The new solution point will always be an interior point because if the solution point is close to the boundaries, at least one of the functions $1/g_i(\mathbf{X})$ or $(-1/x_i)$ will acquire a very large negative value. Because the objective is to maximize $p(\mathbf{X}, t)$, such solution points are automatically excluded. The main result is that successive solution points will always be interior points. Consequently, the problem can always be treated as an unconstrained case.

Once the optimum solution corresponding to a given value of t is obtained, a new value of t is generated and the optimization process (using the steepest ascent method) is repeated. If t' is the current value of t, the next value, t'', must be selected such that $0 < t'' < t'$.

The SUMT procedure is terminated if, for two successive values of t, the corresponding *optimum* values of $\mathbf{X}$ obtained by maximizing $p(\mathbf{X}, t)$ are approximately the same. At this point further trials will produce little improvement.

Actual implementation of SUMT involves more details than have been presented here. Specifically, the selection of an initial value of t is an important factor that can affect the speed of convergence. Further, the determination of an initial interior point may require special techniques. These details can be found in Fiacco and McCormick (1968).

21.3 SUMMARY

The solution methods of nonlinear programming can generally be classified as either *direct* or *indirect* procedures. Examples of direct methods are the gradient algorithms, wherein the maximum (minimum) of a problem is sought by following the fastest rate of increase (decrease) of the objective function at a point. In indirect methods, the original problem is replaced by an auxiliary one from which the optimum is determined. Examples of these situations include quadratic programming, separable programming, and stochastic programming.

SELECTED REFERENCES

BAZARAA, M., H. SHERALL, and C. SHETTY, *Nonlinear Programming, Theory and Algorithms*, 2nd ed., Wiley, New York, 1993.

BEIGHTLER, C., D. PHILLIPS, and D. WILDE, *Foundations of Optimization*, 2nd ed., Prentice Hall, Upper Saddle River, N.J., 1979

FIACCO, A., and G. MCCORMICK, *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*, Wiley, New York, 1968.

Appendix A

Review of Vectors and Matrices

A.1 VECTORS

A.1.1 Definition Of A Vector

Let $p_1, p_2, \ldots, p_n$ be any n real numbers and $\mathbf{P}$ an ordered set of these real numbers—that is,

$$\mathbf{P} = (p_1, p_2, \ldots, p_n)$$

Then $\mathbf{P}$ is an n-vector (or simply a vector). The ith components of $\mathbf{P}$ is given by p_i. For example, $\mathbf{P} = (1, 2)$ is a two-dimensional vector.

A.1.2 Addition (Subtraction) of Vectors

Let

$$\mathbf{P} = (p_1, p_2, \ldots, p_n) \quad \text{and} \quad \mathbf{Q} = (q_1, q_2, \ldots, q_n)$$

be two n-vectors. Then the components of the vector $\mathbf{R} = (r_1, r_2, \ldots, r_n)$ such that $\mathbf{R} = \mathbf{P} \pm \mathbf{Q}$ are given by

$$r_i = p_i \pm q_i$$

In general, given the vectors $\mathbf{P}$, $\mathbf{Q}$, and $\mathbf{S}$,

$$\mathbf{P} \pm \mathbf{Q} = \mathbf{Q} \pm \mathbf{P} \qquad \text{(Commutative law)}$$

$$(\mathbf{P} + \mathbf{Q}) + \mathbf{S} = \mathbf{P} + (\mathbf{Q} + \mathbf{S}) \qquad \text{(Associative law)}$$

$$\mathbf{P} + (-\mathbf{P}) = \mathbf{0} \qquad \text{(zero or null vector)}$$

A.1.3 Multiplication of Vectors by Scalars

Given a vector **P** and a scalar (constant) quantity θ, the new vector

$$\mathbf{Q} = \theta\mathbf{P} = (\theta p_1, \theta p_2, \ldots, \theta p_n)$$

is the scalar product of **P** and θ

In general, given the vectors **P** and **S** and the scalars θ and γ,

$$\theta(\mathbf{P} + \mathbf{S}) = \theta\mathbf{P} + \theta\mathbf{S} \qquad \text{(Distributive law)}$$

$$\theta(\gamma\mathbf{P}) = (\theta\gamma)\mathbf{P} \qquad \text{(Associative law)}$$

A.1.4 Linearly Independent Vectors

The vectors $\mathbf{P}_1, \mathbf{P}_2, \ldots, \mathbf{P}_n$ are *linearly independent* if and only if, for all real θ_j,

$$\sum_{j=1}^{n} \theta_j \mathbf{P}_j = \mathbf{0}$$

implies that all $\theta_j = 0$. If

$$\sum_{j=1}^{n} \theta_j \mathbf{P}_j = \mathbf{0}$$

for some $\theta \neq 0$, the vectors are said to be *linearly dependent*. For example, the vectors

$$\mathbf{P}_1 = (1, 2) \quad \text{and} \quad \mathbf{P}_2 = (2, 4)$$

are linearly dependent because there exist nonzero $\theta_1 = 2$ and $\theta_2 = -1$ for which

$$\theta_1\mathbf{P}_1 + \theta_2\mathbf{P}_2 = \mathbf{0}$$

A.2 MATRICES

A.2.1 Definition of a Matrix

A matrix is a rectangular array of elements. The element a_{ij} of the matrix **A** occupies the ith row and jth column of the array. A matrix with m rows and n columns is said to have the order or size $m \times n$. For example,

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{bmatrix} = \|a_{ij}\|_{4\times 3}$$

is a (4×3)-matrix.

A.2.2 Types of Matrices

1. A *square* matrix has $m = n$.

2. An identity matrix is a square matrix in which all the diagonal elements are one and all the off-diagonal elements are zero. For example, a (3×3) identity matrix is given by

$$\mathbf{I}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

3. A *row vector* is a matrix with one row and n columns.
4. A *column vector* is a matrix with m rows and one column.
5. The matrix $\mathbf{A}^T$ is the **transpose** of **A** if the element a_{ij} in **A** is equal to element a_{ji} in $\mathbf{A}^T$ for all i and j. For example, if

$$\mathbf{A} = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$$

then

$$\mathbf{A}^T = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$$

6. A matrix $\mathbf{B} = \mathbf{0}$ is called a **zero matrix** if every element of **B** is zero.
7. Two matrices $\mathbf{A} = \|a_{ij}\|$ and $\mathbf{B} = \|b_{ij}\|$ are equal if and only if they have the same size and $a_{ij} = b_{ij}$ for all i and j.

A.2.3 Matrix Arithmetic Operations

In matrices only addition (subtraction) and multiplication are defined. The division, though not defined, is replaced by inversion (see Section A.2.6).

Addition (subtraction) of matrices. Two matrices $\mathbf{A} = \|a_{ij}\|$ and $\mathbf{B} = \|b_{ij}\|$ can be added together if they are of the same size $(m \times n)$. The sum $\mathbf{D} = \mathbf{A} + \mathbf{B}$ is obtained by adding the corresponding elements. Thus,

$$\|d_{ij}\|_{m\times n} = \|a_{ij} + b_{ij}\|_{m\times n}$$

If the matrices A, B, and C have the same size, then

$$\mathbf{A} \pm \mathbf{B} = \mathbf{B} \pm \mathbf{A} \qquad \text{(Commutative law)}$$

$$\mathbf{A} \pm (\mathbf{B} \pm \mathbf{C}) = (\mathbf{A} \pm \mathbf{B}) \pm \mathbf{C} \qquad \text{(Associative law)}$$

$$(\mathbf{A} \pm \mathbf{B})^T = \mathbf{A}^T \pm \mathbf{B}^T$$

Product of matrices. The product **AB** of two matrices $\mathbf{A} = \|a_{ij}\|$ and $\mathbf{B} = \|b_{ij}\|$ is defined if and only if the number of columns of **A** equals the number of rows of **B**. Thus, if **A** is of the size $(m \times r)$, then **B** is of the size $(r \times n)$, where m an n are arbitrary sizes.

Let $\mathbf{D} = \mathbf{AB}$. Then $\mathbf{D}$ is of the size $(m \times n)$, and its elements d_{ij} are given by

$$d_{ij} = \sum_{k=1}^{r} a_{ik} b_{kj}, \qquad \text{for all } i \text{ and } j$$

For example, if

$$\mathbf{A} = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} \quad \text{and} \quad \mathbf{B} = \begin{bmatrix} 5 & 7 & 9 \\ 6 & 8 & 0 \end{bmatrix}$$

then

$$\mathbf{D} = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} 5 & 7 & 9 \\ 6 & 8 & 0 \end{bmatrix} = \begin{bmatrix} (1 \times 5 + 3 \times 6) & (1 \times 7 + 3 \times 8) & (1 \times 9 + 3 \times 0) \\ (2 \times 5 + 4 \times 6) & (2 \times 7 + 4 \times 8) & (2 \times 9 + 4 \times 0) \end{bmatrix}$$

$$= \begin{bmatrix} 23 & 31 & 9 \\ 34 & 46 & 18 \end{bmatrix}$$

In general, $\mathbf{AB} \neq \mathbf{BA}$ even if $\mathbf{BA}$ is defined.

Matrix multiplication follows these general properties:

$$\mathbf{I}_m \mathbf{A} = \mathbf{A} \mathbf{I}_n = \mathbf{A}, \quad \text{where } \mathbf{I} \text{ is an identity matrix}$$

$$(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$$

$$\mathbf{C}(\mathbf{A} + \mathbf{B}) = \mathbf{CA} + \mathbf{CB}$$

$$(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}$$

$$\alpha(\mathbf{AB}) = (\alpha \mathbf{A})\mathbf{B} = \mathbf{A}(\alpha \mathbf{B}), \quad \alpha \text{ is a scalar}$$

Multiplication of partitioned matrices. Let $\mathbf{A}$ be an $(m \times r)$-matrix and $\mathbf{B}$ an $(r \times n)$-matrix. If $\mathbf{A}$ and $\mathbf{B}$ are partitioned in to the following submatrices

$$\mathbf{A} = \left[\begin{array}{c|c|c} \mathbf{A}_{11} & \mathbf{A}_{12} & \mathbf{A}_{13} \\ \hline \mathbf{A}_{21} & \mathbf{A}_{22} & \mathbf{A}_{23} \end{array}\right] \quad \text{and} \quad \mathbf{B} = \left[\begin{array}{c|c} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \hline \mathbf{B}_{21} & \mathbf{B}_{22} \\ \hline \mathbf{B}_{31} & \mathbf{B}_{32} \end{array}\right]$$

such that the number of columns of $\mathbf{A}_{ij}$ is equal to the number of rows of $\mathbf{B}_{ji}$ for all i and j, then

$$\mathbf{A} \times \mathbf{B} = \left[\begin{array}{c|c} \mathbf{A}_{11}\mathbf{B}_{11} + \mathbf{A}_{12}\mathbf{B}_{21} + \mathbf{A}_{13}\mathbf{B}_{31} & \mathbf{A}_{11}\mathbf{B}_{12} + \mathbf{A}_{12}\mathbf{B}_{22} + \mathbf{A}_{13}\mathbf{B}_{32} \\ \hline \mathbf{A}_{21}\mathbf{B}_{11} + \mathbf{A}_{22}\mathbf{B}_{21} + \mathbf{A}_{23}\mathbf{B}_{31} & \mathbf{A}_{21}\mathbf{B}_{12} + \mathbf{A}_{22}\mathbf{B}_{22} + \mathbf{A}_{23}\mathbf{B}_{32} \end{array}\right]$$

For example,

$$\left[\begin{array}{c|cc} 1 & 2 & 3 \\ \hline 1 & 0 & 5 \\ 2 & 5 & 6 \end{array}\right] \left[\begin{array}{c} 4 \\ \hline 1 \\ 8 \end{array}\right] = \left[\begin{array}{c} (1)\,(4) + (2 \quad 3)\begin{bmatrix} 1 \\ 8 \end{bmatrix} \\ \hline \begin{bmatrix} 1 \\ 2 \end{bmatrix}(4) + \begin{bmatrix} 0 & 5 \\ 5 & 6 \end{bmatrix}\begin{bmatrix} 1 \\ 8 \end{bmatrix} \end{array}\right] = \left[\begin{array}{c} 4 + 2 + 24 \\ \hline \begin{bmatrix} 4 \\ 8 \end{bmatrix} + \begin{bmatrix} 40 \\ 53 \end{bmatrix} \end{array}\right] = \begin{bmatrix} 30 \\ 44 \\ 61 \end{bmatrix}$$

A.2.4 Determinant of a square matrix

Given the n-square matrix

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

consider the product

$$P_{j_1 j_2 \cdots j_n} = a_{1j_1} a_{2j_2} \cdots a_{nj_n}$$

the elements of which are selected such that each column and each row of $\mathbf{A}$ is represented exactly once among the subscripts of $P_{j_1 j_2 \ldots j_n}$. Next define $\epsilon_{j_1 j_2 \ldots j_n}$ equal to $+1$ if $j_1 j_2 \ldots j_n$ is an even permutation and -1 if $j_1 j_2 \ldots j_n$ is an odd permutation. Thus the scalar

$$\sum_{\rho} \epsilon_{j_1 j_2 \cdots j_n} \rho_{j_1 j_2 \cdots j_n}$$

is called the determinant of $\mathbf{A}$, where ρ represents the summation over all $n!$ permutations. The notation det $\mathbf{A}$ or $|\mathbf{A}|$ is used to represent the determinant of $\mathbf{A}$.

To illustrate, consider

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

Then

$$|\mathbf{A}| = a_{11}(a_{22}\,a_{33} - a_{23}\,a_{32}) - a_{12}(a_{21}\,a_{33} - a_{31}\,a_{23}) + a_{13}(a_{21}\,a_{32} - a_{22}a_{31})$$

The properties of determinants include the following:

1. If every element of a column or a row is zero, then the value of the determinant is zero.
2. The value of the determinant is not changed if its rows and columns are interchanged.
3. If $\mathbf{B}$ is obtained from $\mathbf{A}$ by interchanging any two of its rows (or columns), then $|\mathbf{B}| = -|\mathbf{A}|$.
4. If two rows (or columns) of $\mathbf{A}$ are identical, then $|\mathbf{A}| = 0$.
5. The value of $|\mathbf{A}|$ remains the same if scalar α times a column (row) vector is added to another column (row) vector.
6. If every element of a column (or a row) of a determinant is multiplied by a scalar α, the value of the determinant is multiplied by α.
7. If $\mathbf{A}$ and $\mathbf{B}$ are two n-square matrices, then

$$|\mathbf{AB}| = |\mathbf{A}|\,|\mathbf{B}|$$

Definition of the Minor of a Determinant. The minor M_{ij} of the element a_{ij} in the determinant $|\mathbf{A}|$ is obtained from the matrix $\mathbf{A}$ by striking out the ith row and jth column of $\mathbf{A}$. For example, for

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

$$M_{11} = \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}, \quad M_{22} = \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix}, \ldots$$

Definition of the Adjoint Matrix. Let $A_{ij} = (-1)^{i+j} M_{ij}$ be defined as the **cofactor** of the element a_{ij} of the square matrix $\mathbf{A}$. Then, the adjoint matrix of $\mathbf{A}$ is defined as

$$\text{adj } \mathbf{A} = \|A_{ij}\|^T = \begin{bmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & & \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{bmatrix}$$

For example, if

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 2 \\ 3 & 3 & 4 \end{bmatrix}$$

then, $A_{11} = (-1)^2(3 \times 4 - 2 \times 3) = 6, A_{12} = (-1)^3(2 \times 4 - 2 \times 3) = -2, \ldots$, or

$$\text{adj } \mathbf{A} = \begin{bmatrix} 6 & 1 & -5 \\ -2 & -5 & 4 \\ -3 & 3 & -1 \end{bmatrix}$$

A.2.5 Nonsingular Matrix

A matrix is of a rank r if the largest *square* array in the matrix with nonvanishing determinants is of size r. A *square* matrix whose determinant does not vanish is called a **full-rank** or a **nonsingular** matrix. For example,

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 5 & 7 \end{bmatrix}$$

is a singular matrix because

$$|\mathbf{A}| = 1(21 - 20) - 2(14 - 12) + 3(10 - 9) = 0$$

But A has a rank $r = 2$ because

$$\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} = -1 \neq 0$$

A.2.6 The Inverse of a Matrix

If $\mathbf{B}$ and $\mathbf{C}$ are two n-square matrices such that $\mathbf{BC} = \mathbf{CB} = \mathbf{I}$, then $\mathbf{B}$ is called the inverse of $\mathbf{C}$ and $\mathbf{C}$ the inverse of $\mathbf{B}$. The common notation for the inverses is $\mathbf{B}^{-1}$ and $\mathbf{C}^{-1}$.

Theorem

If $\mathbf{BC} = \mathbf{I}$ *and* $\mathbf{B}$ *is* **nonsingular,** *then* $\mathbf{C} = \mathbf{B}^{-1}$, *which means that the inverse is unique.*

Proof. By assumption,

$$\mathbf{BC} = \mathbf{I}$$

then

$$\mathbf{B}^{-1}\mathbf{BC} = \mathbf{B}^{-1}\mathbf{I}$$

or

$$\mathbf{IC} = \mathbf{B}^{-1}$$

or

$$\mathbf{C} = \mathbf{B}^{-1}$$

Two important results can be proved for nonsingular matrices.

1. If $\mathbf{A}$ and $\mathbf{B}$ are nonsingular n-squre matrices, then $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$
2. If $\mathbf{A}$ is nonsingular, then $\mathbf{AB} = \mathbf{AC}$ implies that $\mathbf{B} = \mathbf{C}$.

Matrix inversion is used to solve n linearly independent equations. Consider

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

where x_i represents the unknowns and a_{ij} and b_i are constants. These n equations can be written in matrix form as

$$\mathbf{AX} = \mathbf{b}$$

Because the equations are independent, it follows that $\mathbf{A}$ is nonsingular. Thus

$$\mathbf{A}^{-1}\mathbf{AX} = \mathbf{A}^{-1}\mathbf{b} \quad \text{or} \quad \mathbf{X} = \mathbf{A}^{-1}\mathbf{b}$$

gives the solution of the n unknowns.

A.2.7 Methods of Computing the Inverse of a Matrix

Adjoint matrix method. Given $\mathbf{A}$ a nonsingular matrix of size n,

$$\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|}\,\text{adj}\,\mathbf{A} = \frac{1}{|\mathbf{A}|}\begin{bmatrix} A_{11} & A_{21} & \cdots & A_{n1}\\ A_{12} & A_{22} & \cdots & A_{n2}\\ \vdots & \vdots & & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn}\end{bmatrix}$$

For example, for

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3\\ 2 & 3 & 2\\ 3 & 3 & 4\end{bmatrix}$$

$$\text{adj}\,\mathbf{A} = \begin{bmatrix} 6 & 1 & -5\\ -2 & -5 & 4\\ -3 & 3 & -1\end{bmatrix} \quad\text{and}\quad |\mathbf{A}| = -7$$

Hence

$$\mathbf{A}^{-1} = \frac{1}{-7}\begin{bmatrix} 6 & 1 & -5\\ -2 & -5 & 4\\ -3 & 3 & -1\end{bmatrix} = \begin{bmatrix} -\frac{6}{7} & -\frac{1}{7} & \frac{5}{7}\\ \frac{2}{7} & \frac{5}{7} & -\frac{4}{7}\\ \frac{3}{7} & -\frac{3}{7} & \frac{1}{7}\end{bmatrix}$$

Row operations (Gauss-Jordan) method. Consider the partitioned matrix $(\mathbf{A}|\mathbf{I})$, where $\mathbf{A}$ is nonsingular. By premultiplying this matrix by $\mathbf{A}^{-1}$, we obtain

$$(\mathbf{A}^{-1}\mathbf{A}|\mathbf{A}^{-1}\mathbf{I}) = (\mathbf{I}|\mathbf{A}^{-1})$$

Thus, by multiplying a sequence of row transformations, the matrix $\mathbf{A}$ is changed to $\mathbf{I}$ and $\mathbf{I}$ is changed to $\mathbf{A}^{-1}$.

For example, consider the system of equations:

$$\begin{bmatrix} 1 & 2 & 3\\ 2 & 3 & 2\\ 3 & 3 & 4\end{bmatrix}\begin{bmatrix} x_1\\ x_2\\ x_3\end{bmatrix} = \begin{bmatrix} 3\\ 4\\ 5\end{bmatrix}$$

The solution of $\mathbf{X}$ and the inverse of basis matrix can be obtained directly by considering

$$\mathbf{A}^{-1}(\mathbf{A}|\mathbf{I}|\mathbf{b}) = (\mathbf{I}|\mathbf{A}^{-1}|\mathbf{A}^{-1}\mathbf{b})$$

Thus, by a row transformation operation, we get

$$\left[\begin{array}{ccc|ccc|c} 1 & 2 & 3 & 1 & 0 & 0 & 3\\ 2 & 3 & 2 & 0 & 1 & 0 & 4\\ 3 & 3 & 4 & 0 & 0 & 1 & 5\end{array}\right]$$

Iteration 1

$$\left[\begin{array}{ccc|ccc|c} 1 & 2 & 3 & 1 & 0 & 0 & 3\\ 0 & -1 & -4 & -2 & 1 & 0 & -2\\ 0 & -3 & -5 & -3 & 0 & 1 & -4\end{array}\right]$$

Iteration 2

$$\left[\begin{array}{ccc|ccc|c} 1 & 0 & -5 & -3 & 2 & 0 & -1\\ 0 & 1 & 4 & 2 & -1 & 0 & 2\\ 0 & 0 & 7 & 3 & -3 & 1 & 2\end{array}\right]$$

Iteration 3

$$\left[\begin{array}{ccc|ccc|c} 1 & 0 & 0 & -\frac{6}{7} & -\frac{1}{7} & \frac{5}{7} & \frac{3}{7}\\ 0 & 1 & 0 & \frac{2}{7} & \frac{5}{7} & -\frac{4}{7} & \frac{6}{7}\\ 0 & 0 & 1 & \frac{3}{7} & -\frac{3}{7} & \frac{1}{7} & \frac{2}{7}\end{array}\right]$$

This gives $x_1 = \frac{3}{7}$, $x_2 = \frac{6}{7}$, and $x_3 = \frac{2}{7}$. The inverse of A is given by the right-hand–side matrix, which is the same as obtained by the method of adjoint matrix.

Partitioned matrix method. Let the two nonsingular matrices $\mathbf{A}$ and $\mathbf{B}$ of size n be partitioned as shown subsequently, given that A_{11} is nonsingular.

$$\mathbf{A} = \left[\begin{array}{c|c} \mathbf{A}_{11} & \mathbf{A}_{12}\\ (p\times p) & (p\times q)\\ \hline \mathbf{A}_{21} & \mathbf{A}_{22}\\ (q\times p) & (q\times q)\end{array}\right] \quad\text{and}\quad \mathbf{B} = \left[\begin{array}{c|c} \mathbf{B}_{11} & \mathbf{B}_{12}\\ (p\times p) & (p\times q)\\ \hline \mathbf{B}_{21} & \mathbf{B}_{22}\\ (q\times p) & (q\times q)\end{array}\right]$$

If $\mathbf{B}$ is the inverse of $\mathbf{A}$, then from $\mathbf{AB} = \mathbf{I}_n$, we have

$$\mathbf{A}_{11}\mathbf{B}_{11} + \mathbf{A}_{12}\mathbf{B}_{21} = \mathbf{I}_p$$

$$\mathbf{A}_{11}\mathbf{B}_{12} + \mathbf{B}_{12}\mathbf{A}_{22} = \mathbf{0}$$

Also, from $\mathbf{BA} = \mathbf{I}_n$, we get

$$\mathbf{B}_{21}\mathbf{A}_{11} + \mathbf{B}_{22}\mathbf{A}_{21} = \mathbf{0}$$

$$\mathbf{B}_{21}\mathbf{A}_{12} + \mathbf{B}_{22}\mathbf{A}_{22} = \mathbf{I}_q$$

Because $\mathbf{A}_{11}$ is nonsingular, that is, $|\mathbf{A}_{11}| \neq 0$, solving for $\mathbf{B}_{11}, \mathbf{B}_{12}, \mathbf{B}_{21}$, and $\mathbf{B}_{22}$, we get

$$\mathbf{B}_{11} = \mathbf{A}_{11}^{-1} + (\mathbf{A}_{11}^{-1}\mathbf{A}_{12})\mathbf{D}^{-1}(\mathbf{A}_{21}\mathbf{A}_{11}^{-1})$$
$$\mathbf{B}_{12} = -(\mathbf{A}_{11}^{-1}\mathbf{A}_{12})\mathbf{D}^{-1}$$
$$\mathbf{B}_{21} = -\mathbf{D}^{-1}(\mathbf{A}_{21}\mathbf{A}_{11}^{-1})$$
$$\mathbf{B}_{22} = \mathbf{D}^{-1}$$

where

$$\mathbf{D} = \mathbf{A}_{22} - \mathbf{A}_{21}(\mathbf{A}_{11}^{-1}\mathbf{A}_{12})$$

To illustrate the use of these formulas, partition the matrix

$$\mathbf{A} = \left[\begin{array}{c|cc} 1 & 2 & 3 \\ \hline 2 & 3 & 2 \\ 3 & 3 & 4 \end{array}\right]$$

such that

$$\mathbf{A}_{11} = (1),\ \mathbf{A}_{12} = (2, 3),\ \mathbf{A}_{21} = \begin{bmatrix} 2 \\ 3 \end{bmatrix},\ \text{and}\ \mathbf{A}_{22} = \begin{bmatrix} 3 & 2 \\ 3 & 4 \end{bmatrix}$$

In this case, $\mathbf{A}_{11}^{-1} = 1$ and

$$\mathbf{D} = \begin{bmatrix} 3 & 2 \\ 3 & 4 \end{bmatrix} - \begin{bmatrix} 2 \\ 3 \end{bmatrix} (1)\,(2, 3) = \begin{bmatrix} -1 & -4 \\ -3 & -5 \end{bmatrix}$$

$$\mathbf{D}^{-1} = -\frac{1}{7}\begin{bmatrix} -5 & 4 \\ 3 & -1 \end{bmatrix} = \begin{bmatrix} \frac{5}{7} & -\frac{4}{7} \\ -\frac{3}{7} & \frac{1}{7} \end{bmatrix}$$

Thus,

$$\mathbf{B}_{11} = \begin{bmatrix} -\frac{6}{7} \end{bmatrix} \quad \text{and} \quad \mathbf{B}_{12} = \begin{bmatrix} -\frac{1}{7} & \frac{5}{7} \end{bmatrix}$$

$$\mathbf{B}_{21} = \begin{bmatrix} \frac{2}{7} \\ \frac{3}{7} \end{bmatrix} \quad \text{and} \quad \mathbf{B}_{22} = \begin{bmatrix} \frac{5}{7} & -\frac{4}{7} \\ -\frac{3}{7} & \frac{1}{7} \end{bmatrix}$$

which directly give $\mathbf{B} = \mathbf{A}^{-1}$

A.3 QUADRATIC FORMS

Given

$$\mathbf{X} = (x_1, x_2, \ldots, x_n)^T$$

and

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

the function

$$Q(\mathbf{X}) = \mathbf{X}^T\mathbf{A}\mathbf{X} = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}x_ix_j$$

is called a quadratic form. The matrix $\mathbf{A}$ can always be assumed symmetric because each element of every pair of coefficients a_{ij} and a_{ji} $(i \neq j)$ can be replaced by $\frac{(a_{ij} + a_{ji})}{2}$ without changing the value of $Q(\mathbf{X})$. This assumption has advantages and hence is taken as a requirement.

To illustrate, the quadratic form

$$Q(\mathbf{X}) = (x_1, x_2, x_3)\begin{bmatrix} 1 & 0 & 1 \\ 2 & 7 & 6 \\ 3 & 0 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

is the same as

$$Q(\mathbf{X}) = (x_1, x_2, x_3)\begin{bmatrix} 1 & 1 & 2 \\ 1 & 7 & 3 \\ 2 & 3 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Note that $\mathbf{A}$ is symmetric in the second case.

The quadratic form is said to be

1. *Positive definite* if $Q(\mathbf{X}) > 0$ for every $\mathbf{X} \neq \mathbf{0}$.
2. *Positive semidefinite* if $Q(\mathbf{X}) \geq 0$ for every $\mathbf{X}$, and there exists $\mathbf{X} \neq \mathbf{0}$ such that $Q(\mathbf{X}) = 0$.
3. *Negative definite* if $-Q(\mathbf{X})$ is positive definite.
4. *Negative semidefinite* if $-Q(\mathbf{X})$ is positive semidefinite.
5. *Indefinite* in all other cases.

It can be proved that the necessary and sufficient conditions for the realization of the preceding cases are

1. $Q(\mathbf{X})$ is positive definite (semidefinite) if the values of the principal minor determinants of $\mathbf{A}$ are positive (nonnegative).† In this case, $\mathbf{A}$ is said to be positive definite (semidefinite).
2. $Q(\mathbf{X})$is negative definite if the value of kth principal minor determinants of $\mathbf{A}$ has the sign of $(-1)^k$, $k = 1, 2, \ldots, n$. In this case, $\mathbf{A}$ is called negative-definite.
3. $Q(\mathbf{X})$is a negative-semidefinite if the kth principal minor determinant of $\mathbf{A}$ is either zero or has the sign of $(-1)^k$, $k = 1, 2, \ldots, n$.

A.4 CONVEX AND CONCAVE FUNCTIONS

A function $f(\mathbf{X})$ is said to be strictly convex if, for any two distinct points $\mathbf{X}_1$ and $\mathbf{X}_2$,

$$f(\lambda\mathbf{X}_1 + (1-\lambda)\mathbf{X}_2) < \lambda f(\mathbf{X}_1) + (1-\lambda)f(\mathbf{X}_2)$$

where $0 < \lambda < 1$. Conversely, a function $f(\mathbf{X})$ is strictly concave if $-f(\mathbf{X})$ is strictly convex.

A special case of the convex (concave) function is the quadratic form (see Section A.3)

$$f(\mathbf{X}) = \mathbf{CX} + \mathbf{X}^T\mathbf{AX}$$

where $\mathbf{C}$ is a constant vector and $\mathbf{A}$ is a symmetric matrix. It can be proved that $f(\mathbf{X})$ is strictly convex if $\mathbf{A}$ is positive definite and $f(\mathbf{X})$ is strictly concave if $\mathbf{A}$ is negative definite.

SELECTED REFERENCES

HADLEY, G., *Matrix Algebra,* Addison-Wesley, Reading, Mass., 1961.
HOHN, F., *Elementary Matrix Algebra,* 2nd ed., Macmillan, New York, 1964.

PROBLEMS

■ **A–1.** Show that the following vectors are linearly dependent.

(a) $\begin{bmatrix}1\\-2\\3\end{bmatrix} \begin{bmatrix}-2\\4\\-2\end{bmatrix} \begin{bmatrix}1\\-2\\-1\end{bmatrix}$

†The kth *principal minor* determinant of $\mathbf{A}_{n\times n}$ is defined by

$$\begin{vmatrix}a_{11} & a_{12} & \cdots & a_{1k}\\ a_{21} & a_{22} & \cdots & a_{2k}\\ \vdots & \vdots & & \vdots\\ a_{k1} & a_{k2} & & a_{kk}\end{vmatrix}, \quad k = 1, 2, \ldots, n$$

(b) $\begin{bmatrix}2\\-3\\4\\5\end{bmatrix} \begin{bmatrix}4\\-6\\8\\10\end{bmatrix}$

■ **A–2.** Given

$$\mathbf{A} = \begin{bmatrix}1 & 4 & 9\\ 2 & 5 & -8\\ 3 & 7 & 2\end{bmatrix} \quad \text{and} \quad \mathbf{B} = \begin{bmatrix}7 & -1 & 2\\ 9 & 4 & 8\\ 3 & 6 & 10\end{bmatrix}$$

find
(a) $\mathbf{A} + 7\mathbf{B}$
(b) $2\mathbf{A} - 3\mathbf{B}$
(c) $(\mathbf{A} + 7\mathbf{B})^T$

■ **A–3.** In Problem A–2, show that $\mathbf{AB} \neq \mathbf{BA}$

■ **A–4.** Given the partitioned matrices

$$\mathbf{A} = \left[\begin{array}{c|cc}1 & 5 & 7\\ 2 & -6 & 9\\ \hline 3 & 7 & 2\\ 4 & 9 & 1\end{array}\right] \quad \text{and} \quad \mathbf{B} = \left[\begin{array}{cccc}2 & 3 & -4 & 5\\ \hline 1 & 2 & 6 & 7\\ 3 & 1 & 0 & 9\end{array}\right]$$

find $\mathbf{AB}$ using partitioning.

■ **A–5.** In Problem A–2, find $\mathbf{A}^{-1}$ and $\mathbf{B}^{-1}$ using the following:
(a) Adjoint matrix method
(b) Row operations method
(c) Partitioned matrix method

■ **A–6.** Verify the formulas given in Section A.2.7 for obtaining the inverse of a partitioned matrix.

■ **A–7.** Find the inverse of

$$\mathbf{A} = \begin{pmatrix}1 & \mathbf{G}\\ \mathbf{H} & \mathbf{B}\end{pmatrix}$$

where $\mathbf{B}$ is a nonsingular matrix.

■ **A–8.** Show that the quadratic form

$$Q(x_1, x_2) = 6x_1 + 3x_2 - 4x_1x_2 - 2x_1^2 - 3x_2^2 - \frac{27}{4}$$

is negative definite.

■ **A–9.** Show that the quadratic form

$$Q(x_1, x_2, x_3) = 2x_1^2 + 2x_2^2 + 3x_3^2 + 2x_1x_2 + 2x_2x_3$$

is positive definite.

■ **A–10.** Show that the function $f(x) = e^x$ is strictly convex over all real values of x.

■ **A–11.** Show that the quadratic function

$$f(x_1, x_2, x_3) = 5x_1^2 + 5x_2^2 + 4x_3^2 + 4x_1x_2 + 2x_2x_3$$

is strictly convex.

■ **A–12.** In Problem A–11, show that $-f(x_1, x_2, x_3)$ is strictly concave.

Appendix B

Introduction to Simnet II†

B.1 MODELING FRAMEWORK

The design of SIMNET II views discrete simulation models as queueing systems. The language is based on a **network approach** that uses three suggestive nodes: a **source** from which transactions (customers) arrive, a **queue** where waiting may occur, and a **facility** where service is performed. A fourth node, **auxiliary,** is added to enhance the capabilities of the language.

Nodes in SIMNET II are connected by **branches.** As transactions (also called entities) traverse the branches, they perform important functions that include (1) controlling transaction flow anywhere in the network, (2) collecting pertinent statistics, and (3) performing arithmetic calculations.

During the simulation execution, SIMNET II keeps track of the transactions by placing them in **files.** A file can be thought of as a two-dimensional array in which each row stores information about a single transaction. The columns of the array represent the **attributes** that allow the modeler to keep track of the unique characteristics of each transaction.

SIMNET II uses three types of files.

1. Event calendar
2. Queue
3. Facility

†This appendix provides the basic features of SIMNET II. Space limitation does not allow the presentation of the intermediate and advanced features of the language. A complete documentation of SIMNET II is given in Hamdy A. Taha, *Simulation with SIMNET II,* second edition, SimTec, Inc., Fayetteville, AR, 1995