Chapter 7

NEAREST POINT PROBLEMS ON SIMPLICIAL CONES

Let $\mathbf{\Gamma} = \{B_{.1}, \dots, B_{.n}\}$ 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 $B = (B_{.1} \vdots \dots \vdots B_{.n})$. 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 $\alpha = (\alpha_1, \dots, \alpha_n)^T \geqq 0$. This is the quadratic program: Minimize $-b^T B\alpha + \frac{1}{2}\alpha^T(B^T B)\alpha$, subject to $\alpha = (\alpha_1, \dots, \alpha_n)^T \geqq 0$. The solution of this can be obtained by solving the following LCP :

$$\begin{aligned} &u - (B^T B)\alpha = -B^T b \\ &u \geqq 0, \quad \alpha \geqq 0 \\ &u^T \alpha = 0 \end{aligned}$$

where $u = (u_1, \dots, u_n)^T$ is a column vector of variables in $\mathbf{R}^n$. Let $D = B^T B$. 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

$$\begin{aligned} &w - Mz = q \\ &w \geqq 0, \; z \geqq 0 \\ &w^T z = 0 \end{aligned} \tag{7.1}$$

where M is a positive definite symmetric matrix of order n. Let F be a nonsingular matrix such that $F^T F = M$ (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 $[F; -(F^{-1})^T q]$. Conversely if z^* is the optimum combination vector for the nearest point problem $[F; -(F^{-1})^T q]$, 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 $B^T B = M$ (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 $B^T B = M$ is available, and $b = -(B^{-1})^T q$, and $\mathbf{\Gamma} = \{B_{.1}, \dots, B_{.n}\}$. 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 $B^T B = M$ and $b = -(B^{-1})^T q$. 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 $\mathbf{S} = \{B_{.j_1}, \ldots, B_{.j_r}\} \subset \mathbf{\Gamma}$. 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 $\mathbf{H}(\mathbf{S}) =$ Pos$(\mathbf{S}) = \{0\}$. For any $\mathbf{S} \subset \mathbf{\Gamma}$, Pos$(\mathbf{S})$ is a face of Pos$(\mathbf{\Gamma})$. 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* $\mathbf{S} \subset \mathbf{\Gamma}$ *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, $b(\mathbf{S}) = B(\mathbf{S})\overline{\gamma} = B(\mathbf{S})\big(B(\mathbf{S})^T B(\mathbf{S})\big)^{-1} (B(\mathbf{S}))^T b$.

$\square$

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 Pos(B).

Theorem 7.2 *For $\mathbf{S} \subset \mathbf{\Gamma}$, 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 $||b - x||^\mathbf{2} = ||b - b(\mathbf{S})||^\mathbf{2} + ||b(\mathbf{S}) - x||^\mathbf{2}$. Since Pos$(\mathbf{S}) \subset \mathbf{H}(\mathbf{S})$, this equality obviously holds for all $x \in$ Pos$(\mathbf{S})$. Hence the theorem follows.

$\square$

Theorem 7.3 *Let $\mathbf{S} \subset \mathbf{\Gamma}$, $\mathbf{S} \neq \emptyset$. 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 Pos$(\mathbf{S})$ if and only if $b(\mathbf{S}) = B(\mathbf{S})\overline{\gamma}$, where $\overline{\gamma} > 0$. As long as $b(\mathbf{S}) \in$ 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 $b(\mathbf{S}) \notin$ 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 Pos$(\mathbf{S})$.

$\square$

Definition — Projection Face
Let $\mathbf{S} \subset \mathbf{\Gamma}$. Pos$(\mathbf{S})$ is a face of Pos$(\mathbf{\Gamma})$ of dimension $|\mathbf{S}|$. Pos$(\mathbf{S})$ is said to be a **Projection face** of Pos$(\mathbf{\Gamma})$, if $b(\mathbf{S}) \in$ 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 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 Pos$\{B_{.1}, B_{.3}\}$ is not a projection face.

On the other hand, consider the face 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 Pos$\{B_{.2}, B_{.3}\}$. So Pos$\{B_{.2}, B_{.3}\}$ is a projection face of Pos(B).

Theorem 7.4 *Let $x^* = B\alpha^*$ be the optimum solution of $[\mathbf{\Gamma}; b]$. Let $\mathbf{I}(\mathbf{S}) = \{j_1, \ldots, j_r\} = \{j : j$ 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$(\mathbf{\Gamma})$.*

Proof. Obviously $x^* \in$ Pos$(\mathbf{S})$. Since x^* is the nearest point in Pos$(\mathbf{\Gamma})$ to b, and since Pos$(\mathbf{S}) \subset$ Pos$(\mathbf{\Gamma})$, clearly x^* is the nearest point in Pos$(\mathbf{S})$ to b. However, by the definition of $\mathbf{S}$, x^* is in the relative interior of Pos$(\mathbf{S})$. Hence, by Theorem 7.3, x^* must be the projection of b in $\mathbf{H}(\mathbf{S})$. Since $x^* \in$ Pos$(\mathbf{S})$, this implies that Pos$(\mathbf{S})$ is a projection face of Pos$(\mathbf{\Gamma})$.

$\square$

Exercises

7.1 Prove that the problem of finding the nearest point in the face $\text{Pos}(\mathbf{S})$ of $\text{Pos}(\mathbf{\Gamma})$ 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 $\text{Pos}(\mathbf{S})$ to b or $b(\mathbf{S})$; and conversely. Also prove that the face $\text{Pos}(\mathbf{S})$ of $\text{Pos}(\mathbf{\Gamma})$ 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 $\text{Pos}(\mathbf{S})$ is a projection face of $\text{Pos}(\mathbf{\Gamma})$, prove that $b(\mathbf{S})$ is the nearest point in $\text{Pos}(\mathbf{\Gamma})$ 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 $\text{Pos}(\mathbf{\Gamma})$ 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

$$\mathbf{N}(\bar{x}) = \{j : j \text{ such that } B^T_{.j}(b - \bar{x}) > 0\} .$$

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.

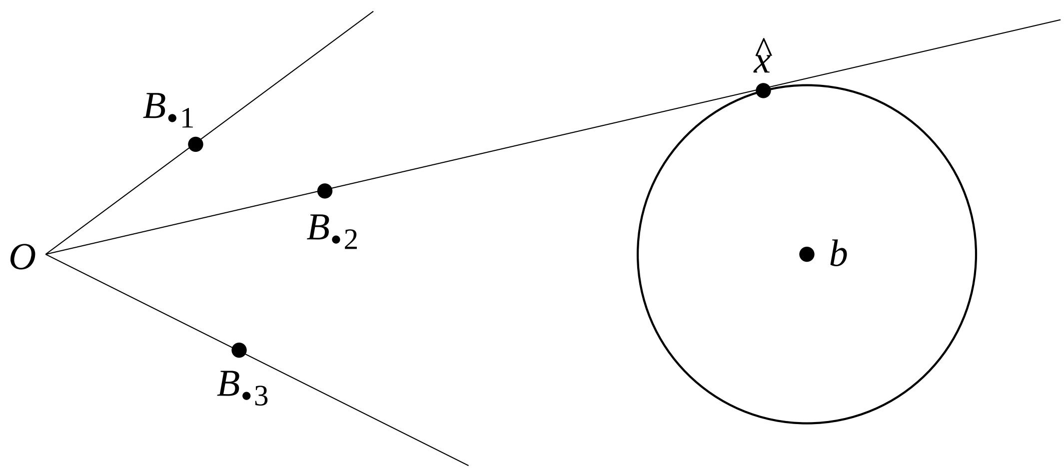

Figure 7.1

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 $b^T B_{\cdot j} = 0, -1, -1$ respectively for $j = 1, 2, 3$. So, the nearest point on the ray of $B_{\cdot j}$ is $V^j = 0$ for all $j = 1, 2, 3$. Hence in this case 0 is the nearest point in $\text{Pos}(B)$ to b.

Thus 0 is the nearest point to b in $\text{Pos}(B)$ iff $b^T B_{\cdot j} \leqq 0$ for all $j = 1$ to n. So, in the sequel, we will assume that $b^T B_{\cdot j} > 0$ 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*

$$\begin{aligned} &0 \in \mathbf{T}(b; \bar{x}) \text{ and} \\ &(b - \bar{x})^T B_{\cdot j} \leqq 0, \quad \text{for all } j = 1 \text{ to } n. \end{aligned} \tag{7.2}$$

Proof. Suppose $\bar{x}$ is the nearest point in $\text{Pos}(\mathbf{\Gamma})$ 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 $\text{Pos}(\mathbf{\Gamma})$. So $(b - \bar{x})^T B_{\cdot j} \leqq 0$ for all $j = 1$ to n.

Conversely suppose $\bar{x} \in \text{Pos}(\mathbf{\Gamma})$ 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 $\text{Pos}(\mathbf{\Gamma})$. So, under these conditions, $\bar{x}$ is the nearest point in $\text{Pos}(\mathbf{\Gamma})$ 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 $(b - \bar{x})^T B_{\cdot 3} = (\frac{3}{2}) > 0$, and hence $\bar{x}$ is not the nearest point in $\text{Pos}(B)$ to b.

Let $\hat{x} = (0, -4, 7)^T$, the orthogonal projection of b in the linear hull of $\{B_{\cdot 2}, B_{\cdot 3}\}$, which is the nearest point in the face $\text{Pos}\{B_{\cdot 2}, B_{\cdot 3}\}$ of $\text{Pos}(B)$ 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 $\{B_{\cdot 2}, B_{\cdot 3}\}$ itself) and hence the origin 0. Also, it can be verified that $(b - \hat{x})^T B_{\cdot j} = -3, 0, 0 \leqq 0$, for $j = 1, 2, 3$. So $\mathbf{N}(\hat{x}) = \emptyset$ and $\hat{x}$ is the nearest point in $\text{Pos}(B)$ to b in this example. See Figure 7.2.

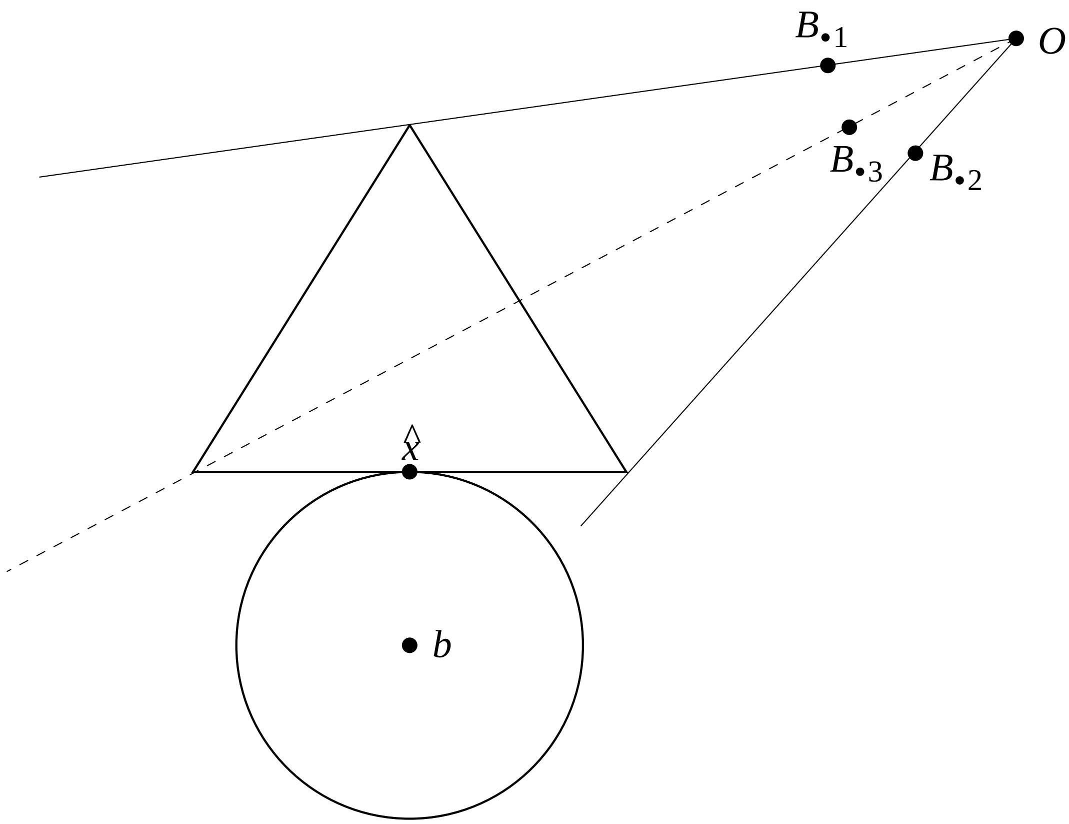

Figure 7.2 $\hat{x}$ is the nearest point in Pos(B) to b.

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	$\dots$	w_n	z_1	z_2	$\dots$	z_n	
$-\overline{m}_{11}$	0	$\dots$	0	1	$-\overline{m}_{12}$	$\dots$	$-\overline{m}_{1n}$	$\bar{q}_1$
$-\overline{m}_{21}$	1	$\dots$	0	0	$-\overline{m}_{22}$	$\dots$	$-\overline{m}_{2n}$	$\bar{q}_2$
$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
$-\overline{m}_{n1}$	0	$\dots$	1	0	$-\overline{m}_{n2}$	$\dots$	$-\overline{m}_{nn}$	$\bar{q}_n$

Let $\overline{\mathcal{M}} = (\overline{m}_{ij} : 2 \leqq i,\ j \leqq n)$ be the matrix of order $n-1$, and $\bar{q} = (\bar{q}_2, \dots, \bar{q}_n)^T$, 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 $\omega = (w_2, \dots, w_n)$ and $\xi = (z_2, \dots, z_n)$, 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 $(y_2, \dots, y_n)$, where $y_j \in \{w_j, z_j\}$, is a complementary feasible basic vector for $(\bar{q}, \overline{\mathcal{M}})$, then, since $1 \in \mathbf{J}$, $(z_1, y_2, \dots, y_n)$ 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 $n-1$.

$\square$

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 $\text{Pos}(\mathbf{\Gamma} \cup \{-B_{\cdot 1}\})$. Define $\bar{b} = b - \frac{B_{\cdot 1}(b^T B_{\cdot 1})}{||B_{\cdot 1}||^2}$, $\bar{B}_{\cdot j} = B_{\cdot j} - \frac{B_{\cdot 1}(B_{\cdot 1})^T B_{\cdot j}}{||B_{\cdot 1}||^2}$, for $j = 2, \dots, n$. Let $\overline{\mathbf{\Gamma}} = \{\bar{B}_{\cdot 2}, \dots, \bar{B}_{\cdot n}\}$. For $2 \leqq j \leqq n$, $\bar{B}_{\cdot j}$ is orthogonal to $B_{\cdot 1}$ and the cone $\text{Pos}(\mathbf{\Gamma} \cup \{-B_{\cdot 1}\})$ is the direct sum of the full line generated by $B_{\cdot 1}$ 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 $x^* = \bar{x}^* + \frac{B_{\cdot 1}(b^T B_{\cdot 1})}{||B_{\cdot 1}||^2}$ 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. $\quad w_j z_j = 0$ 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. $\quad w_j z_j = 0$ 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 $0 \neq \bar{x} \in \text{Pos}(\mathbf{\Gamma})$ satisfying $0 \in \mathbf{T}(b; \bar{x})$, if for some $i \in \{1, \dots, n\}$, 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 $\text{Pos}\{\bar{x}, B_{.i}\}$.

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 $\text{Pos}\{\bar{x}, B_{.i}\}$.

$\square$

Theorem 7.9 *Let $\emptyset \neq \mathbf{S} \subset \mathbf{\Gamma}$ be such that $\bar{x} = b(\mathbf{S}) \in \text{Pos}(\mathbf{S})$. 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 $\text{Pos}(\mathbf{\Gamma})$ 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 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 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 Pos$(\mathbf{\Gamma})$ to b. Let $\hat{x}$ be the nearest point in 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 $b \in$ Pos$(\mathbf{\Gamma})$, the nearest point in 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.

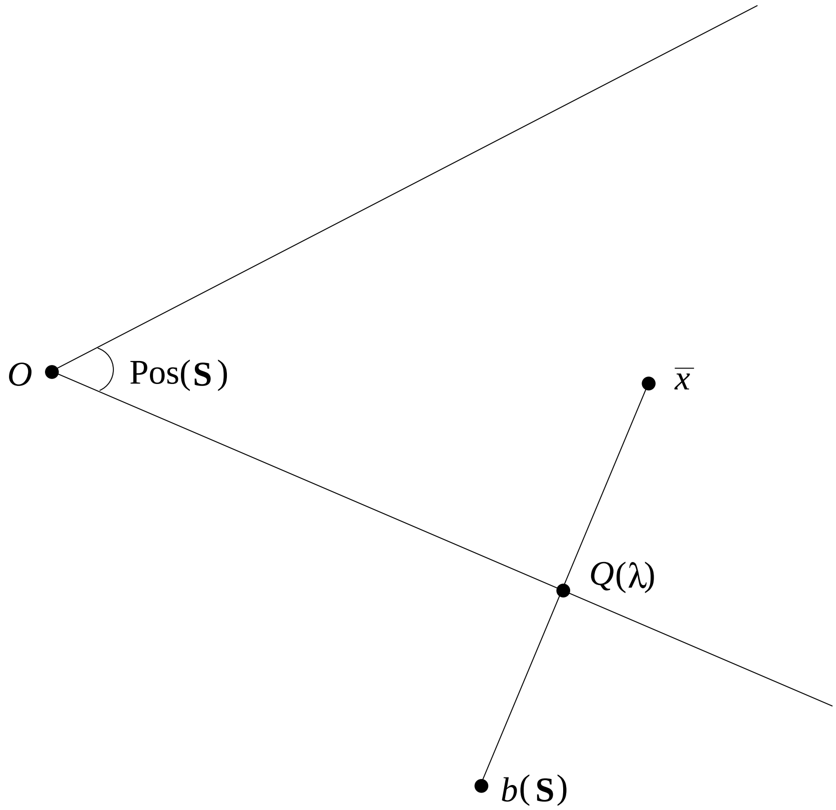

Figure 7.3

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 $\bar{\lambda} = \min\{\frac{\alpha_j}{(\alpha_j - \gamma_j)} : j \text{ such that } \gamma_j < 0\}$, 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 $f(x) = (f_1(x), \dots, f_m(x))^T$ 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 $\mathbf{J} \subset \{1, \dots, n\}$, 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 $\mathbf{J} \subset \{1, \dots, n\}$, 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 $\{B_{.1}, \dots, B_{.i-1}, B_{.i+1}, \dots, B_{.n}\}$. 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

$$\begin{aligned} h(y) = &\text{ maximum value of } y^T x, \text{ over } x \in \mathbf{K} \\ s(y) = &\text{ a point in } \mathbf{\Gamma} \text{ which maximizes } y^T x \text{ over } x \in \mathbf{K}. \text{ So, } h(y) = y^T s(y). \end{aligned}$$

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 $\|x\|^2 + h(-x) \geqq 0$. Also prove that $\|x\|^2 + h(-x) = 0$ 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 $\|x\|^2 + h(-x) > 0$, 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 $\|\overline{x}\|^2 + h(-\overline{x})$. If $\|\overline{x}\|^2 + h(-\overline{x}) = 0$, then $x^* = \overline{x}$, terminate the algorithm.

If $\overline{x} \neq 0$ and $\|\overline{x}\|^2 + h(-\overline{x}) > 0$, 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.