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

$$w_j = |x_j| - x_j,\ \ z_j = |x_j| + x_j, \qquad \text{for each } j = 1 \text{ to } n \tag{9.3}$$

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 $r \geqq 1$, $\|x^* - x^{r+1}\| \leqq \big(\frac{\rho^{\mathbf{r}}}{1-\rho}\big)\|x^2 - x^1\|$.*

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 $\|x^* - x^1\| \leqq \sum_{t=1}^{\infty} \|x^{t+1} - x^t\| \leqq \|x^2 - x^1\|\big(\sum_{t=0}^{\infty} \rho^t\big)$ (from (9.6)) $= \frac{\|x^2-x^1\|}{(1-\rho)}$. Using this in (9.5) leads to $\|x^* - x^{r+1}\| \leqq \big(\frac{\rho^{\mathbf{r}}}{1-\rho}\big)\|x^2 - x^1\|$ for $r \geqq 1$.

□

Theorem 9.5 *If* $x^1 = 0$, *we have* $\|x^* - x^{r+1}\| \leqq \rho^{\mathbf{r}}\big(\frac{\|b\|}{1-\rho}\big)$.

Proof. Follows from Theorem (9.4).

□

Theorem 9.6 *If* $x^1 = 0$, *we have for* $r \geqq 1$, $\|x^{r+1}\| \geqq \|b\|\big(\frac{1}{1+\rho} - \frac{\rho^{\mathbf{r}}}{1-\rho}\big)$.

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

$$|x_i^{r+1}| \geqq \frac{\|b\|}{\sqrt{n}}\Big(\frac{1}{1+\rho} - \frac{\rho^{\mathbf{r}}}{1-\rho}\Big)\ . \tag{9.7}$$

But from Theorem 9.5, for the same i, we must have $|x_i^* - x_i^{r+1}| \leqq \|b\|\big(\frac{\rho^{\mathbf{r}}}{1-\rho}\big)$. So if r is such that $\frac{1}{\sqrt{n}}\big(\frac{1}{1+\rho} - \frac{\rho^{\mathbf{r}}}{1-\rho}\big) > \frac{\rho^{\mathbf{r}}}{(1-\rho)}$, that is $r > N = \big\lceil \log\big(\frac{(1-\rho)}{(1+\sqrt{n})(1+\rho)}\big)/\log\rho \big\rceil$ 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

$$z^{r+1} = \lambda\big(z^r - \omega E^r (M z^r + q + K^r (z^{r+1} - z^r))\big)^+ + (1-\lambda) z^r \tag{9.9}$$

for $r = 0, 1, \dots$, 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 $\{E^r : r = 0, 1, \dots\}$, $\{K^r : r = 0, 1, \dots\}$ are bounded sequences of matrices. When K^r is strictly lower triangular, (9.9) yields,

$$\begin{aligned} z_1^{r+1} &= \lambda\big(z_1^r - \omega E_{11}^r (M_{1.} z^r + q_1)\big)^+ + (1-\lambda) z_1^r, \text{ and} \\ z_j^{r+1} &= \lambda\Big(z_j^r - \omega E_{jj}^r \big(M_{j.} z^r + q_j + \sum_{l=1}^{j-1} K_{jl}^r (z_l^{r+1} - z_l^r)\big)\Big)^+ + (1-\lambda) z_j^r, \text{ for } j = 2 \text{ to } n, \end{aligned}$$

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 $j = 1, 2, \dots, n$. If K^r is strictly upper triangular, (9.9) yields

$$\begin{aligned} z_n^{r+1} =& \lambda\big(z_n^r - \omega E_{nn}^r (M_{n.} z^r + q_n)\big)^+ + (1-\lambda) z_n^r, \text{ and} \\ z_j^{r+1} =& \lambda\Big(z_j^r - \omega E_{jj}^r \big(M_{j.} z^r + q_j + \sum_{l=j+1}^{n} K_{jl}^r (z_l^{r+1} - z_l^r)\big)\Big)^+ + \\ & (1-\lambda) z_j^r, \text{ for } j = n-1 \text{ to } 1, \end{aligned}$$

and so in this case z_{j+1}^r can be computed very conveniently in the specific order $j = n, n-1, \dots, 2, 1$.

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*

$$\big(z - \omega E(Mz + q)\big)^+ - z = 0, \text{ for some or all } \omega > 0 \ . \tag{9.12}$$

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), $0 = \big(\bar{z}_j - \omega E_{jj}(M_{j.}\bar{z} + q_j)\big)^+ - \bar{z}_j = -\omega E_{jj}(M_{j.}\bar{z} + q_j)$, 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).

$\square$

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.

$\square$

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 $0 = \lim_{t \to +\infty} \big(f(z^{r_t}) - f(z^{1+r_t})\big) \geqq \lim_{t \to +\infty} \gamma \|z^{1+r_t} - z^{r_t}\|^2 \geqq 0$. 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

$$\begin{aligned} 0 &= \lim_{t \to +\infty} \|z^{1+r_t} - z^{r_t}\| \\ &= \lambda \lim_{t \to +\infty} \|(z^{r_t} - \omega E^{r_t}(M z^{r_t} + q + K^{r_t}(z^{1+r_t} - z^{r_t})))^+ - z^{r_t}\| \\ &= \lambda \|(\bar{z} - \omega E(M\bar{z} + q))^+ - \bar{z}\| \ . \end{aligned}$$

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 $\bar{y} > 0$, $\bar{y}^T M \bar{y} = 0$, $q^T \bar{y} \leqq 0$. 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

$$\frac{\alpha}{\|z^{s_t}\|^2} \geqq \frac{f(z^{s_t})}{\|z^{s_t}\|^2} = \frac{1}{2}(y^{s_t})^T M y^{s_t} + \frac{q^T y^{s_t}}{\|z^{s_t}\|} \ .$$

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 $\{x^r : r = 0, 1, \dots\}$ 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 $\{x^r : r = 0, 1, \dots\}$ is not a continuum (i. e., a closed set which cannot be written as the union of two nonempty disjoint closed sets), then $\{x^r : r = 0, 1, \dots\}$ 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 $\{z^r : r = 0, 1, \dots\}$ 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 $\{z^r : r = 0, 1, \dots\}$ 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 $\{f(z^r) : r = 0, 1, \dots\}$ 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 $\{z^r : r = 0, 1, \dots\}$ 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 $\{z^r : r = 0, 1, \dots\}$ 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 $\{z^r : r = 0, 1, \dots\}$ 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 $\{z^r : r = 0, 1, \dots\}$ 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

$$
\begin{aligned}
z^0 &\geqq 0, \text{ an initial point} \\
z^{r+1} &= \lambda(z^r - \omega E(Mz^r + q))^+ + (1 - \lambda)z^r, \ r = 0, 1, \dots
\end{aligned}
\tag{9.17}
$$

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.

$$\begin{aligned} z^0 &\geqq 0, \text{ an initial point} \\ z^{r+1} &= \lambda(z^r - \omega E(Mz^r + q + K^r(z^{r+1} - z^r)))^+ + (1-\lambda)z^r, \ r = 0, 1, \dots \end{aligned} \tag{9.18}$$

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.

$$\begin{aligned} z^0 &\geqq 0, \text{ an initial point.} \\ z^{r+1} &= \lambda(z^r - \omega E(Mz^r + q + L(z^{r+1} - z^r)))^+ + (1-\lambda)z^r, \ r = 0, 2, 4, \dots \\ &= \lambda(z^r - \omega E(Mz^r + q + U(z^{r+1} - z^r)))^+ + (1-\lambda)z^r, \ r = 1, 3, 5, \dots \end{aligned} \tag{9.20}$$

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

$$\begin{array}{ll} \text{Minimize} & Q(x) = cx + \frac{1}{2}x^T Dx \\ \text{Subject to} & Ax \geqq b \\ & Fx = d \end{array} \tag{9.28}$$

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

$$
\begin{aligned}
&(u^0, \xi^0) \text{ selected arbitrarily to satisfy } u^0 \geqq 0. \\
&\begin{pmatrix} u^{r+1} \\ \xi^{r+1} \end{pmatrix} = \begin{pmatrix} u^r \\ \xi^r \end{pmatrix} - \omega E \Bigg[\begin{pmatrix} A \\ F \end{pmatrix} D^{-1} \begin{pmatrix} A^T & F^T \end{pmatrix} \begin{pmatrix} u^r \\ \xi^r \end{pmatrix} \\
&- \begin{pmatrix} A \\ F \end{pmatrix} D^{-1} c^T - \begin{pmatrix} b \\ d \end{pmatrix} + K^r \left(\begin{pmatrix} u^{r+1} \\ \xi^{r+1} \end{pmatrix} - \begin{pmatrix} u^r \\ \xi^r \end{pmatrix} \right) \Bigg]^*
\end{aligned}
\tag{9.30}
$$

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 $r = 0, 1, \dots$ 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 $\{z^r : r = 0, 1, \dots\}$ 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 $\{z^r : r = 0, 1, \dots\}$ 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 $\{z^r : r = 0, 1 \dots\}$ 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 $\{z^r : r = 0, 1, \dots\}$ 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 $\{z^r : r = 0, 1, \dots\}$ generated by (9.33) is unbounded, by Theorem 9.10, there exists a $\bar{y} \in \mathbf{R}^n$ satisfying: $\bar{y} \geqq 0$, $M\bar{y} = 0$, $q^T\bar{y} \leqq 0$. 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.

$$\begin{aligned} u_i^{r+1} &= \left(u_i^r - \frac{\omega}{\|A_{i.}D^{-\frac{1}{2}}\|^2} \Big((A_{i.}D^{-1}(\gamma^{i,r+1} + \sum_{t=i}^{m}(A_{t.})^T u_t^r + v^r - c^T)) - b_i \Big) \right)^+ \\ v^{r+1} &= \left(v^r - \omega(A^T u^{r+1} + v^r - c^T)\right)^+ \end{aligned} \tag{9.37}$$

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).

$\square$

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

$$z^{r+1} = \Big(z^r - \omega E\big(Mz^r + q + K(z^{r+1} - z^r)\big)\Big)^+, \quad r = 0, 1, \ldots \tag{9.38}$$

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 $\{z^r : r = 0, 1, \dots\}$ 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 $\{z^r : r = 0, 1, \dots\}$ obtained using (9.32) satisfy for each $r = 1, 2, \dots$*

$$|z^{r+1} - z^r| \leqq (I - \omega E|K|)^{-1}|I - \omega E(M - K)|{\cdot}|z^r - z^{r-1}| \, . \tag{9.39}$$

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.

$\square$

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 $\{z^r : r = 0, 1, \dots\}$ 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 $|z^r - z^0| \leqq |z^r - z^{r-1}| + \dots + |z^1 - z^0| \leqq (Q^r + \dots + I)|z^1 - z^0| \leqq (I - Q)^{-1}|z^1 - z^0|$, (since $\|Q\| < 1$) = a constant vector independent of r. So the sequence $\{z^r : r = 0, 1, \dots\}$ is bounded. So it has a subsequence $\{z^{r_t} : t = 1, 2, \dots\}$ which converges to a limit, $\bar{z}$, say. So $\lim\limits_{t\to\infty} |z^{r_t+1} - \bar{z}| < \lim\limits_{t\to\infty} |z^{r_t+1} - z^{r_t}| + \lim\limits_{t\to\infty} |z^{r_t} - \bar{z}| = 0$, 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 $\{z^r : r = 0, 1, \dots\}$ itself converges in the limit to $\bar{z}$.

$\square$

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 $0 < \omega < 1/\max\{M_{jj}E_{jj} : j = 1$ 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

$$\begin{aligned}
&f(z^{r+1}) - f(z^r) = \\
&= (q_{\Gamma} + A_{\Gamma\Gamma} z^r_{\Gamma})^T (z^{r+1}_{\Gamma} - z^r_{\Gamma}) + (z^{r+1}_{\Gamma} - z^r_{\Gamma})^T A_{\Gamma\Gamma} (z^{r+1}_{\Gamma} - z^r_{\Gamma})/2 \\
&= (q_{\Gamma} + C_{\Gamma\Gamma} z^r_{\Gamma} + (V + C^T)_{\Gamma\Gamma} z^{r+1}_{\Gamma})^T (z^{r+1}_{\Gamma} - z^r_{\Gamma}) \\
&\quad - (z^{r+1}_{\Gamma} - z^r_{\Gamma})^T V_{\Gamma\Gamma} (z^{r+1}_{\Gamma} - z^r_{\Gamma})/2 \\
&= (q_{\Gamma} + C_{\Gamma\Gamma} z^r_{\Gamma} + (V + C^T)_{\Gamma\Gamma} z^{r+1}_{\Gamma} + U_{\Gamma\overline{\Gamma}} z^{r+1}_{\overline{\Gamma}})^T (z^{r+1}_{\Gamma} - z^r_{\Gamma}) \\
&\quad - (z^{r+1}_{\overline{\Gamma}})^T (U^T_{\Gamma\overline{\Gamma}} z^{r+1}_{\Gamma} - U^T_{\Gamma\overline{\Gamma}} z^r_{\Gamma}) - (z^{r+1}_{\Gamma} - z^r_{\Gamma})^T V_{\Gamma\Gamma} (z^{r+1}_{\Gamma} - z^r_{\Gamma})/2 \\
&= (q_{\Gamma} + C_{\Gamma\Gamma} z^r_{\Gamma} + (V + C^T)_{\Gamma\Gamma} z^{r+1}_{\Gamma} + U_{\Gamma\overline{\Gamma}} z^{r+1}_{\overline{\Gamma}})^T ((z^{r+1} - E z^r) + (E - I) z^r)_{\Gamma} \\
&\quad + (z^{r+1}_{\overline{\Gamma}})^T ((q_{\overline{\Gamma}} - U^T_{\Gamma\overline{\Gamma}} z^{r+1}_{\Gamma}) - (q_{\overline{\Gamma}} - U^T_{\Gamma\overline{\Gamma}} z^r_{\Gamma})) - (z^{r+1}_{\Gamma} - z^r_{\Gamma})^T V_{\Gamma\Gamma} (z^{r+1}_{\Gamma} - z^r_{\Gamma})/2,
\end{aligned}$$

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.

$\square$

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

$$\begin{gathered} c^T + Dx + y \geqq 0, \quad x \geqq 0, \quad y \geqq 0 \\ b - x \geqq 0 \\ x^T(c^T + Dx + y) = y^T(b - x) = 0 \end{gathered} \tag{9.48}$$

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 $i_0 = 1$, $i_1 = 2$, $\ldots$, $i_m = 1$, $i_{m+1} = 2$, 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).