Interior Point Methods for LP

Katta G. Murty, IOE 510, LP, U. Of Michigan, Ann Arbor, Winter 1997.

Simplex Method - A Boundary Method:
Starting at an extreme point of the feasible set, the simplex method walks along its edges, until it

- either finds an optimum extreme point
- or an unbounded edge along which the objective function diverges

and then it terminates. Since all its action takes place on the boundary of the feasible set, it is labeled as a *boundary method.*

Interior Point Methods:
start with a point in the *(relative) interior* of the feasible set, continue in the interior until they reach a near optimum solution. Each iteration here consists of 2 steps:

Step 1: Determine the *search direction*, i.e., the direction to move at the current interior solution.

Step 2: Determine the *step length* of the move.

There are 2 classes of interior point methods.

Class 1: In these (affine scaling methods, Karmarkar's projective scaling method, etc.) search direction determined by the solution of a modified (approximating) problem constructed around the current interior feasible solution.

Class 2: These methods apply variants of Newton's method (for solving systems of nonlinear eqs.) to the optimality conds. consisting of primal and dual feasibility and complementary slackness conds.

Minimizing a Linear Function Over a Ball or an Ellipsoid

These problems are easy, the answer can be explicitly written down directly. The Class 1 interior point methods use these results.

1. Consider $\min z(x) = cx$ s. to $(x - x^0)^T(x - x^0) \leq \rho^2$.

Case 1: $c = 0$. Every point in the ball is optimal.

Case 2: $c \neq 0$. Optimum solution is $x^0 + \rho(-c^T/||c||)$, obtained by moving from the center x^0, a step length of the radius ρ, in the direction of the negative gradient of the objective function.

2. Consider $\min z(x) = cx$ s. to $(x-x^0)^T D^{-2}(x-x^0) \leq \rho^2$ where $x^0 = (x^0_j)$ is such that $x^0_j \neq 0$ for all j, and $D = \text{diag}(x^0_1, \ldots, x^0_n)$.

Set of feasible solutions of this is an ellipsoid with x^0 as center.

An *affine scaling transformation*

$$y = D^{-1}x$$

converts the ellipsoid into a ball with $e = (1, \ldots, 1)^T$ as center and ρ as radius. Assuming $c \neq 0$ and using this transformation, it can be verified that the opt. sol. of this problem is:

$$x^0 - \rho \frac{D^2 c^T}{||Dc^T||}$$

.

3. Consider $\min z(x) = cx$ s. to $Ax = b$, and $(x - x^0)^T(x - x^0) \leq \rho^2$ where $A_{m \times n}$ has rank m, and x^0 is a point in the affine space $H = \{x : \ Ax = b\}$.

Let $B = \{x : \ (x - x^0)^T(x - x^0) \leq \rho^2\}$ be the ball. Since the center $x^0 \in H$, the intersection $H \cap B$ is another ball in H with radius ρ and center x^0. Optimum sol. of this problem is obtained by following procedure.

- Project c^T = gradient of $z(x)$ into the affine space H. This gives Pc^T where P is the projection matrix corresponding to the affine space H,

$$P = I - A^T(AA^T)^{-1}A$$

- Optimum sol. is obtained by moving a step length of ρ from the center x^0 in the direction of the negative projected gradient, $-Pc^T$. Hence the opt. sol. is $x^0 - \rho \frac{Pc^T}{||Pc^T||}$.

Actually if $Pc^T = 0$, then c is a linear combination of row vwctors of A, in this case every feasible sol. is optimal.

The Primal Affine Scaling Method

We consider the LP in standard form: $\min z = cx$, subject to $Ax = b, x \geq 0$ where $A_{m\times n}$ and rank m. Let K denote the set of feasible slutions of this problem.

A feasible sol. x is said to be an *interior feasible solution*, or *strict feasible solution* if $x > 0$.

Let $\bar{x} = (\bar{x}_j) > 0$ be the current interior feasible sol. Let $\bar{X} = \text{diag}\{\bar{x}_1, \ldots, \bar{x}_n\}$.

The method now looks at the approximating problem obtained by replacing the constraints "$x \geq 0$" in original LP by the constraint

$$x \in E = \{x : (x - \bar{x})^T \bar{X}^{-2}(x - \bar{x}) \leq 1\}$$

So the approximating problem is: $\min z = cx$ s. to $Ax = b$, $x \in E = \{x : (x - \bar{x})^T \bar{X}^{-2}(x - \bar{x}) \leq \rho^2\}$ where $\rho = 1$.

THEOREM: If $0 < \rho \leq 1$, set of feasible solutions of approximating problem is $\subset K$.

The optimum solution of the approximating problem is:

$$x^* = \bar{x} - \rho \frac{\bar{X}P\bar{X}c^T}{||P\bar{X}c^T||}$$

where $P = I - \bar{X}A^T(A\bar{X}^2A^T)^{-1}A\bar{X}$ is the projection matrix.

THEOREM: With $\rho = 1$, if x^* is on the boundary of K, i.e., if $x_j^* = 0$ for some j, then x^* is an opt. sol. of original LP and y^T is an opt. dual sol, where

$$y = (A\bar{X}^2A^T)^{-1}A\bar{X}^2c^T$$

and $s = c^T - A^Ty$ is the dual slack vector at y.

The method essentially consists of starting at an interior feasible sol. $\bar{x}$, moving to x^* with $\rho = 1$ (or moving from $\bar{x}$ in the direction $x^* - \bar{x}$ with a step length that is a certain percentage (typically 95%) of the maximum step length while maintaining feasibility), and repeating the whole process with the new interior feasible solution.

PRIMAL AFFINE SCALING METHOD

INPUT NEEDED: Problem in standard form, and an interior feasible solution. Let α be the step length fraction parameter

($\alpha = 0.95$ typically).

GENERAL STEP: Let $\bar{x} > 0$ be the current interior feasible sol. and $\bar{X} = \text{diag}\{\bar{x}_1, \ldots, \bar{x}_n\}$.

Compute $\bar{y}$ = tentative dual sol. $= (A\bar{X}^2A^T)^{-1}A\bar{X}^2c^T$.

Compute tentative dual slack vector $\bar{s} = c^T - A^T\bar{y}$.

If $\bar{s} \leq 0$, z is unbounded below in original LP, terminate.

Compute opt. sol. of approximating problem

$$x^* = \bar{x} - \frac{\bar{X}^2\bar{s}}{||\bar{X}\bar{s}}$$

If $x_j^* = 0$ for some j, x^* optimal to the LP and $\bar{y}$ opt. to the dual, terminate. Otherwise, compute step length

$$\theta = \min\{\frac{||\bar{X}\bar{s}||}{\bar{x}_j\bar{s}_j} : \quad \text{over } j \text{ s. th. } \bar{s}_j > 0\}$$

Take the next point to be

$$\hat{x} = \bar{x} - 0.95\theta\frac{\bar{X}^2\bar{s}}{||\bar{X}\bar{s}}$$

Verify that $c\hat{x} = c\bar{x} - 0.95\theta||\bar{X}\bar{s}||$. Go to the next step with the new point $\hat{x}$.

Convergence Results

Let $\{x^k\}$, $\{y^k\}$, $\{s^k\}$, be the sequences generated by the affine scaling method. Assume that $A_{m\times n}$ has rank m, the LP has an optimum solution, and that c is not in the linear hull of the set of row vwctors of A.

THEOREM: The primal objective value cx^k is strictly monotone decreasing.

THEOREM: The sequence $\{x^k\}$ converges to an optimum solution, x^* of the LP.

THEOREM: If the LP is nondegenerate, all three sequences converge to x^*, y^*, s^* say, where x^* is optimum to primal, and y^* is optimal to the dual, and s^* is the dual slack vector corresponding to y^*.

If the LP is degenerate, the dual sequence may not converge, counterexamples are known. However, if $\alpha \leq 2/3$, then the dual sequence has been shown to converge to an optimum dual solution even under degeneracy.

Primal-Dual Path Following Interior Point Methods

We consider the LP in standard form: min $z = cx$, subject to $Ax = b, x \geq 0$ where $A_{m \times n}$ and rank m.

Let $y = (y_1, \ldots, y_m)^T$ be the column vector of dual variables, and $s = (s_1, \ldots, s_n)^T$ the column vector of slack variables. Let $e = (1, \ldots, 1)^T \in R^n$.

Define $X = \text{diag}(x_1, \ldots, x_n)$, $S = \text{diag}(s_1, \ldots, s_n)$.

From optimality conds., solving the LP is equivalent to finding a solution (x, y, s) satisfying $(x, s) \geq 0$, to the system of $2n + m$ equations in $2n + m$ unknowns:

$$F(x, y, s) = \begin{bmatrix} A^T y + s - c \\ Ax - b \\ XSe \end{bmatrix} = 0$$

Let

$$\mathcal{F} = \{(x, y, s) : Ax = b, A^T y + s = c, (x, s) \geq 0\}$$

$$\mathcal{F}^0 = \{(x, y, s) : Ax = b, A^T y + s = c, (x, s) > 0\}$$

The primal-dual interior point methods generate iterates $(x^k, y^k, s^k) \in$

$\mathcal{F}^0$ based on modified Newton methods for solving the square system of equations $F(x, y, s) = 0$.

The Central Path

This path, $\mathcal{C}$ is an arc in $\mathcal{F}^0$ parametrized by a positive parameter $\tau > 0$. For each $\tau > 0$, the point $(x^\tau, y^\tau, s^\tau) \in \mathcal{C}$ satisfies: $(x^\tau, s^\tau) > 0$ and

$$\begin{aligned} A^T y^\tau + s^\tau &= c^T \\ A x^\tau &= b \\ x_j^\tau s_j^\tau &= \tau, \quad j = 1, \ldots, n \end{aligned}$$

If $\tau = 0$, the above eqs. define the optimality conditions for the LP. For each $\tau > 0$, the solution (x^τ, y^τ, s^τ) is unique, and as $\tau \downarrow 0$ the central path converges to the center of the optimum face.

Starting at an interior feasible solution (x, y, s) (i.e., a feasible solution with $(x, s) > 0$, these methods take steps in modified Newton directions towards points on $\mathcal{C}$. For any interior feasible solution, define:

Centering parameter $\sigma \in [0, 1]$

Duality measure $\mu = x^T s / n$

The direction for the move is: $(\triangle x, \triangle y, \triangle s)$ obtained as the solution of the system of linear equations

$$\begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{bmatrix}, \begin{bmatrix} \triangle x \\ \triangle y \\ \triangle s \end{bmatrix} = p$$

where $p = (0, 0, -XSe + \sigma\mu e)$.

If $\sigma = 1$, the direction obtained will be a centering direction, which is a Newton direction towards the point (x^μ, y^μ, s^μ) on $\mathcal{C}$ at which all pairwise products $x_j s_j$ are $= \mu$. Many algorithms choose σ from open interval $(0, 1)$ to trade off between twin goals of reducing μ and improving centrality.

General Primal-Dual Path Following Method

INPUT NEEDED: Problem in standard form, initial interior feasible solution (x, y, s), i.e., one with $(x, s) > 0$.

When (x, y, s) is current interior feasible solution, compute direction $(\triangle x, \triangle y, \triangle s)$ as above. Take the next point to be $(\hat{x}, \hat{y}, \hat{s}) = (x, y, s) + \alpha(\triangle x, \triangle y, \triangle s)$, where α is step length selected so that $(\hat{x}, \hat{s})$ remains > 0.