}$.

(b) If $f(x)$ is convex on $\mathbf{J}$, then $B_n(f, x)$ is convex on $\mathbf{J}$. In this case, $B_{n-1}(f, x) \geqq B_n(f, x)$ for $0 < x < 1$ and $n \geqq 2$.

(c) If $f(x)$ is bounded on $\mathbf{J}$, then $B_n(f, x) \to f(x)$ as $n \to \infty$ at any point x in $\mathbf{J}$ at which $f(x)$ is continuous. Furthermore, if $f(x)$ is continuous on $\mathbf{J}$, then this convergence is uniform on $\mathbf{J}$. Hence conclude that the class of nondecreasing (convex) polynomials on $\mathbf{J}$ are dense in the class of continuous nondecreasing (convex) functions on $\mathbf{J}$ when the uniform norm $\|f\| = \max\{|f(x)| : x \in \mathbf{J}\}$ is used to generate a metric for the set of continuous functions $f(x)$.

(P. J. Davis, *"Interpolation and Approximation"*, Dover, New York, 1975).

89. Let $\mathbf{\Gamma}$ be a convex subset of $\mathbf{R}^n$ and $f(x)$ a real valued function defined on $\mathbf{\Gamma}$. $f(x)$ is said to be a quasiconvex function if $\{x \in \mathbf{\Gamma} : f(x) \leqq \alpha\}$ is a convex set for all real α.

A real valued function $g(x)$, defined on a convex set is said to be quasiconcave, if $-g(x)$ is quasiconvex.

Show that $f(x)$ is quasi-convex on $\mathbf{\Gamma}$ iff

$$f(\lambda x + (1-\lambda)y) \leqq \max\{f(x), f(y)\}$$

holds for all $x, y \in \mathbf{\Gamma}$, all $0 \leqq \lambda \leqq 1$.

90. The following result is well known:

Let $\mathbf{\Gamma} \subset \mathbf{R}^n$ and $\mathbf{\Delta} \subset \mathbf{R}^m$ be compact convex subsets. Let $h(x, y)$ be a continuous real valued function defined on $\mathbf{\Gamma} \times \mathbf{\Delta}$ be such that, for each $y \in \mathbf{\Delta}$, $h(x, y)$ is a quasiconcave function of x; and for each $x \in \mathbf{\Gamma}$, $h(x, y)$ is a quasiconvex function of y. Then,

$$\min_{y\in\Delta}\max_{x\in\Gamma} h(x, y) = \max_{x\in\Gamma}\min_{y\in\Delta} h(x, y).$$

(See, e. g., H. Nikaidô, "On Von Neumann's minimax theorem", *Pacific Journal of Mathematics*, 4 (1954), 65–72, for the above result and M. Sion, "On general minimax theorems", *Pacific Journal of Mathematics*, 8 (1958), 171–176, for more general versions.) Using the above result, derive the following:

Let $\mathbf{K}$, $\mathbf{P}$ be bounded subsets of $\mathbf{R}^2$ with the property that there exists a $\delta > 0$ such that $u_1 \geqq \delta$ for all $u = (u_1, u_2)^T \in \mathbf{K}$ and $v_1 \geqq \delta$ for all $v = (v_1, v_2)^T \in \mathbf{P}$. Then,

$$\inf_{v\in\mathbf{P}}\Big\{\sup_{u\in\mathbf{K}}\Big\{\frac{u_2+v_2}{u_1+v_1}\Big\}\Big\} = \sup_{u\in\mathbf{K}}\Big\{\inf_{v\in\mathbf{P}}\Big\{\frac{u_2+v_2}{u_1+v_1}\Big\}\Big\} \qquad \text{(contd.)}$$

(V. A. Ubhaya, "Almost monotone approximation in L_∞", *Journal of Mathematical Analysis and Applications*, 49 (1975), 659–679).

91. A metric on $\mathbf{R}^n$ is a real valued function $d(x, y)$ defined over ordered pairs of points in $\mathbf{R}^n$ satisfying the following properties.

$$\begin{aligned} d(x,y) &\geqq 0, \quad \text{for all } x, y \in \mathbf{R}^n \\ d(x,y) &= 0, \quad \text{iff } x = y \\ d(x,y) &> 0, \quad \text{iff } x \neq y \\ d(x,y) &= d(y,x), \quad \text{for all } x, y \in \mathbf{R}^n \\ d(x,y) + d(y,z) &\geqq d(x,z), \quad \text{for all } x, y, z \in \mathbf{R}^n. \end{aligned}$$

Let $d(x, y)$ be a metric on $\mathbf{R}^n$ and $\mathbf{F}$ be a nonempty subset of $\mathbf{R}^n$. For each x in $\mathbf{R}^n$, let $f(x)$ denote the minimum distance between x and $\mathbf{F}$, viz.,

$$f(x) = \inf\{d(x,u) : u \in \mathbf{F}\}.$$

Show that

$$|f(x) - f(y)| \leqq d(x,y)$$

for all x, y in $\mathbf{R}^n$. Thus f is nonexpansive.

92. Let

$$d'(f,g) = \max\{w_i|f_i - g_i| : 1 \leqq i \leqq n\},$$

denote the distance between two vectors $f = (f_1, f_2, \ldots, f_n)$ and $g = (g_1, g_2, \ldots, g_n)$, where $w = (w_1, w_2, \ldots, w_n) > 0$ is a given weight vector. A vector g is called isotonic if $g_i \leqq g_{i+1}$, $1 \leqq i < n$. Given a vector f, the problem is to find an isotonic vector g which minimizes $d'(f, g)$. Such a g, called an optimal vector, is not unique in general. Denote the minimum of $d'(f, g)$ over isotonic vectors g, for given f, by Δ.

Define the following quantities:

$$\begin{aligned} \theta &= \max\Big\{\frac{w_i w_j}{w_i + w_j}(f_i - f_j) : 1 \leqq i \leqq j \leqq n\Big\}, \\ \underline{g}_i &= \max\{f_j - \theta/w_j : 1 \leqq j \leqq i\}, 1 \leqq i \leqq n, \\ \overline{g}_i &= \min\{f_j + \theta/w_j : 1 \leqq j \leqq n\}, 1 \leqq i \leqq n. \end{aligned}$$

Show the following: (a) Duality: $\Delta = \theta$, (b) Optimality: $\underline{g}$ and $\overline{g}$ are optimal vectors with $\underline{g} \leqq \overline{g}$. Furthermore, an isotonic g is an optimal vector iff $\underline{g} \leqq g \leqq \overline{g}$. (V. A. Ubhaya, "Isotone optimization, I, II", *Journal of Approximation Theory*, 12 (1974), 146–159 and 315–331).

93. Consider Exercise 92 with $w_i = 1$ for all i and define

$$d(f,g) = \max\{|f_i - g_i| : 1 \leqq i \leqq n\}.$$

$$\underline{h}_i = \max\{f_j : 1 \leqq j \leqq i\},\ 1 \leqq i \leqq n,$$
$$\overline{h}_i = \min\{f_j : i \leqq j \leqq n\},\ 1 \leqq i \leqq n.$$

Show the following:

$$\theta = \max\{(\underline{h}_i - f_i) : 1 \leqq i \leqq n\} = \max\{(f_i - \overline{h}_i) : 1 \leqq i \leqq n\}$$

and

$$\underline{g}_i = \underline{h}_i - \theta,\ \overline{g}_i = \overline{h}_i + \theta.$$

Construct an $O(n)$ algorithm for computing optimal vectors $\underline{g}$ and $\overline{g}$.

94. Let $d(f, g)$, as defined in Exercise 93, denote the distance between two vectors f and g. A vector g is called **convex** if it satisfies

$$g_{i-1} - 2g_i + g_{i+1} \geqq 0,\ 1 < i < n,$$

or more generally,

$$a_{i-1}g_{i-1} - (a_{i-1} + a_i)g_i + a_i g_{i+1} \geqq 0,\ 1 < i < n,$$

where a_i, $1 \leqq i < n$, are given positive numbers. Given a vector f, the problem is to find a convex vector g, called an optimal vector, which minimizes $d(f, g)$. Let Δ denote the minimum of $d(f, g)$ over convex vectors g, for given f.

The greatest convex minorant $\overline{h} = (\overline{h}_1, \overline{h}_2, \ldots, \overline{h}_n)$ of f is the largest convex vector (i. e. satisfying the above condition) which does not exceed f. (See Exercise 85). Show the following: $\Delta = (1/2)d(f, \overline{h})$ and $\overline{g} = \overline{h} + e\Delta$ is the maximal optimal vector, i. e., for all optimal vectors g it is true that $\overline{g} \geqq g$. Construct an $O(n)$ algorithm for computing $\overline{h}$ and then $\overline{g}$. (V. A. Ubhaya, "An $O(n)$ algorithm for discrete n-point convex approximation with applications to continuous case", *Journal of Mathematical Analysis and Applications*, 72 (1979), 338–354.)

95. In connection with Exercise 94 consider the following LP.

$$\begin{array}{ll} \text{minimize} & \sum_{i=1}^{n} x_i \\ \text{subject to} & -x_{i-1} + 2x_i - x_{i+1} \geqq -f_{i-1} + 2f_i - f_{i+1},\ 1 < i < n \\ & x_i \geqq 0,\ 1 \leqq i \leqq n. \end{array}$$

Show that the LP has a unique optimal solution x^* and the quantities defined in Exercise 94 for the first convexity constraint are given by

$$\Delta = (1/2)\max\{x_i^* : 1 \leqq i \leqq n\},$$
$$\overline{g}_i = \Delta - x_i^* + f_i,\ 1 \leqq i \leqq n.$$

Devise a special pivoting strategy in conjunction with the Dual Simplex Algorithm of linear programming to solve the above LP in $O(n)$ computing time. (V. A. Ubhaya,

"Linear time algorithms for convex and monotone approximation", *Computers and Mathematics with Applications*, An International Journal, 9 (1983), 633–643.)

96. A vector $g = (g_1, g_2, \ldots, g_n)$ is called **quasiconvex** if

$$g_j \leqq \max\{g_p, g_q\},$$

for all j with $p \leqq j \leqq q$ and for all $1 \leqq p \leqq q \leqq n$. Show that g is quasiconvex iff there exists $1 \leqq r \leqq n$ such that $g_i \geqq g_{i+1}$ for $1 \leqq i < r$ and $g_i \leqq g_{i+1}$ for $r \leqq i < n$. Show that the set of all quasiconvex vectors is a closed nonconvex cone, but the set of all isotonic or convex vectors is a closed convex cone.

Let $d(f, g)$ be as defined in Exercise 93. Given a vector f, consider the problem of finding a quasiconvex vector g, called an optimal vector, which minimizes $d(f, g)$. Show that there exist two optimal vectors $\underline{g}$ and $\overline{g}$ with $\underline{g} \leqq \overline{g}$ so that any quasiconvex vector g with $\underline{g} \leqq g \leqq \overline{g}$ is also an optimal vector. Furthermore, $\overline{g}$ is the maximal optimal vector, i. e., for all optimal vectors g it is true that $\overline{g} \geqq g$. Construct an $O(n)$ algorithm to compute $\underline{g}$ and $\overline{g}$. (V. A. Ubhaya, "Quasi-convex optimization", *Journal of Mathematical Analysis and Applications*, 116 (1986), 439–449.)

97. Exercise 93 to 96 involved finding an isotonic, convex or quasiconvex vector g minimizing $d(f, g)$ given the vector f. Such an optimal vector g is not unique in general. For each f it is of interest to select an optimal vector f' (in each of three cases) so that f' is least sensitive to perturbations in f. Specifically, the following two conditions may be imposed on the selection f' for f.

(i) $d(f', h') \leqq C\ d(f, h)$ holds for all vectors f, h for some least number C. This makes the mapping T defined by $T(f) = f'$ Lipschitzian with constant C.

(ii) The selection f' is such that the number C is smallest among all selections of optimal vectors for f. This makes T optimal.

Thus a mapping T satisfying (i) and (ii) may be called an optimal Lipschitzian selection operator.

Show that optimal Lipschitzian selections are possible for the three problems as shown below. Here $\underline{g}$ and $\overline{g}$ are as defined in Exercises 93, 94 and 96.

(a) Isotonic problem: $T(f) = f' = (1/2)(\underline{g} + \overline{g})$ and $C = 1$.

(b) Convex problem: $T(f) = f' = \overline{g}$ and $C = 2$.

(c) Quasiconvex problem: $T(f) = f' = \overline{g}$ and $C = 2$.

(V. A. Ubhaya, "Lipschitz condition on minimum norm problems on bounded functions", *Journal of Approximation Theory*, 45 (1985), 201–218, also "Optimal Lipschitzian selection operator in quasi-convex optimization", *Journal of Mathematical Analysis and Applications*, to appear).

98. Prove that the functions $\log x$ and $x \log x$ are respectively concave and convex on the interval $0 < x < \infty$. Using this, eastablish the following inequality: if $x > 0$, $y > 0$,

both $x, y \in \mathbf{R}^1$, then

$$\log\frac{x+y}{2} \leqq \frac{x\log x + y\log y}{x+y} \leqq \log\frac{x^{\mathbf{2}}+y^{\mathbf{2}}}{x+y}.$$

---

99. (i) Let $\theta(x)$, $h_i(x)$, $i = 1$ to m be continuously differentiable real valued functions defined over $\mathbf{R}^n$. Consider the nonlinear program.

$$\begin{array}{ll} \text{minimize} & \theta(x) \\ \text{subject to} & h_i(x) = 0,\ i = 1 \text{ to } m. \end{array}$$

Prove that if $\hat{x}$ is a feasible solution to this nonlinear program which is a local minimum for this NLP, then the set of vectors $\{\nabla\theta(\hat{x}); \nabla h_i(\hat{x}),\ i = 1 \text{ to } m\}$ must be a linearly dependent set.

(ii) Consider the following NLP

$$\begin{array}{ll} \text{minimize} & \theta(x) \\ \text{subject to} & h_i(x) = 0,\ i = 1 \text{ to } m \\ & g_p(x) \geqq 0,\ p = 1 \text{ to } t \end{array}$$

where $\theta(x)$, $h_i(x)$, $g_p(x)$ are all continuously differentiable real valued functions defined over $\mathbf{R}^n$. Let $\overline{x}$ be a feasible solution to this NLP. Define $\mathbf{P}(\overline{x}) = \{p : p = 1 \text{ to } t, g_p(\overline{x}) = 0\}$. If $\overline{x}$ is a local minimum for this NLP, prove that the set of vectors

$$\{\nabla\theta(\overline{x})\} \cup \{\nabla h_i(\overline{x}) : i = 1 \text{ to } m\} \cup \{\nabla g_p(\overline{x}) : p \in \mathbf{P}(\overline{x})\}$$

must be a linearly dependent set. In addition, prove that there must exist a linear dependence relation for this set of vectors of the form

$$\delta_0\nabla\theta(\overline{x}) - \sum_{i=1}^{m}\mu_i\nabla h_i(\overline{x}) - \sum_{p\in\mathbf{P}(\overline{x})}\pi_p g_p(\overline{x}) = 0$$

where $(\delta_0,\ \mu_i \text{ for } i = 1 \text{ to } m;\ \pi_p \text{ to } p \in \mathbf{P}(\overline{x})) \neq 0$ and $(\delta_0, \pi_p : p \in \mathbf{P}(\overline{x})) \geqq 0$.

100. Consider the following general QP

$$\begin{array}{lrl} \text{minimize} & Q(x) & = cx + (1/2)x^T Dx \\ \text{subject to} & Ax & \geqq b \\ & x & \geqq 0. \end{array}$$

Define the following:

$\mathbf{K}$ = Set of feasible solutions of this problem.

$\mathbf{L}$ = Set of all local minima for this problem.

$\mathbf{G}$ = Set of all global minima for this problem.

If **K** is bounded, prove that each of the sets **L** and **G**, is a union of a finite number of convex polyhedra. Is this result also true when **K** is not bounded?

101. Maximum Area Hexagon of Diameter One: A problem which has long intrigued mathematicians if finding the maximum area convex polygon in $\mathbf{R}^2$ with an even number of sides, and an upper bound on its diameter. The diameter of a convex polygon is defined to be the maximum distance between any pair of points in it. When the number of sides is odd, the regular polygon has the maximum area; but this may not be true when the number of sides is even.

Consider the special case of this problem, of finding the maximum area hexagon of diameter one. Clearly, without any loss of generality, one can assume that two of the vertices of the hexagon are $(0, 0)$ and $(0, x_1)$; and that the other vertices have coordinates and positions as entered in the following figure,

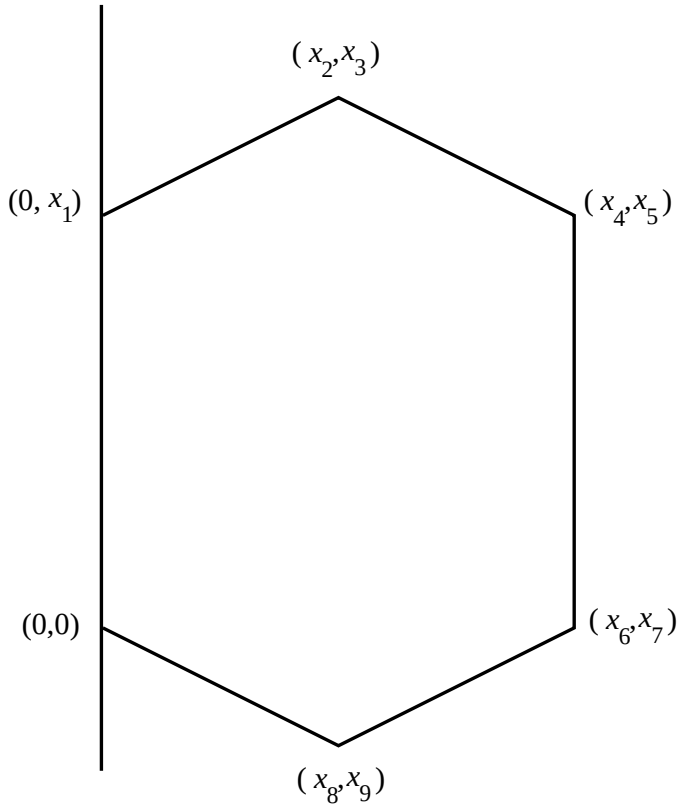


Figure 23

where x_2, x_4, x_6, x_8 are all $\geqq 0$. Formulate the problem of finding the maximum area hexagon of diameter one, as a nonlinear program in terms of the variables x_1 to x_9. Check whether your model is a convex or nonconvex NLP. Write down the necessary optimality conditions for your problem. Solve it on a computer using one of the algorithms discussed in this text.

102. Let $g_i(x)$ be a differentiable convex function defined on $\mathbf{R}^n$ for $i = 1$ to m. Let $\overline{x}$ be a feasible solution of the system

$$g_i(x) \leqq 0, \ \ i = 1 \text{ to } m$$

and let $\mathbf{J}(\overline{x}) = \{i : g_i(\overline{x}) = 0\}$. Prove that the system: $g_i(x) < 0$, $i = 1$ to m, has a feasible solution iff the objective value in the following LP, in which the variables are λ, $d = (d_1, \ldots, d_n)^T$, is unbounded above.

$$\begin{array}{ll} \text{minimize} & \lambda \\ \text{subject to} & \nabla g_i(\overline{x})d + \lambda \leqq 0, \ \ i \in \mathbf{J}(\overline{x}). \end{array}$$

103. Let $g_i(x)$ be a differentiable convex function defined on $\mathbf{R}^n$ for $i = 1$ to m. Let $\overline{x}$ be a feasible solution of the system

$$g_i(x) \leqq 0, \ \ i = 1 \text{ to } m.$$

Prove that the system: $g_i(x) < 0$, $i = 1$ to m, has a feasible solution, iff the following system has no feasible solution $\pi = (\pi_1, \ldots, \pi_m)$.

$$\sum_{i=1}^{m} \pi_i \nabla g_i(\overline{x}) = 0$$

$$\pi \geq 0.$$

104. Let $A(m \times n)$, $B(m \times p)$, $a(1 \times n)$, $b(1 \times p)$ be given matrices. Prove that exactly one of the following two systems (I), (II) has a feasible solution, and the other is infeasible.

(I)	(II)
$Ax + By = 0$	$\pi A \leqq a$
$ax + by < 0$	$\pi B = b$
$x \geqq 0$	

105. Let $A(m \times n)$, $B(m \times p)$, $a(1 \times n)$, $b(1 \times p)$ be given matrices where b is in the linear hull of the row vectors of B. Prove that exactly one of the following systems (I), (II) has a feasible solution, and the other is infeasible.

(I)	(II)
$Ax + By = 0$	$\pi A < a$
$ax + by \leqq 0$	$\pi B = b$
$x \geq 0$	

106. Consider the following NLP involving the vectors of decision variables $x \in \mathbf{R}^{n_1}$, and $y \in \mathbf{R}^{n_2}$

$$\begin{array}{lll} \text{minimize} & h(x,y) = \theta(x) + cy & \\ \text{subject to} & g(x) + ay = b^1 & (m_1 \text{ constraints}) \\ & Bx + Dy = b^2 & (m_2 \text{ constraints}) \\ & l^1 \leqq x \leqq u^1 & \\ & l^2 \leqq y \leqq u^2 & \end{array}$$

where $\theta(x)$, $g(x)$ are continuously differentiable functions. Given $\overline{x} \in \mathbf{R}^{n_1}$, $\overline{y} \in \mathbf{R}^{n_2}$ satisfying the bound constraints in the NLP, define the following LP, which comes from a linearization of the NLP around $(\overline{x}, \overline{y})$.

$$\begin{array}{ll} \text{minimize} & \nabla\theta(\overline{x})d + cy \\ \text{subject to} & \nabla g(\overline{x})d + Ay = b^1 - g(\overline{x}) \\ & Bd + Dy = b^2 - B\overline{x} \end{array}$$

$$\max\{l^1_j - \overline{x}_j, s_j\} \leqq d_j \leqq \min\{u^1_j - \overline{x}_j, s_j\}, \ \ j = 1 \text{ to } n_1$$
$$l^2 \leqq y \leqq u^2$$

where $d = x - \overline{x}$, $s = (s_j) \in \mathbf{R}^{n_1}$, $s > 0$ is a vector of small positive numbers used to bound d in the LP to keep the linearization reasonably accurate. Prove the following

i) If $(\overline{x}, \overline{y})$ is feasible to the NLP, $(d, y) = (0, \overline{y})$ is feasible to the above LP for any $s > 0$.

ii) If the constraint matrix of the LP has full row rank, and $(\overline{x}, \overline{y})$ is a feasible solution of the NLP, then $(0, \overline{y})$ is an optimum solution of the above LP iff $(\overline{x}, \overline{y})$ is a KKT point for the NLP.

iii) Let $(\overline{x}, \overline{y})$ be a feasible solution for the NLP, and suppose $(0, \overline{y})$ is not an optimum solution for the above LP. If (d^0, y^0) is an optimum solution for the LP, then $\nabla\theta(x)d^0 + c(y^0 - \overline{y}) < 0$, that is, $(d^0, y^0 - \overline{y})$ is a descent direction for the NLP at the point $(\overline{x}, \overline{y})$. (See F. Palacios-Gomez, L. Lasdon and M. Engquist, "Nonlinear optimization by successive linear programming", *Management Science*, 28, 10 (October 1982) 1106–1120.)

107. Consider the following NLP

$$\begin{array}{ll} \text{minimize} & Q(x) = cx + \frac{1}{2}x^T Dx \\ \text{subject to} & \|x\| \leqq \delta \end{array}$$

where D is a PD symmetric matrix of order n and $\delta > 0$. Write down the KKT optimality conditions for this problem. Prove that the optimum solution of this problem is $x(\lambda) = -(D + \lambda l)^{-1}c^T$ for the unique $\lambda \geqq 0$ such that $\|x(\lambda)\| = \delta$; unless $\|x(0)\| \leqq \delta$, in which case, $x(0)$ is the optimum solution.

108. Let $f(x) = f_1(x), \ldots, f_n(x))^T$ where each $f_j(x)$ is a continuous function defined ofer $\mathbf{R}^n$. Let $\mathbf{K}$ be a closed convex cone in $\mathbf{R}^n$. Define the polar cone of $\mathbf{K}$ to be $\mathbf{K}^* = \{y : y \in \mathbf{R}^n, y^T x \geqq 0 \text{ for all } x \in \mathbf{K}\}$. (For example, if $\mathbf{K}$ is the nonnegative orthant, $\mathbf{K}^*$ is tha same. Let $\mathbf{J} \subset \{1, \ldots, n\}$. If $\mathbf{K}$ is the orthant $\{x : x = (x_j) \in \mathbf{R}^n, x_j \geqq 0 \text{ for } j \notin \mathbf{J}, x_j \leqq 0 \text{ for } j \in \mathbf{J}\}$, then $\mathbf{K}^*$ is again $\mathbf{K}$ itself.)

The generalized complementary problem corresponding to $f(x)$ and $\mathbf{K}$ is to find x satisfying

$$x \in \mathbf{K},\ f(x) \in \mathbf{K}^*,\ x^T f(x) = 0 \tag{101}$$

using the hypothesis that $\mathbf{K}$ is a closed convex cone, prove that the generalized complementarity problem (101) is equivalent to the variational inequality problem: find $x^* \in \mathbf{K}$ satisfying

$$(x - x^*)^T f(x^*) \geqq 0 \text{ for all } x \in \mathbf{K} \tag{102}$$

(see Karamardian [1.14]).

109. Let $\mathbf{K}$, $\mathbf{K}^*$, $f(x)$ be defined as in the previous Exercise 108. For any $x \in \mathbf{R}^n$ define $P_{\mathbf{K}}(x)$ to be the projection of x into $\mathbf{K}$ (i. e., the nearest point in $\mathbf{K}$ to x, in terms of the usual Euclidean distance). Prove that a solution $x^* \in \mathbf{K}$ to the variational inequality problem (102), can be characterized by the relation

$$x^* = P_{\mathbf{K}}(x^* - \rho f(x^*))$$

where ρ is a positive constant. Using this, show that the generalized complementarity problem (101) can be formulated as the fixed point problem of finding $x \in \mathbf{K}$ satisfying

$$x = g(x) \tag{103}$$

where $g(x) = \lambda P_{\mathbf{K}}(x - \rho f(x)) + (1 - \lambda)x$, with a constant $\rho > 0$ and $0 < \lambda \leqq 1$. Here λ is known as the relaxation factor used after the projection.

Study the application of the successive substitution method for solving (103). This method will begin with a given $x^0 \in \mathbf{K}$, and generate the sequence of points $\{x^r : r = 0, 1, \ldots\}$ using the iteration, $x^{r+1} = g(x^r)$. The iterative methods discussed in Sections 9.3, 9.4, 9.5 are special cases of this general approach. Study the convergence properties of the sequence of points generated under this method (M. Aslam Noor, and K. Inayat Noor, "Iterative methods for variational inequalities and nonlinear programming", *Operations Research Verf.*, 31 (1979) 455–463).

110. Let $\mathbf{K} \subset \mathbf{R}^n$ be convex and let $f(x) = (f_1(x), \ldots, f_n(x))^T$, where each $f_i(x)$ is a continuous real valued function defined over $\mathbf{K}$. Define a point $\overline{x} \in \mathbf{K}$ to be a **critical point** for the pair $(f, \mathbf{K})$ if $y = \overline{x}$ minimizes $(f(\overline{x}))^T y$ over $y \in \mathbf{K}$. Let $\mathbf{\Gamma}(f, \mathbf{K})$ denote the set of all critical points for the pair $(f, \mathbf{K})$.

Let $\mathbf{\Delta}(f, \mathbf{K})$ denote the set of all points $\overline{x} \in \mathbf{K}$ such that $y = \overline{x}$ minimizes $\|y - \overline{x} + f(\overline{x})\|$ over $y \in \mathbf{K}$. Prove that $\mathbf{\Delta}(f, \mathbf{K}) \subset \mathbf{\Gamma}(f, \mathbf{K})$.

Let $\theta(x)$ be a real valued continuously differentiable function defined over $\mathbf{K}$. Consider the NLP

$$\begin{array}{ll} \text{minimize} & \theta(x) \\ \text{over} & x \in \mathbf{K}. \end{array}$$

Prove that every stationary point for this NLP is a critical point for the pair $(\nabla\theta(x), \mathbf{K})$. If $\mathbf{K} = \mathbf{R}^n_+ = \{x : x \in \mathbf{R}^n, x \geqq 0\}$, prove that the problem of finding a critical point for the pair $(f, \mathbf{R}^n_+)$ is equivalent to the nonlinear complementarity problem (NLCP): find $x \in \mathbf{R}^n$ satisfying

$$x \geqq 0,\ f(x) \geqq 0,\ x^T f(x) = 0.$$

Let $d \in \mathbf{R}^n$, $d > 0$ be a given vector. Let $\mathbf{D}(\alpha) = \{x : x \in \mathbf{R}^n, x \geqq 0, \text{ and } d^T x \leqq \alpha\}$, for each $\alpha \geqq 0$. If $\mathbf{K} = \mathbf{D}(\alpha)$ for some $\alpha \geqq 0$, prove that $\overline{x} \in \mathbf{D}(\alpha)$ is a critical point for the pair $(f, \mathbf{D}(\alpha))$, iff there is a $w \in \mathbf{R}^n_+$ and $z_0 \geqq 0$ such that,

$$\begin{aligned} f(\overline{x}) = w - dz_0,\ \overline{x}^T w = 0 \\ z_0(\alpha - d^T\overline{x}) = 0. \end{aligned}$$

Also, prove that if $\overline{x}$ is a critical point of $(f, \mathbf{D}(\alpha))$ and $d^T\overline{x} < \alpha$, then $\overline{x}$ is a critical point of $(f, \mathbf{R}^n_+)$. Conversely if $\overline{x} \in \mathbf{\Gamma}(f, \mathbf{R}^n_+)$ and $d^T\overline{x} \leqq \alpha$, then $\overline{x} \in \mathbf{\Gamma}(f, \mathbf{D}(\alpha))$.

Consider the case where $\mathbf{K}$ is nonempty, compact and convex. In this case, for each $x \in \mathbf{K}$, define $h(x)$ to be the y that minimizes $\|y - x + f(x)\|$ over $y \in \mathbf{K}$. Using $h(x)$ and Brower's fixed point theorem show that $(f, \mathbf{K})$ has a critical point.
(B. C. Eaves [3.20])

111. Let $f(x) = (f_1(x), \ldots, f_n(x))^T$ where each $f_i(x)$ is a continuous real valued function defined over $\mathbf{R}^n$. Consider the NLCP: find x satisfying

$$x \geqq 0,\ f(x) \geqq 0,\ x^T f(x) = 0.$$

For each $x \geqq 0$, define $h(x)$ to be the y that minimizes $\|y - x + f(x)\|$ over $y \geqq 0$. If $h(x) = (h_i(x))$, show that

$$h_i(x) = \begin{cases} 0, & \text{if } f_i(x) - x_i \geqq 0 \\ x_i - f_i(x), & \text{if } f_i(x) - x_i \leqq 0 \end{cases}$$

Prove that the following conditions are equivalent

i) x solves the NLCP given above,
ii) $x \geqq 0$ and $h(x) = x$,
iii) $x \geqq 0$ and $x \in \mathbf{\Gamma}(f, \mathbf{R}^n_+)$

where $\mathbf{\Gamma}(f, \mathbf{R}^n_+)$ is defined in the previous Exercise 110.

Suppose there is a compact convex set $\mathbf{S} \subset \mathbf{R}^n$ such that for each $x \in \mathbf{R}^n_+ \setminus \mathbf{S}$, there is a $y \in \mathbf{S}$ satisfying $(y - x)^T f(x) < 0$. Under this condition, prove that every fixed point of $h(x)$ lies in the set $\mathbf{S}$.
(R. Saigal and C. Simon [3.67], B. C. Eaves [3.20])

112. (Research Problem): In Section 11.4.1, subsection 5, we described a special direct procedure for obtaining a true optimum solution for an LP, from a given near optimum solution for it. Consider the QP (1.11). Assuming that D is PSD, and that a near optimum feasible solution, $\overline{x}$, is given for it, develop a special direct procedure to obtain a true optimum solution for the QP (1.11), from $\overline{x}$.

113. An economic model leads to the following optimization problem. The decision variables in this problem are $x \in \mathbf{R}^n$, $y \in \mathbf{R}^n$ and $z \in \mathbf{R}^p$. The problem is

$$\begin{array}{ll} \text{minimize} & cx + dy + az \\ \text{subject to} & A_1x + A_2y + A_3z = b \\ & x, y, z \geqq 0 \\ & \text{and } x^T y = 0 \end{array}$$

where A_1, A_2, A_3 are given matrices of order $m \times n$, $m \times n$, $m \times p$ respectively, and c, d, a, b are given vectors of appropriate dimensions. Formulate this as a mixed integer linear programming problem.

114. Let $x \in \mathbf{R}^1$. Define $\mathbf{F}(x) = \{x^3 + 3x^2 - 9x - 24\}$. Compute a Kakutani fixed point of $\mathbf{F}(x)$ using the algorithm discussed in Section 2.7.8.

7. References

A1 F. H. Clarke, "Generalized Gradients and Applications", *Transactions of the American Mathematical Society*, 205 (1975) 247–262.

A2 R. W. Cottle, F. Giannessi and J. L. Lions, Editors, "Variational Inequalities and Complementarity Problems", *Proceedings of International Summer School held in Sicily*, John Wiley & Sons, Inc., 1978.

A3 J. E. Dennis, Jr. and J. J. More', "Quasi-Newton Methods, Motivation and Theory", *SIAM Review*, 19 (1977) 46–89.

A4 W. Karush, "Minima of Functions of Several Variables with Inequalities as Side Conditions", M. S. Dissertation, Department of Mathematics, University of Chicago, Chicago, Illinois (1939).

A5 H. W. Kuhn, "Solvability and Consistency for Linear Equations and Inequalities", *American Mathematical Monthly*, 63 (1956) 217–232.

A6 H. W. Kuhn and A. W. Tucker, "Nonlinear Programming", *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, J. Neyman (editor), University of California Press, Berkeley, California (1951) 481–492.

A7 H. W. Kuhn, "Nonlinear Programming: A Historical View", in: R. W. Cottle and C. E. Lemke (editors), *Nonlinear Programming: Proceedings of the Joint SIAM-AMS Symposium on Applied Mathematics*, American Mathematical Society, Providence, Rhode Island (1976).

A8 D. G. Luenberger, *Introduction to Linear and Nonlinear Programming*, Second ed., Addison-Wesley, Reading, USA (1984).

A9 A. Majthay, "Optimality Conditions for Quadratic Programming", *Mathematical Programming*, 1 (1971) 359–365.

A10 O. L. Mangasarian, *Nonlinear Programming*, McGraw-Hill, New York (1969).

A11 O. L. Mangasarian and T. H. Shiau, "Lipschitz Continuity of Solutions of Linear Inequalities, Programs and Complementarity Problems", Computer Science Technical Report # 599, University of Wisconsin, Madison, June 1985.

A12 G. P. McCormick, "Second Order Conditions for Constrained Minima", *SIAM Journal on Applied Mathematics*, 15 (1967) 641–652.

A13 N. Z. Shor, *Minimization Methods for Nondifferentiable Functions*, Springer Series in Computational Mathematics 3, Springer-Verlag, New York, 1985.

A14 K. Sydsaeter, "Letter to the Editor on Some Frequently Occuring Errors in the Economic Literature Concerning Problems of Maxima and Minima", *Journal of Economic Theory*, 9 (1974) 464–466.

A15 L. Watson, "Solving the Nonlinear Complementarity Problem by a Homotopy Method", *SIAM Journal on Control and Optimization*, 17 (1979) 36–46.

Index