2

CLASSICAL OPTIMIZATION TECHNIQUES

2.1 INTRODUCTION

The classical methods of optimization are useful in finding the optimum solution of continuous and differentiable functions. These methods are analytical and make use of the techniques of differential calculus in locating the optimum points. Since some of the practical problems involve objective functions that are not continuous and/or differentiable, the classical optimization techniques have limited scope in practical applications. However, a study of the calculus methods of optimization forms a basis for developing most of the numerical techniques of optimization presented in subsequent chapters. In this chapter we present the necessary and sufficient conditions in locating the optimum solution of a single-variable function, a multivariable function with no constraints, and a multivariable function with equality and inequality constraints.

2.2 SINGLE-VARIABLE OPTIMIZATION

A function of one variable $f(x)$ is said to have a *relative* or *local minimum* at $x = x^*$ if $f(x^*) \leq f(x^* + h)$ for all sufficiently small positive and negative values of h. Similarly, a point x^* is called a *relative* or *local maximum* if $f(x^*) \geq f(x^* + h)$ for all values of h sufficiently close to zero. A function $f(x)$ is said to have a *global* or *absolute minimum* at x^* if $f(x^*) \leq f(x)$ for all x, and not just for all x close to x^*, in the domain over which $f(x)$ is defined. Similarly, a point x^* will be a global maximum of $f(x)$ if $f(x^*) \geq f(x)$ for all x in the domain. Figure 2.1 shows the difference between the local and global optimum points.

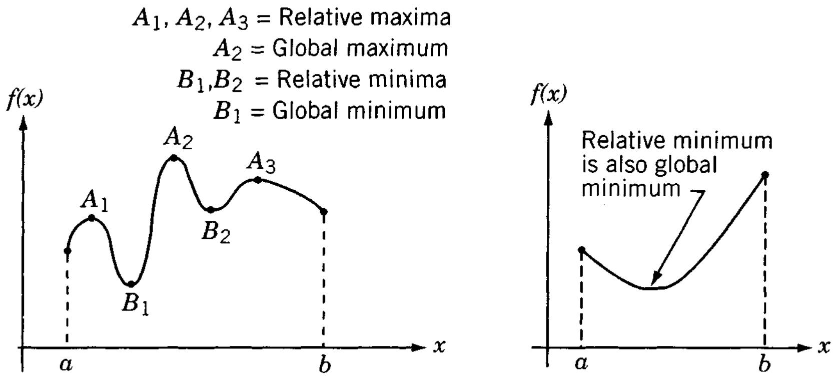

Figure 2.1 Relative and global minima.

A *single-variable optimization problem* is one in which the value of $x = x^*$ is to be found in the interval $[a,b]$ such that x^* minimizes $f(x)$. The following two theorems provide the necessary and sufficient conditions for the relative minimum of a function of a single variable.

Theorem 2.1: Necessary Condition If a function $f(x)$ is defined in the interval $a \leq x \leq b$ and has a relative minimum at $x = x^*$, where $a < x^* < b$, and if the derivative $df(x)/dx = f'(x)$ exists as a finite number at $x = x^*$, then $f'(x^*) = 0$.

Proof: It is given that

$$f'(x^*) = \lim_{h \to 0} \frac{f(x^* + h) - f(x^*)}{h} \tag{2.1}$$

exists as a definite number, which we want to prove to be zero. Since x^* is a relative minimum, we have

$$f(x^*) \leq f(x^* + h)$$

for all values of h sufficiently close to zero. Hence

$$\frac{f(x^* + h) - f(x^*)}{h} \geq 0 \quad \text{if } h > 0$$

$$\frac{f(x^* + h) - f(x^*)}{h} \leq 0 \quad \text{if } h < 0$$

Thus Eq. (2.1) gives the limit as h tends to zero through positive values as

$$f'(x^*) \geq 0 \tag{2.2}$$

while it gives the limit as h tends to zero through negative values as

$$f'(x^*) \leq 0 \tag{2.3}$$

The only way to satisfy both Eqs. (2.2) and (2.3) is to have

$$f'(x^*) = 0 \tag{2.4}$$

This proves the theorem.

Notes:

1. This theorem can be proved even if x^* is a relative maximum.
2. The theorem does not say what happens if a minimum or maximum occurs at a point x^* where the derivative fails to exist. For example, in Fig. 2.2,

$$\lim_{h\to 0} \frac{f(x^* + h) - f(x^*)}{h} = m^+ \text{ (positive) or } m^- \text{ (negative)}$$

depending on whether h approaches zero through positive or negative values, respectively. Unless the numbers m^+ and m^- are equal, the derivative $f'(x^*)$ does not exist. If $f'(x^*)$ does not exist, the theorem is not applicable.
3. The theorem does not say what happens if a minimum or maximum occurs at an endpoint of the interval of definition of the function. In this

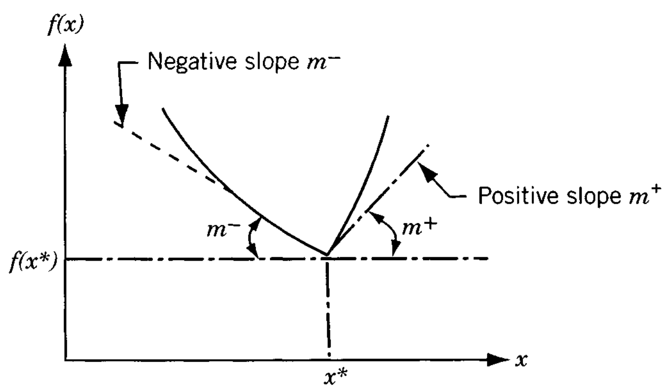

Figure 2.2 Derivative undefined at x^*.

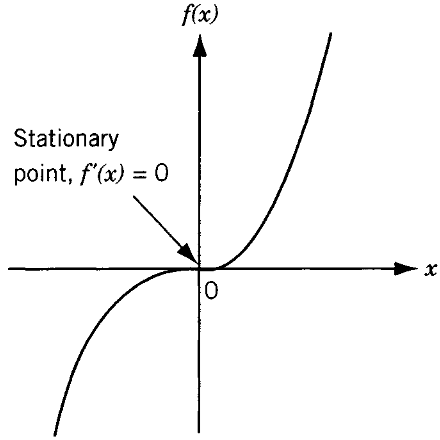

Figure 2.3 Stationary (inflection) point.

case

$$\lim_{h \to 0} \frac{f(x^* + h) - f(x^*)}{h}$$

exists for positive values of h only or for negative values of h only, and hence the derivative is not defined at the endpoints.

4. The theorem does not say that the function necessarily will have a minimum or maximum at every point where the derivative is zero. For example, the derivative $f'(x) = 0$ at $x = 0$ for the function shown in Fig. 2.3. However, this point is neither a minimum nor a maximum. In general, a point x^* at which $f'(x^*) = 0$ is called a *stationary point*.

If the function $f(x)$ possesses continuous derivatives of every order that come in question, in the neighborhood of $x = x^*$, the following theorem provides the sufficient condition for the minimum or maximum value of the function.

Theorem 2.2: Sufficient Condition Let $f'(x^*) = f''(x^*) = \cdots = f^{(n-1)}(x^*) = 0$, but $f^{(n)}(x^*) \neq 0$. Then $f(x^*)$ is (i) a minimum value of $f(x)$ if $f^{(n)}(x^*) > 0$ and n is even; (ii) a maximum value of $f(x)$ if $f^{(n)}(x^*) < 0$ and n is even; (iii) neither a maximum nor a minimum if n is odd.

Proof: Applying Taylor's theorem with remainder after n terms, we have

$$f(x^* + h) = f(x^*) + hf'(x^*) + \frac{h^2}{2!}f''(x^*) + \cdots + \frac{h^{n-1}}{(n-1)!}f^{(n-1)}(x^*)$$

$$+ \frac{h^n}{n!}f^{(n)}(x^* + \theta h) \quad \text{for} \quad 0 < \theta < 1 \tag{2.5}$$

Since $f'(x^*) = f''(x^*) = \cdots = f^{(n-1)}(x^*) = 0$, Eq. (2.5) becomes

$$f(x^* + h) - f(x^*) = \frac{h^n}{n!} f^{(n)}(x^* + \theta h)$$

As $f^{(n)}(x^*) \neq 0$, there exists an interval around x^* for every point x of which the nth derivative $f^{(n)}(x)$ has the same sign, namely, that of $f^{(n)}(x^*)$. Thus for every point $x^* + h$ of this interval, $f^{(n)}(x^* + \theta h)$ has the sign of $f^{(n)}(x^*)$. When n is even, $h^n/n!$ is positive irrespective of whether h is positive or negative, and hence $f(x^* + h) - f(x^*)$ will have the same sign as that of $f^{(n)}(x^*)$. Thus x^* will be a relative minimum if $f^{(n)}(x^*)$ is positive and a relative maximum if $f^{(n)}(x^*)$ is negative. When n is odd, $h^n/n!$ changes sign with the change in the sign of h and hence the point x^* is neither a maximum nor a minimum. In this case the point x^* is called a *point of inflection*.

Example 2.1 Determine the maximum and minimum values of the function

$$f(x) = 12x^5 - 45x^4 + 40x^3 + 5$$

SOLUTION Since $f'(x) = 60(x^4 - 3x^3 + 2x^2) = 60x^2(x - 1)(x - 2)$, $f'(x) = 0$ at $x = 0$, $x = 1$, and $x = 2$. The second derivative is

$$f''(x) = 60(4x^3 - 9x^2 + 4x)$$

At $x = 1, f''(x) = -60$ and hence $x = 1$ is a relative maximum. Therefore,

$$f_{\max} = f(x = 1) = 12$$

At $x = 2, f''(x) = 240$ and hence $x = 2$ is a relative minimum. Therefore,

$$f_{\min} = f(x = 2) = -11$$

At $x = 0, f''(x) = 0$ and hence we must investigate the next derivative.

$$f'''(x) = 60(12x^2 - 18x + 4) = 240 \quad \text{at} \quad x = 0$$

Since $f'''(x) \neq 0$ at $x = 0$, $x = 0$ is neither a maximum nor a minimum, and it is an inflection point.

Example 2.2 In a two-stage compressor, the working gas leaving the first stage of compression is cooled (by passing it through a heat exchanger) before it enters the second stage of compression to increase the efficiency [2.13]. The total work input to a compressor (W) for an ideal gas, for isentropic compression, is given by

$$W = c_p T_1 \left[\left(\frac{p_2}{p_1} \right)^{(k-1)/k} + \left(\frac{p_3}{p_2} \right)^{(k-1)/k} - 2 \right]$$

where c_p is the specific heat of the gas at constant pressure, k is the ratio of specific heat at constant pressure to that at constant volume of the gas, and T_1 is the temperature at which the gas enters the compressor. Find the pressure, p_2, at which intercooling should be done to minimize the work input to the compressor. Also determine the minimum work done on the compressor.

SOLUTION The necessary condition for minimizing the work done on the compressor is:

$$\frac{dW}{dp_2} = c_p T_1 \frac{k}{k-1} \left[\left(\frac{1}{p_1} \right)^{(k-1)/k} \frac{k-1}{k} (p_2)^{-1/k} + (p_3)^{(k-1)/k} \frac{-k+1}{k} (p_2)^{(1-2k)/k} \right] = 0$$

which yields

$$p_2 = (p_1 p_3)^{1/2}$$

The second derivative of W with respect to p_2 gives

$$\frac{d^2W}{dp_2^2} = c_p T_1 \left[-\left(\frac{1}{p_1} \right)^{(k-1)/k} \frac{1}{k} (p_2)^{-(1+k)/k} - (p_3)^{(k-1)/k} \frac{1-2k}{k} (p_2)^{(1-3k)/k} \right]$$

$$\left(\frac{d^2W}{dp_2^2} \right)_{p_2 = (p_1 p_2)^{1/2}} = \frac{2c_p T_1 \frac{k-1}{k}}{p_1^{(3k-1)/2k} p_3^{(k+1)/2k}}$$

Since the ratio of specific heats k is greater than 1, we get

$$\frac{d^2W}{dp_2^2} > 0 \quad \text{at} \quad p_2 = (p_1 p_3)^{1/2}$$

and hence the solution corresponds to a relative minimum. The minimum work done is given by

$$W_{\min} = 2c_p T_1 \frac{k}{k-1} \left[\left(\frac{p_3}{p_1} \right)^{(k-1)/2k} - 1 \right]$$

2.3 MULTIVARIABLE OPTIMIZATION WITH NO CONSTRAINTS

In this section we consider the necessary and sufficient conditions for the minimum or maximum of an unconstrained function of several variables. Before seeing these conditions, we consider the Taylor's series expansion of a multivariable function.

***Definition:* r*th Differential of* f** If all partial derivatives of the function f through order $r \geq 1$ exist and are continuous at a point $\mathbf{X}^*$, the polynomial

$$d^r f(\mathbf{X}^*) = \underbrace{\sum_{i=1}^{n} \sum_{j=1}^{n} \cdots \sum_{k=1}^{n}}_{r \text{ summations}} h_i h_j \cdots h_k \frac{\partial^r f(\mathbf{X}^*)}{\partial x_i \, \partial x_j \cdots \partial x_k} \tag{2.6}$$

is called the rth differential of f at $\mathbf{X}^*$. Notice that there are r summations and one h_i is associated with each summation in Eq. (2.6).

For example, when $r = 2$ and $n = 3$, we have

$$d^2 f(\mathbf{X}^*) = d^2 f(x_1^*, x_2^*, x_3^*) = \sum_{i=1}^{3} \sum_{j=1}^{3} h_i h_j \frac{\partial^2 f(\mathbf{X}^*)}{\partial x_i \, \partial x_j}$$

$$= h_1^2 \frac{\partial^2 f}{\partial x_1^2}(\mathbf{X}^*) + h_2^2 \frac{\partial^2 f}{\partial x_2^2}(\mathbf{X}^*) + h_3^2 \frac{\partial^2 f}{\partial x_3^2}(\mathbf{X}^*)$$

$$+ 2h_1 h_2 \frac{\partial^2 f}{\partial x_1 \, \partial x_2}(\mathbf{X}^*) + 2h_2 h_3 \frac{\partial^2 f}{\partial x_2 \, \partial x_3}(\mathbf{X}^*) + 2h_1 h_3 \frac{\partial^2 f}{\partial x_1 \, \partial x_3}(\mathbf{X}^*)$$

The Taylor's series expansion of a function $f(\mathbf{X})$ about a point $\mathbf{X}^*$ is given by

$$f(\mathbf{X}) = f(\mathbf{X}^*) + df(\mathbf{X}^*) + \frac{1}{2!} d^2 f(\mathbf{X}^*) + \frac{1}{3!} d^3 f(\mathbf{X}^*)$$

$$+ \cdots + \frac{1}{N!} d^N f(\mathbf{X}^*) + R_N(\mathbf{X}^*, \mathbf{h}) \tag{2.7}$$

where the last term, called the *remainder*, is given by

$$R_N(\mathbf{X}^*, \mathbf{h}) = \frac{1}{(N+1)!} d^{N+1} f(\mathbf{X}^* + \theta \mathbf{h}) \tag{2.8}$$

where $0 < \theta < 1$ and $\mathbf{h} = \mathbf{X} - \mathbf{X}^*$.

Example 2.3 Find the second-order Taylor's series approximation of the function

$$f(x_1,x_2,x_3) = x_2^2 x_3 + x_1 e^{x_3}$$

about the point $\mathbf{X}^* = \begin{Bmatrix} 1 \\ 0 \\ -2 \end{Bmatrix}$.

SOLUTION The second-order Taylor's series approximation of the function f about point $\mathbf{X}^*$ is given by

$$f(\mathbf{X}) = f\begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} + df\begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} + \frac{1}{2!}\, d^2 f\begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}$$

where

$$f\begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} = e^{-2}$$

$$df\begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} = h_1 \frac{\partial f}{\partial x_1}\begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} + h_2 \frac{\partial f}{\partial x_2}\begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} + h_3 \frac{\partial f}{\partial x_3}\begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}$$

$$= [h_1 e^{x_3} + h_2(2x_2x_3) + h_3 x_2^2 + h_3 x_1 e^{x_3}]\begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} = h_1 e^{-2} + h_3 e^{-2}$$

$$d^2 f\begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} = \sum_{i=1}^{3}\sum_{j=1}^{3} h_i h_j \frac{\partial^2 f}{\partial x_i\, \partial x_j}\begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} = \left(h_1^2 \frac{\partial^2 f}{\partial x_1^2} + h_2^2 \frac{\partial^2 f}{\partial x_2^2} + h_3^2 \frac{\partial^2 f}{\partial x_3^2} \right.$$

$$\left. + 2h_1h_2 \frac{\partial^2 f}{\partial x_1\, \partial x_2} + 2h_2h_3 \frac{\partial^2 f}{\partial x_2\, \partial x_3} + 2h_1h_3 \frac{\partial^2 f}{\partial x_1\, \partial x_3} \right)\begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}$$

$$= [h_1^2(0) + h_2^2(2x_3) + h_3^2(x_1 e^{x_3}) + 2h_1h_2(0) + 2h_2h_3(2x_2)$$

$$+ 2h_1h_3(e^{x_3})]\begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} = -4h_2^2 + e^{-2}h_3^2 + 2h_1h_3e^{-2}$$

Thus the Taylor's series approximation is given by

$$f(\mathbf{X}) \simeq e^{-2} + e^{-2}(h_1 + h_3) + \frac{1}{2!}(-4h_2^2 + e^{-2}h_3^2 + 2h_1 h_3 e^{-2})$$

where $h_1 = x_1 - 1$, $h_2 = x_2$, and $h_3 = x_3 + 2$.

Theorem 2.3: Necessary Condition If $f(\mathbf{X})$ has an extreme point (maximum or minimum) at $\mathbf{X} = \mathbf{X}^*$ and if the first partial derivatives of $f(\mathbf{X})$ exist at $\mathbf{X}^*$, then

$$\frac{\partial f}{\partial x_1}(\mathbf{X}^*) = \frac{\partial f}{\partial x_2}(\mathbf{X}^*) = \cdots = \frac{\partial f}{\partial x_n}(\mathbf{X}^*) = 0 \tag{2.9}$$

Proof: The proof given for Theorem 2.1 can easily be extended to prove the present theorem. However, we present a different approach to prove this theorem. Suppose that one of the first partial derivatives, say the kth one, does not vanish at $\mathbf{X}^*$. Then, by Taylor's theorem,

$$f(\mathbf{X}^* + \mathbf{h}) = f(\mathbf{X}^*) + \sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(\mathbf{X}^*) + R_1(\mathbf{X}^*, \mathbf{h})$$

that is,

$$f(\mathbf{X}^* + \mathbf{h}) - f(\mathbf{X}^*) = h_k \frac{\partial f}{\partial x_k}(\mathbf{X}^*) + \frac{1}{2!} d^2 f(\mathbf{X}^* + \theta\mathbf{h}), \qquad 0 < \theta < 1$$

Since $d^2 f(\mathbf{X}^* + \theta\mathbf{h})$ is of order h_i^2, the terms of order $\mathbf{h}$ will dominate the higher-order terms for small $\mathbf{h}$. Thus the sign of $f(\mathbf{X}^* + \mathbf{h}) - f(\mathbf{X}^*)$ is decided by the sign of $h_k\, \partial f(\mathbf{X}^*)/\partial x_k$. Suppose that $\partial f(\mathbf{X}^*)/\partial x_k > 0$. Then the sign of $f(\mathbf{X}^* + \mathbf{h}) - f(\mathbf{X}^*)$ will be positive for $h_k > 0$ and negative for $h_k < 0$. This means that $\mathbf{X}^*$ cannot be an extreme point. The same conclusion can be obtained even if we assume that $\partial f(\mathbf{X}^*)/\partial x_k < 0$. Since this conclusion is in contradiction with the original statement that $\mathbf{X}^*$ is an extreme point, we may say that $\partial f/\partial x_k = 0$ at $\mathbf{X} = \mathbf{X}^*$. Hence the theorem is proved.

Theorem 2.4: Sufficient Condition A sufficient condition for a stationary point $\mathbf{X}^*$ to be an extreme point is that the matrix of second partial derivatives (Hessian matrix) of $f(\mathbf{X})$ evaluated at $\mathbf{X}^*$ is (i) positive definite when $\mathbf{X}^*$ is a relative minimum point, and (ii) negative definite when $\mathbf{X}^*$ is a relative maximum point.

Proof: From Taylor's theorem we can write

$$f(\mathbf{X}^* + \mathbf{h}) = f(\mathbf{X}^*) + \sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(\mathbf{X}^*) + \frac{1}{2!}\sum_{i=1}^{n}\sum_{j=1}^{n} h_i h_j \left.\frac{\partial^2 f}{\partial x_i\, \partial x_j}\right|_{\mathbf{X}=\mathbf{X}^*+\theta\mathbf{h}},$$

$$0 < \theta < 1 \tag{2.10}$$

Since $\mathbf{X}^*$ is a stationary point, the necessary conditions give (Theorem 2.3)

$$\frac{\partial f}{\partial x_i} = 0, \qquad i = 1,2,\ldots,n$$

Thus Eq. (2.10) reduces to

$$f(\mathbf{X}^* + \mathbf{h}) - f(\mathbf{X}^*) = \frac{1}{2!} \sum_{i=1}^{n} \sum_{j=1}^{n} h_i h_j \left.\frac{\partial^2 f}{\partial x_i\, \partial x_j}\right|_{\mathbf{X}=\mathbf{X}^*+\theta \mathbf{h}}, \qquad 0 < \theta < 1$$

Therefore, the sign of

$$f(\mathbf{X}^* + \mathbf{h}) - f(\mathbf{X}^*)$$

will be same as that of

$$\sum_{i=1}^{n} \sum_{j=1}^{n} h_i h_j \left.\frac{\partial^2 f}{\partial x_i\, \partial x_j}\right|_{\mathbf{X}=\mathbf{X}^*+\theta \mathbf{h}}$$

Since the second partial derivative of $\partial^2 f(\mathbf{X})/\partial x_i\, \partial x_j$ is continuous in the neighborhood of $\mathbf{X}^*$,

$$\left.\frac{\partial^2 f}{\partial x_i\, \partial x_j}\right|_{\mathbf{X}=\mathbf{X}^*+\theta \mathbf{h}}$$

will have the same sign as $(\partial^2 f/\partial x_i\, \partial x_j)|\mathbf{X} = \mathbf{X}^*$ for all sufficiently small $\mathbf{h}$. Thus $f(\mathbf{X}^* + \mathbf{h}) - f(\mathbf{X}^*)$ will be positive, and hence X^* will be a relative minimum, if

$$Q = \sum_{i=1}^{n} \sum_{j=1}^{n} h_i h_j \left.\frac{\partial^2 f}{\partial x_i\, \partial x_j}\right|_{\mathbf{X}=\mathbf{X}^*} \tag{2.11}$$

is positive. This quantity Q is a quadratic form and can be written in matrix form as

$$Q = \mathbf{h}^T \mathbf{J} \mathbf{h}|_{\mathbf{X}=\mathbf{X}^*} \tag{2.12}$$

where

$$\mathbf{J}|_{\mathbf{X}=\mathbf{X}^*} = \left[\left.\frac{\partial^2 f}{\partial x_i\, \partial x_j}\right|_{\mathbf{X}=\mathbf{X}^*} \right] \tag{2.13}$$

is the matrix of second partial derivatives and is called the *Hessian matrix* of $f(\mathbf{X})$.

It is known from matrix algebra that the quadratic form of Eq. (2.11) or (2.12) will be positive for all **h** if and only if [**J**] is positive definite at **X** = **X***. This means that a sufficient condition for the stationary point **X*** to be a relative minimum is that the Hessian matrix evaluated at the same point be positive definite. This completes the proof for the minimization case. By proceeding in a similar manner, it can be proved that the Hessian matrix will be negative definite if **X*** is a relative maximum point.

Note: A matrix **A** will be positive definite if all its eigenvalues are positive; that is, all the values of λ that satisfy the determinantal equation

$$|\mathbf{A} - \lambda \mathbf{I}| = 0 \tag{2.14}$$

should be positive. Similarly, the matrix [**A**] will be negative definite if its eigenvalues are negative.

Another test that can be used to find the positive definiteness of a matrix **A** of order n involves evaluation of the determinants

$$A = |a_{11}|,$$

$$A_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix},$$

$$A_3 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{32} \end{vmatrix}, \ldots, \qquad A_n = \begin{vmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & & & & \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{vmatrix} \tag{2.15}$$

The matrix **A** will be positive definite if and only if all the values A_1, A_2, A_3, . . . , A_n are positive. The matrix **A** will be negative definite if and only if the sign of A_j is $(-1)^j$ for $j = 1,2,\ldots,n$. If some of the A_j are positive and the remaining A_j are zero, the matrix **A** will be positive semidefinite.

Example 2.4 Figure 2.4 shows two frictionless rigid bodies (carts) A and B connected by three linear elastic springs having spring constants k_1, k_2, and k_3. The springs are at their natural positions when the applied force P is zero. Find the displacements x_1 and x_2 under the force P by using the principle of minimum potential energy.

SOLUTION According to the principle of minimum potential energy, the system will be in equilibrium under the load P if the potential energy is a minimum. The potential energy of the system is given by

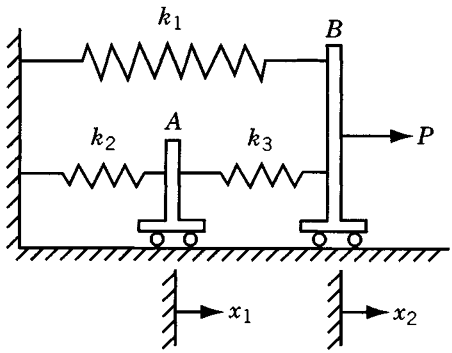

Figure 2.4 Spring–cart system.

potential energy (U)

$$= \text{strain energy of springs} - \text{work done by external forces}$$
$$= [\tfrac{1}{2} k_2 x_1^2 + \tfrac{1}{2} k_3 (x_2 - x_1)^2 + \tfrac{1}{2} k_1 x_2^2] - P x_2$$

The necessary conditions for the minimum of U are

$$\frac{\partial U}{\partial x_1} = k_2 x_1 - k_3(x_2 - x_1) = 0 \qquad (E_1)$$

$$\frac{\partial U}{\partial x_2} = k_3(x_2 - x_1) + k_1 x_2 - P = 0 \qquad (E_2)$$

The values of x_1 and x_2 corresponding to the equilibrium state, obtained by solving Eqs. (E_1) and (E_2) are given by

$$x_1^* = \frac{P k_3}{k_1 k_2 + k_1 k_3 + k_2 k_3}$$

$$x_2^* = \frac{P(k_2 + k_3)}{k_1 k_2 + k_1 k_3 + k_2 k_3}$$

The sufficiency conditions for the minimum at (x_1^*, x_2^*) can also be verified by testing the positive definiteness of the Hessian matrix of U. The Hessian matrix of U evaluated at (x_1^*, x_2^*) is

$$\mathbf{J}\big|_{(x_1^*, x_2^*)} = \begin{bmatrix} \dfrac{\partial^2 U}{\partial x_1^2} & \dfrac{\partial^2 U}{\partial x_1\, \partial x_2} \\ \dfrac{\partial^2 U}{\partial x_1\, \partial x_2} & \dfrac{\partial^2 U}{\partial x_2^2} \end{bmatrix}_{(x_1^*, x_2^*)} = \begin{bmatrix} k_2 + k_3 & -k_3 \\ -k_3 & k_1 + k_3 \end{bmatrix}$$

The determinants of the square submatrices of **J** are

$$J_1 = |k_2 + k_3| = k_2 + k_3 > 0$$

$$J_2 = \begin{vmatrix} k_2 + k_3 & -k_3 \\ -k_3 & k_1 + k_3 \end{vmatrix} = k_1k_2 + k_1k_3 + k_2k_3 > 0$$

since the spring constants are always positive. Thus the matrix **J** is positive definite and hence (x_1^*, x_2^*) corresponds to the minimum of potential energy.

2.3.1 Semidefinite Case

We now consider the problem of determining the sufficient conditions for the case when the Hessian matrix of the given function is semidefinite. In the case of a function of a single variable, the problem of determining the sufficient conditions for the case when the second derivative is zero was resolved quite easily. We simply investigated the higher-order derivatives in the Taylor's series expansion. A similar procedure can be followed for functions of n variables. However, the algebra becomes quite involved, and hence we rarely investigate the stationary points for sufficiency in actual practice. The following theorem, analogous to Theorem 2.2, gives the sufficiency conditions for the extreme points of a function of several variables.

Theorem 2.5 Let the partial derivatives of f of all orders up to the order $k \geq 2$ be continuous in the neighborhood of a stationary point $\mathbf{X}^*$, and

$$d^r f|_{\mathbf{X}=\mathbf{X}^*} = 0, \qquad 1 \leq r \leq k - 1$$

$$d^k f|_{\mathbf{X}=\mathbf{X}^*} \neq 0$$

so that $d^k f|_{\mathbf{X}=\mathbf{X}^*}$ is the first nonvanishing higher-order differential of f at $\mathbf{X}^*$. If k is even, then (i) $\mathbf{X}^*$ is a relative minimum if $d^k f|_{\mathbf{X}=\mathbf{X}^*}$ is positive definite, (ii) $\mathbf{X}^*$ is a relative maximum if $d^k f|_{\mathbf{X}=\mathbf{X}^*}$ is negative definite, and (iii) if $d^k f|_{\mathbf{X}=\mathbf{X}^*}$ is semidefinite (but not definite), no general conclusion can be drawn. On the other hand, if k is odd, $\mathbf{X}^*$ is not an extreme point of $f(\mathbf{X})$.

Proof: A proof similar to that of Theorem 2.2 can be found in Ref. [2.5].

2.3.2 Saddle Point

In the case of a function of two variables, $f(x,y)$, the Hessian matrix may be neither positive nor negative definite at a point (x^*,y^*) at which

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = 0$$

In such a case, the point (x^*,y^*) is called a *saddle point*. The characteristic of a saddle point is that it corresponds to a relative minimum or maximum of $f(x,y)$ with respect to one variable, say, x (the other variable being fixed at $y = y^*$) and a relative maximum or minimum of $f(x,y)$ with respect to the second variable y (the other variable being fixed at x^*).

As an example, consider the function $f(x,y) = x^2 - y^2$. For this function,

$$\frac{\partial f}{\partial x} = 2x \quad \text{and} \quad \frac{\partial f}{\partial y} = -2y$$

These first derivatives are zero at $x^* = 0$ and $y^* = 0$. The Hessian matrix of f at (x^*,y^*) is given by

$$\mathbf{J} = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}$$

Since this matrix is neither positive definite nor negative definite, the point $(x^* = 0, y^* = 0)$ is a saddle point. The function is shown graphically in Fig. 2.5. It can be seen that $f(x,y^*) = f(x,0)$ has a relative minimum and $f(x^*,y) = f(0,y)$ has a relative maximum at the saddle point (x^*,y^*). Saddle points may exist for functions of more than two variables also. The characteristic of the saddle point stated above still holds provided that x and y are interpreted as vectors in multidimensional cases.

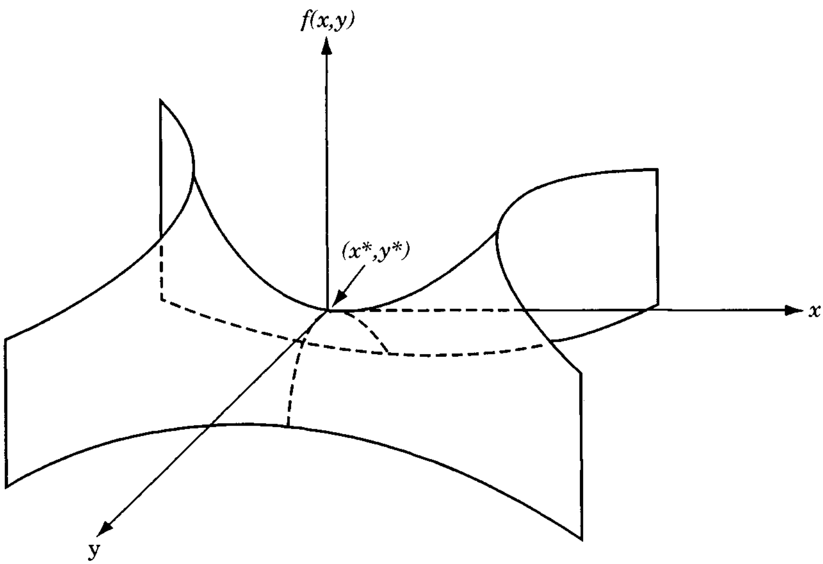

Figure 2.5 Saddle point of the function $f(x,y) = x^2 - y^2$.

Example 2.5 Find the extreme points of the function

$$f(x_1,x_2) = x_1^3 + x_2^3 + 2x_1^2 + 4x_2^2 + 6$$

SOLUTION The necessary conditions for the existence of an extreme point are

$$\frac{\partial f}{\partial x_1} = 3x_1^2 + 4x_1 = x_1(3x_1 + 4) = 0$$

$$\frac{\partial f}{\partial x_2} = 3x_2^2 + 8x_2 = x_2(3x_2 + 8) = 0$$

These equations are satisfied at the points

$$(0,0), \quad (0,-\tfrac{8}{3}), \quad (-\tfrac{4}{3},0), \quad \text{and} \quad (-\tfrac{4}{3},-\tfrac{8}{3})$$

To find the nature of these extreme points, we have to use the sufficiency conditions. The second-order partial derivatives of f are given by

$$\frac{\partial^2 f}{\partial x_1^2} = 6x_1 + 4$$

$$\frac{\partial^2 f}{\partial x_2^2} = 6x_2 + 8$$

$$\frac{\partial^2 f}{\partial x_1\, \partial x_2} = 0$$

The Hessian matrix of f is given by

$$\mathbf{J} = \begin{bmatrix} 6x_1 + 4 & 0 \\ 0 & 6x_2 + 8 \end{bmatrix}$$

If $J_1 = |6x_1 + 4|$ and $J_2 = \begin{vmatrix} 6x_1 + 4 & 0 \\ 0 & 6x_2 + 8 \end{vmatrix}$, the values of J_1 and J_2 and the nature of the extreme point are as given below.

Point $\mathbf{X}$	Value of J_1	Value of J_2	Nature of $\mathbf{J}$	Nature of $\mathbf{X}$	$f(\mathbf{X})$
$(0,0)$	+4	+32	Positive definite	Relative minimum	6
$(0,-\frac{8}{3})$	+4	−32	Indefinite	Saddle point	418/27
$(-\frac{4}{3},0)$	−4	−32	Indefinite	Saddle point	194/27
$(-\frac{4}{3}, -\frac{8}{3})$	−4	+32	Negative definite	Relative maximum	50/3

2.4 MULTIVARIABLE OPTIMIZATION WITH EQUALITY CONSTRAINTS

In this section we consider the optimization of continuous functions subjected to equality constraints:

$$\begin{aligned}&\text{Minimize } f = f(\mathbf{X})\\&\text{subject to}\\&g_j(\mathbf{X}) = 0, \qquad j = 1,2,\ldots,m\end{aligned} \tag{2.16}$$

where

$$\mathbf{X} = \begin{Bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{Bmatrix}$$

Here m is less than or equal to n; otherwise (if $m > n$), the problem becomes overdefined and, in general, there will be no solution. There are several methods available for the solution of this problem. The methods of direct substitution, constrained variation, and Lagrange multipliers are discussed in the following sections.

2.4.1 Solution by Direct Substitution

For a problem with n variables and m equality constraints, it is theoretically possible to solve simultaneously the m equality constraints and express any set of m variables in terms of the remaining $n - m$ variables. When these expressions are substituted into the original objective function, there results a new objective function involving only $n - m$ variables. The new objective function is not subjected to any constraint, and hence its optimum can be found by using the unconstrained optimization techniques discussed in Section 2.3.

This method of direct substitution, although it appears to be simple in theory, is not convenient from practical point of view. The reason for this is that the constraint equations will be nonlinear for most of practical problems, and often, it becomes impossible to solve them and express any m variables in terms of the remaining $n - m$ variables. However, the method of direct substitution might prove to be very simple and direct for solving simpler problems, as shown by the following example.

Example 2.6 Find the dimensions of a box of largest volume that can be inscribed in a sphere of unit radius.

SOLUTION Let the origin of the Cartesian coordinate system x_1, x_2, x_3 be at the center of the sphere and the sides of the box be $2x_1$, $2x_2$, and $2x_3$. The volume of the box is given by

$$f(x_1,x_2,x_3) = 8x_1x_2x_3 \tag{E$_1$}$$

Since the corners of the box lie on the surface of the sphere of unit radius, x_1, x_2, and x_3 have to satisfy the constraint

$$x_1^2 + x_2^2 + x_3^2 = 1 \tag{E$_2$}$$

This problem has three design variables and one equality constraint. Hence the equality constraint can be used to eliminate any one of the design variables from the objective function. If we choose to eliminate x_3, Eq. (E$_2$) gives

$$x_3 = (1 - x_1^2 - x_2^2)^{1/2} \tag{E$_3$}$$

Thus the objective function becomes

$$f(x_1,x_2) = 8x_1x_2(1 - x_1^2 - x_2^2)^{1/2} \tag{E$_4$}$$

which can be maximized as an unconstrained function in two variables.

The necessary conditions for the maximum of f give

$$\frac{\partial f}{\partial x_1} = 8x_2\left[(1 - x_1^2 - x_2^2)^{1/2} - \frac{x_1^2}{(1 - x_1^2 - x_2^2)^{1/2}}\right] = 0 \tag{E$_5$}$$

$$\frac{\partial f}{\partial x_2} = 8x_1\left[(1 - x_1^2 - x_2^2)^{1/2} - \frac{x_2^2}{(1 - x_1^2 - x_2^2)^{1/2}}\right] = 0 \tag{E$_6$}$$

Equations (E$_5$) and (E$_6$) can be simplified to obtain

$$1 - 2x_1^2 - x_2^2 = 0$$

$$1 - x_1^2 - 2x_2^2 = 0$$

from which it follows that $x_1^* = x_2^* = 1/\sqrt{3}$ and hence $x_3^* = 1/\sqrt{3}$. This solution gives the maximum volume of the box as

$$f_{\max} = \frac{8}{3\sqrt{3}}$$

To find whether the solution found corresponds to a maximum or a minimum, we apply the sufficiency conditions to $f(x_1,x_2)$ of Eq. (E$_4$). The second-order partial derivatives of f at (x_1^*,x_2^*) are given by

$$\frac{\partial^2 f}{\partial x_1^2} = -\frac{8x_1x_2}{(1 - x_1^2 - x_2^2)^{1/2}} - \frac{8x_2}{1 - x_1^2 - x_2^2}\left[\frac{x_1^3}{(1 - x_1^2 - x_2^2)^{1/2}} + 2x_1 (1 - x_1^2 - x_2^2)^{1/2}\right]$$

$$= -\frac{32}{\sqrt{3}} \text{ at } (x_1^*,x_2^*)$$

$$\frac{\partial^2 f}{\partial x_2^2} = -\frac{8x_1x_2}{(1 - x_1^2 - x_2^2)^{1/2}} - \frac{8x_1}{1 - x_1^2 - x_2^2}\left[\frac{x_2^3}{(1 - x_1^2 - x_2^2)^{1/2}} + 2x_2 (1 - x_1^2 - x_2^2)^{1/2}\right]$$

$$= -\frac{32}{\sqrt{3}} \text{ at } (x_1^*,x_2^*)$$

$$\frac{\partial^2 f}{\partial x_1\, \partial x_2} = 8(1 - x_1^2 - x_2^2)^{1/2} - \frac{8x_2^2}{(1 - x_1^2 - x_2^2)^{1/2}} - \frac{8x_1^2}{1 - x_1^2 - x_2^2} \cdot \left[(1 - x_1^2 - x_2^2)^{1/2} + \frac{x_2^2}{(1 - x_1^2 - x_2^2)^{1/2}}\right]$$

$$= -\frac{16}{\sqrt{3}} \text{ at } (x_1^*,x_2^*)$$

Since

$$\frac{\partial^2 f}{\partial x_1^2} < 0 \quad \text{and} \quad \frac{\partial^2 f}{\partial x_1^2}\frac{\partial^2 f}{\partial x_2^2} - \left(\frac{\partial^2 f}{\partial x_1\, \partial x_2}\right)^2 > 0$$

the Hessian matrix of f is negative definite at (x_1^*,x_2^*). Hence the point (x_1^*,x_2^*) corresponds to the maximum of f.

2.4.2 Solution by the Method of Constrained Variation

The basic idea used in the method of constrained variation is to find a closed-form expression for the first-order differential of $f(df)$ at all points at which the constraints $g_j(\mathbf{X}) = 0$, $j = 1,2,\ldots,m$, are satisfied. The desired optimum points are then obtained by setting the differential df equal to zero. Before presenting the general method, we indicate its salient features through the following simple problem with $n = 2$ and $m = 1$.

$$\text{Minimize } f(x_1,x_2) \tag{2.17}$$

subject to

$$g(x_1,x_2) = 0 \tag{2.18}$$

A necessary condition for f to have a minimum at some point (x_1^*,x_2^*) is that the total derivative of $f(x_1,x_2)$ with respect to x_1 must be zero at (x_1^*,x_2^*). By setting the total differential of $f(x_1,x_2)$ equal to zero, we obtain

$$df = \frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2 = 0 \tag{2.19}$$

Since $g(x_1^*,x_2^*) = 0$ at the minimum point, any variations dx_1 and dx_2 taken about the point (x_1^*,x_2^*) are called *admissible variations* provided that the new point lies on the constraint:

$$g(x_1^* + dx_1, x_2^* + dx_2) = 0 \tag{2.20}$$

The Taylor's series expansion of the function in Eq. (2.20) about the point (x_1^*,x_2^*) gives

$$g(x_1^* + dx_1, x_2^* + dx_2)$$
$$\simeq g(x_1^*,x^*2) + \frac{\partial g}{\partial x_1}(x_1^*,x_2^*)\, dx_1 + \frac{\partial g}{\partial x_2}(x_1^*,x_2^*)\, dx_2 = 0 \tag{2.21}$$

where dx_1 and dx_2 are assumed to be small. Since $g(x_1^*,x_2^*) = 0$, Eq. (2.21) reduces to

$$dg = \frac{\partial g}{\partial x_1} dx_1 + \frac{\partial g}{\partial x_2} dx_2 = 0 \quad \text{at} \quad (x_1^*,x_2^*) \tag{2.22}$$

Thus Eq. (2.22) has to be satisfied by all admissible variations. This is illustrated in Fig. 2.6, where PQ indicates the curve at each point of which Eq.

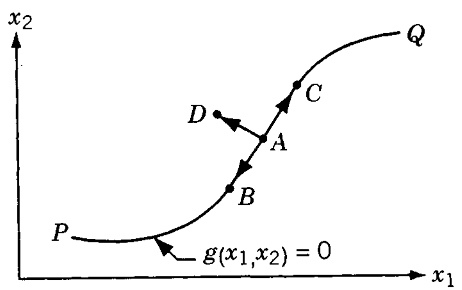

Figure 2.6 Variations about A.

(2.18) is satisfied. If A is taken as the base point (x_1^*,x_2^*), the variations in x_1 and x_2 leading to points B and C are called *admissible variations.* On the other hand, the variations in x_1 and x_2 representing point D are not admissible since point D does not lie on the constraint curve, $g(x_1,x_2) = 0$. Thus any set of variations (dx_1, dx_2) that does not satisfy Eq. (2.22) lead to points such as D which do not satisfy constraint Eq. (2.18).

Assuming that $\partial g/\partial x_2 \neq 0$, Eq. (2.22) can be rewritten as

$$dx_2 = -\frac{\partial g/\partial x_1}{\partial g/\partial x_2}(x_1^*,x_2^*)\, dx_1 \tag{2.23}$$

This relation indicates that once the variation in x_1 (dx_1) is chosen arbitrarily, the variation in x_2 (dx_2) is decided automatically in order to have dx_1 and dx_2 as a set of admissible variations. By substituting Eq. (2.23) in Eq. (2.19), we obtain

$$df = \left(\frac{\partial f}{\partial x_1} - \frac{\partial g/\partial x_1}{\partial g/\partial x_2}\frac{\partial f}{\partial x_2}\right)\bigg|_{(x_1^*,x_2^*)} dx_1 = 0 \tag{2.24}$$

The expression on the left-hand side is called the *constrained variation* of f. Note that Eq. (2.24) has to be satisfied for all values of dx_1. Since dx_1 can be chosen arbitrarily, Eq. (2.24) leads to

$$\left(\frac{\partial f}{\partial x_1}\frac{\partial g}{\partial x_2} - \frac{\partial f}{\partial x_2}\frac{\partial g}{\partial x_1}\right)\bigg|_{(x_1^*,x_2^*)} = 0 \tag{2.25}$$

Equation (2.25) represents a necessary condition in order to have (x_1^*,x_2^*) as an extreme point (minimum or maximum).

Example 2.7 A beam of uniform rectangular cross section is to be cut from a log having a circular cross section of diameter $2a$. The beam has to be used as a cantilever beam (the length is fixed) to carry a concentrated load at the free end. Find the dimensions of the beam that correspond to the maximum tensile (bending) stress carrying capacity.

SOLUTION From elementary strength of materials, we know that the tensile stress induced in a rectangular beam (σ) at any fiber located a distance y from the neutral axis is given by

$$\frac{\sigma}{y} = \frac{M}{I}$$

where M is the bending moment acting and I is the moment of inertia of the cross section about the x axis. If the width and depth of the rectangular beam

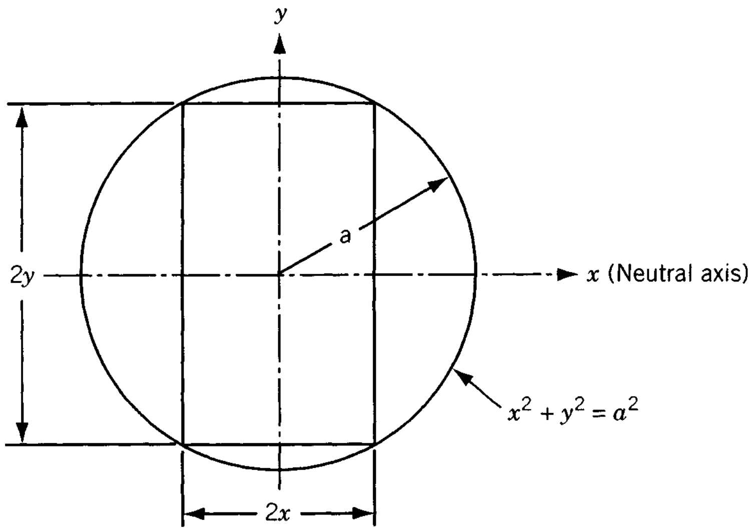

Figure 2.7 Cross section of the log.

shown in Fig. 2.7 are $2x$ and $2y$, respectively, the maximum tensile stress induced is given by

$$\sigma_{\max} = \frac{M}{I} y = \frac{My}{\frac{1}{12}(2x)(2y)^3} = \frac{3}{4}\frac{M}{xy^2}$$

Thus for any specified bending moment, the beam is said to have maximum tensile stress carrying capacity if the maximum induced stress ($\sigma_{\max}$) is a minimum. Hence we need to minimize k/xy^2 or maximize Kxy^2, where $k = 3M/4$ and $K = 1/k$, subject to the constraint

$$x^2 + y^2 = a^2$$

This problem has two variables and one constraint; hence Eq. (2.25) can be applied for finding the optimum solution. Since

$$f = kx^{-1}y^{-2} \tag{E$_1$}$$

$$g = x^2 + y^2 - a^2 \tag{E$_2$}$$

we have

$$\frac{\partial f}{\partial x} = -kx^{-2}y^{-2}$$

$$\frac{\partial f}{\partial y} = -2kx^{-1}y^{-3}$$

$$\frac{\partial g}{\partial x} = 2x$$

$$\frac{\partial g}{\partial y} = 2y$$

Equation (2.25) gives

$$-kx^{-2}y^{-2}(2y) + 2kx^{-1}y^{-3}(2x) = 0 \quad \text{at} \quad (x^*,y^*)$$

that is,

$$y^* = \sqrt{2}\, x^* \tag{E$_3$}$$

Thus the beam of maximum tensile stress carrying capacity has a depth of $\sqrt{2}$ times its breadth. The optimum values of x and y can be obtained from Eqs. (E_3) and (E_2) as

$$x^* = \frac{a}{\sqrt{3}} \quad \text{and} \quad y^* = \sqrt{2}\,\frac{a}{\sqrt{3}}$$

Necessary Conditions for a General Problem. The procedure indicated above can be generalized to the case of a problem in n variables with m constraints. In this case, each constraint equation $g_j(\mathbf{X}) = 0, j = 1,2,\ldots,m$, gives rise to a linear equation in the variations dx_i, $i = 1,2,\ldots,n$. Thus there will be in all m linear equations in n variations. Hence any m variations can be expressed in terms of the remaining $n - m$ variations. These expressions can be used to express the differential of the objective function, df, in terms of the $n - m$ independent variations. By letting the coefficients of the independent variations vanish in the equation $df = 0$, one obtains the necessary conditions for the constrained optimum of the given function. These conditions can be expressed as [2.6]

$$J\left(\frac{f,g_1,g_2,\ldots,g_m}{x_k,x_1,x_2,x_3,\ldots,x_m}\right) = \begin{vmatrix} \dfrac{\partial f}{\partial x_k} & \dfrac{\partial f}{\partial x_1} & \dfrac{\partial f}{\partial x_2} & \cdots & \dfrac{\partial f}{\partial x_m} \\ \dfrac{\partial g_1}{\partial x_k} & \dfrac{\partial g_1}{\partial x_1} & \dfrac{\partial g_1}{\partial x_2} & \cdots & \dfrac{\partial g_1}{\partial x_m} \\ \dfrac{\partial g_2}{\partial x_k} & \dfrac{\partial g_2}{\partial x_1} & \dfrac{\partial g_2}{\partial x_2} & \cdots & \dfrac{\partial g_2}{\partial x_m} \\ \vdots & & & & \\ \dfrac{\partial g_m}{\partial x_k} & \dfrac{\partial g_m}{\partial x_1} & \dfrac{\partial g_m}{\partial x_2} & \cdots & \dfrac{\partial g_m}{\partial x_m} \end{vmatrix} = 0 \tag{2.26}$$

where $k = m + 1, m + 2, \ldots, n$. It is to be noted that the variations of the first m variables $(dx_1, dx_2, \ldots, dx_m)$ have been expressed in terms of the variations of the remaining $n - m$ variables $(dx_{m+1}, dx_{m+2}, \ldots, dx_n)$ in deriving Eqs. (2.26). This implies that the following relation is satisfied:

$$J\left(\frac{g_1, g_2, \ldots, g_m}{x_1, x_2, \ldots, x_m}\right) \neq 0 \tag{2.27}$$

The $n - m$ equations given by Eqs. (2.26) represent the necessary conditions for the extremum of $f(\mathbf{X})$ under the m equality constraints, $g_j(\mathbf{X}) = 0$, $j = 1, 2, \ldots, m$.

Example 2.8

$$\text{Minimize } f(\mathbf{Y}) = \tfrac{1}{2}(y_1^2 + y_2^2 + y_3^2 + y_4^2) \tag{E$_1$}$$

subject to

$$g_1(\mathbf{Y}) = y_1 + 2y_2 + 3y_3 + 5y_4 - 10 = 0 \tag{E$_2$}$$

$$g_2(\mathbf{Y}) = y_1 + 2y_2 + 5y_3 + 6y_4 - 15 = 0 \tag{E$_3$}$$

SOLUTION This problem can be solved by applying the necessary conditions given by Eqs. (2.26). Since $n = 4$ and $m = 2$, we have to select two variables as independent variables. First we show that any arbitrary set of variables cannot be chosen as independent variables since the remaining (dependent) variables have to satisfy the condition of Eq. (2.27).

In terms of the notation of our equations, let us take the independent variables as

$$x_3 = y_3 \quad \text{and} \quad x_4 = y_4 \qquad \text{so that} \quad x_1 = y_1 \quad \text{and} \quad x_2 = y_2$$

Then the Jacobian of Eq. (2.27) becomes

$$J\left(\frac{g_1, g_2}{x_1, x_2}\right) = \begin{vmatrix} \dfrac{\partial g_1}{\partial y_1} & \dfrac{\partial g_1}{\partial y_2} \\[2ex] \dfrac{\partial g_2}{\partial y_1} & \dfrac{\partial g_2}{\partial y_2} \end{vmatrix} = \begin{vmatrix} 1 & 2 \\ 1 & 2 \end{vmatrix} = 0$$

and hence the necessary conditions of Eqs. (2.26) cannot be applied.

Next, let us take the independent variables as $x_3 = y_2$ and $x_4 = y_4$ so that $x_1 = y_1$ and $x_2 = y_3$. Then the Jacobian of Eq. (2.27) becomes

$$J\left(\frac{g_1,g_2}{x_1,x_2}\right) = \begin{vmatrix} \dfrac{\partial g_1}{\partial y_1} & \dfrac{\partial g_1}{\partial y_3} \\ \dfrac{\partial g_2}{\partial y_1} & \dfrac{\partial g_2}{\partial y_3} \end{vmatrix} = \begin{vmatrix} 1 & 3 \\ 1 & 5 \end{vmatrix} = 2 \neq 0$$

and hence the necessary conditions of Eqs. (2.26) can be applied. Equations (2.26) give for $k = m + 1 = 3$,

$$\begin{vmatrix} \dfrac{\partial f}{\partial x_3} & \dfrac{\partial f}{\partial x_1} & \dfrac{\partial f}{\partial x_2} \\ \dfrac{\partial g_1}{\partial x_3} & \dfrac{\partial g_1}{\partial x_1} & \dfrac{\partial g_1}{\partial x_2} \\ \dfrac{\partial g_2}{\partial x_3} & \dfrac{\partial g_2}{\partial x_1} & \dfrac{\partial g_2}{\partial x_2} \end{vmatrix} = \begin{vmatrix} \dfrac{\partial f}{\partial y_2} & \dfrac{\partial f}{\partial y_1} & \dfrac{\partial f}{\partial y_3} \\ \dfrac{\partial g_1}{\partial y_2} & \dfrac{\partial g_1}{\partial y_1} & \dfrac{\partial g_1}{\partial y_3} \\ \dfrac{\partial g_2}{\partial y_2} & \dfrac{\partial g_2}{\partial y_1} & \dfrac{\partial g_2}{\partial y_3} \end{vmatrix}$$

$$= \begin{vmatrix} y_2 & y_1 & y_3 \\ 2 & 1 & 3 \\ 2 & 1 & 5 \end{vmatrix}$$

$$= y_2(5 - 3) - y_1(10 - 6) + y_3(2 - 2)$$

$$= 2y_2 - 4y_1 = 0 \tag{E$_4$}$$

and for $k = m + 2 = n = 4$,

$$\begin{vmatrix} \dfrac{\partial f}{\partial x_4} & \dfrac{\partial f}{\partial x_1} & \dfrac{\partial f}{\partial x_2} \\ \dfrac{\partial g_1}{\partial x_4} & \dfrac{\partial g_1}{\partial x_1} & \dfrac{\partial g_1}{\partial x_2} \\ \dfrac{\partial g_2}{\partial x_4} & \dfrac{\partial g_2}{\partial x_1} & \dfrac{\partial g_2}{\partial x_2} \end{vmatrix} = \begin{vmatrix} \dfrac{\partial f}{\partial y_4} & \dfrac{\partial f}{\partial y_1} & \dfrac{\partial f}{\partial y_3} \\ \dfrac{\partial g_1}{\partial y_4} & \dfrac{\partial g_1}{\partial y_1} & \dfrac{\partial g_1}{\partial y_3} \\ \dfrac{\partial g_2}{\partial y_4} & \dfrac{\partial g_2}{\partial y_1} & \dfrac{\partial g_2}{\partial y_3} \end{vmatrix}$$

$$= \begin{vmatrix} y_4 & y_1 & y_3 \\ 5 & 1 & 3 \\ 6 & 1 & 5 \end{vmatrix}$$

$$= y_4(5 - 3) - y_1(25 - 18) + y_3(5 - 6)$$

$$= 2y_4 - 7y_1 - y_3 = 0 \tag{E$_5$}$$

Equations (E_4) and (E_5) give the necessary conditions for the minimum or the maximum of f as

$$\begin{aligned} y_1 &= \tfrac{1}{2} y_2 \\ y_3 &= 2y_4 - 7y_1 = 2y_4 - \tfrac{7}{2} y_2 \end{aligned} \tag{E$_6$}$$

When Eqs. (E_6) are substituted, Eqs. (E_2) and (E_3) take the form

$$-8y_2 + 11y_4 = 10$$

$$-15y_2 + 16y_4 = 15$$

from which the desired optimum solution can be obtained as

$$y_1^* = -\tfrac{5}{74}$$

$$y_2^* = -\tfrac{5}{37}$$

$$y_3^* = \tfrac{155}{74}$$

$$y_4^* = \tfrac{30}{37}$$

Sufficiency Conditions for a General Problem. By eliminating the first m variables, using the m equality constraints (this is possible, at least in theory), the objective function f can be made to depend only on the remaining variables, $x_{m+1}, x_{m+2}, \ldots, x_n$. Then the Taylor's series expansion of f, in terms of these variables, about the extreme point $\mathbf{X}^*$ gives

$$f(\mathbf{X}^* + d\mathbf{X}) \simeq f(\mathbf{X}^*) + \sum_{i=m+1}^{n} \left(\frac{\partial f}{\partial x_i}\right)_g dx_i + \frac{1}{2!} \sum_{i=m+1}^{n} \sum_{j=m+1}^{n} \left(\frac{\partial^2 f}{\partial x_i\, \partial x_j}\right)_g dx_i\, dx_j \tag{2.28}$$

where $(\partial f/\partial x_i)_g$ is used to denote the partial derivative of f with respect to x_i (holding all the other variables $x_{m+1}, x_{m+2}, \ldots, x_{i-1}, x_{i+1}, x_{i+2}, \ldots, x_n$ constant) when $x_1, x_2, \ldots, x_m$ are allowed to change so that the constraints $g_j(\mathbf{X}^* + d\mathbf{X}) = 0$, $j = 1,2,\ldots,m$, are satisfied; the second derivative, $(\partial^2 f/\partial x_i\, \partial x_j)_g$, is used to denote a similar meaning.

As an example, consider the problem of minimizing

$$f(\mathbf{X}) = f(x_1,x_2,x_3)$$

subject to the only constraint

$$g_1(\mathbf{X}) = x_1^2 + x_2^2 + x_3^2 - 8 = 0$$

Since $n = 3$ and $m = 1$ in this problem, one can think of any of the m variables, say x_1, to be dependent and the remaining $n - m$ variables, namely x_2 and x_3, to be independent. Here the constrained partial derivative $(\partial f/\partial x_2)_g$, for example, means the rate of change of f with respect to x_2 (holding the other independent variable x_3 constant) and at the same time allowing x_1 to change about $\mathbf{X}^*$ so as to satisfy the constraint $g_1(\mathbf{X}) = 0$. In the present case, this means that dx_1 has to be chosen to satisfy the relation

$$g_1(\mathbf{X}^* + d\mathbf{X}) \simeq g_1(\mathbf{X}^*) + \frac{\partial g_1}{\partial x_1}(\mathbf{X}^*)\, dx_1 + \frac{\partial g_1}{\partial x_2}(\mathbf{X}^*)\, dx_2 + \frac{\partial g_1}{\partial x_3}(\mathbf{X}^*)\, dx_3 = 0$$

that is,

$$2x_1^*\, dx_1 + 2x_2^*\, dx_2 = 0$$

since $g_1(\mathbf{X}^*) = 0$ at the optimum point and $dx_3 = 0$ (x_3 is held constant).

Notice that $(\partial f/\partial x_i)_g$ has to be zero for $i = m + 1, m + 2, \ldots, n$ since the dx_i appearing in Eq. (2.28) are all independent. Thus the necessary conditions for the existence of constrained optimum at $\mathbf{X}^*$ can also be expressed as

$$\left(\frac{\partial f}{\partial x_i}\right)_g = 0, \qquad i = m + 1, m + 2, \ldots, n \tag{2.29}$$

Of course, with little manipulation, one can show that Eqs. (2.29) are nothing but Eqs. (2.26). Further, as in the case of optimization of a multivariable function with no constraints, one can see that a sufficient condition for $\mathbf{X}^*$ to be a constrained relative minimum (maximum) is that the quadratic form Q defined by

$$Q = \sum_{i=m+1}^{n} \sum_{j=m+1}^{n} \left(\frac{\partial^2 f}{\partial x_i\, \partial x_j}\right)_g dx_i\, dx_j \tag{2.30}$$

is positive (negative) for all nonvanishing variations dx_i. As in Theorem 2.4, the matrix

$$\begin{bmatrix} \left(\dfrac{\partial^2 f}{\partial x_{m+1}^2}\right)_g & \left(\dfrac{\partial^2 f}{\partial x_{m+1}\, \partial x_{m+2}}\right)_g & \cdots & \left(\dfrac{\partial^2 f}{\partial x_{m+1}\, \partial x_n}\right)_g \\ \vdots & & & \\ \left(\dfrac{\partial^2 f}{\partial x_n\, \partial x_{m+1}}\right)_g & \left(\dfrac{\partial^2 f}{\partial x_n\, \partial x_{m+2}}\right)_g & \cdots & \left(\dfrac{\partial^2 f}{\partial x_n^2}\right)_g \end{bmatrix}$$

has to be positive (negative) definite to have Q positive (negative) for all choices of dx_i. It is evident that computation of the constrained derivatives $(\partial^2 f/\partial x_i \partial x_j)_g$ is a difficult task and may be prohibitive for problems with more than three constraints. Thus the method of constrained variation, although it appears to be simple in theory, is very difficult to apply since the necessary conditions themselves involve evaluation of determinants of order $m + 1$. This is the reason that the method of Lagrange multipliers, discussed in the following section, is more commonly used to solve a multivariable optimization problem with equality constraints.

2.4.3 Solution by the Method of Lagrange Multipliers

The basic features of the Lagrange multiplier method is given initially for a simple problem of two variables with one constraint. The extension of the method to a general problem of n variables with m constraints is given later.

Problem with Two Variables and One Constraint. Consider the problem:

$$\text{Minimize } f(x_1, x_2) \tag{2.31}$$

subject to

$$g(x_1, x_2) = 0$$

For this problem, the necessary condition for the existence of an extreme point at $\mathbf{X} = \mathbf{X}^*$ was found in Section 2.4.2 to be

$$\left.\left(\frac{\partial f}{\partial x_1} - \frac{\partial f/\partial x_2}{\partial g/\partial x_2}\frac{\partial g}{\partial x_1}\right)\right|_{(x_1^*, x_2^*)} = 0 \tag{2.32}$$

By defining a quantity λ, called the *Lagrange multiplier*, as

$$\lambda = -\left.\left(\frac{\partial f/\partial x_2}{\partial g/\partial x_2}\right)\right|_{(x_1^*, x_2^*)} \tag{2.33}$$

Equation (2.32) can be expressed as

$$\left.\left(\frac{\partial f}{\partial x_1} + \lambda\frac{\partial g}{\partial x_1}\right)\right|_{(x_1^*, x_2^*)} = 0 \tag{2.34}$$

and Eq. (2.33) can be written as

$$\left.\left(\frac{\partial f}{\partial x_2} + \lambda\frac{\partial g}{\partial x_2}\right)\right|_{(x_1^*, x_2^*)} = 0 \tag{2.35}$$

In addition, the constraint equation has to be satisfied at the extreme point, that is,

$$g(x_1, x_2)|_{(x_1^*, x_2^*)} = 0 \tag{2.36}$$

Thus Eqs. (2.34) to (2.36) represent the necessary conditions for the point (x_1^*, x_2^*) to be an extreme point.

Notice that the partial derivative $(\partial g/\partial x_2)|_{(x_1^*, x_2^*)}$ has to be nonzero to be able to define λ by Eq. (2.33). This is because the variation dx_2 was expressed in terms of dx_1 in the derivation of Eq. (2.32) [see Eq. (2.23)]. On the other hand, if we choose to express dx_1 in terms of dx_2, we would have obtained the requirement that $(\partial g/\partial x_1)|_{(x_1^*, x_2^*)}$ be nonzero to define λ. Thus the derivation of the necessary conditions by the method of Lagrange multipliers requires that at least one of the partial derivatives of $g(x_1, x_2)$ be nonzero at an extreme point.

The necessary conditions given by Eqs. (2.34) to (2.36) are more commonly generated by constructing a function L, known as the Lagrange function, as

$$L(x_1, x_2, \lambda) = f(x_1, x_2) + \lambda g(x_1, x_2) \tag{2.37}$$

By treating L as a function of the three variables x_1, x_2, and λ, the necessary conditions for its extremum are given by

$$\begin{aligned}
\frac{\partial L}{\partial x_1}(x_1, x_2, \lambda) &= \frac{\partial f}{\partial x_1}(x_1, x_2) + \lambda \frac{\partial g}{\partial x_1}(x_1, x_2) = 0 \\
\frac{\partial L}{\partial x_2}(x_1, x_2, \lambda) &= \frac{\partial f}{\partial x_2}(x_1, x_2) + \lambda \frac{\partial g}{\partial x_2}(x_1, x_2) = 0 \\
\frac{\partial L}{\partial \lambda}(x_1, x_2, \lambda) &= g(x_1, x_2) = 0
\end{aligned} \tag{2.38}$$

Equations (2.38) can be seen to be same as Eqs. (2.34) to (2.36). The sufficiency conditions are given later

Example 2.9 Find the solution of Example 2.7 using the Lagrange multiplier method:

$$\text{Minimize } f(x,y) = kx^{-1}y^{-2}$$

subject to

$$g(x,y) = x^2 + y^2 - a^2 = 0$$

SOLUTION The Lagrange function is

$$L(x,y,\lambda) = f(x,y) + \lambda g(x,y) = kx^{-1}y^{-2} + \lambda(x^2 + y^2 - a^2)$$

The necessary conditions for the minimum of $f(x, y)$ [Eqs. (2.38)] give

$$\frac{\partial L}{\partial x} = -kx^{-2}y^{-2} + 2x\lambda = 0 \tag{E$_1$}$$

$$\frac{\partial L}{\partial y} = -2kx^{-1}y^{-3} + 2y\lambda = 0 \tag{E$_2$}$$

$$\frac{\partial L}{\partial \lambda} = x^2 + y^2 - a^2 = 0 \tag{E$_3$}$$

Equations (E_1) and (E_2) yield

$$2\lambda = \frac{k}{x^3y^2} = \frac{2k}{xy^4}$$

from which the relation $x^* = (1/\sqrt{2})\, y^*$ can be obtained. This relation, along with Eq. (E_3), gives the optimum solution as

$$x^* = \frac{a}{\sqrt{3}} \quad \text{and} \quad y^* = \sqrt{2}\,\frac{a}{\sqrt{3}}$$

Necessary Conditions for a General Problem. The equations derived above can be extended to the case of a general problem with n variables and m equality constraints:

$$\text{Minimize } f(\mathbf{X})$$

subject to (2.39)

$$g_j(\mathbf{X}) = 0, \qquad j = 1,2,\ldots,m$$

The Lagrange function, L, in this case is defined by introducing one Lagrange multiplier λ_j for each constraint $g_j(\mathbf{X})$ as

$$L(x_1,x_2,\ldots,x_n,\lambda_1,\lambda_2,\ldots,\lambda_m) = f(\mathbf{X}) + \lambda_1 g_1(\mathbf{X}) + \lambda_2 g_2(\mathbf{X}) + \cdots + \lambda_m g_m(\mathbf{X}) \tag{2.40}$$

By treating L as a function of the $n + m$ unknowns, $x_1, x_2, \ldots, x_n, \lambda_1, \lambda_2, \ldots, \lambda_m$, the necessary conditions for the extremum of L, which also correspond to the solution of the original problem stated in Eq. (2.39), are given by

$$\frac{\partial L}{\partial x_i} = \frac{\partial f}{\partial x_i} + \sum_{j=1}^{m} \lambda_j \frac{\partial g_j}{\partial x_i} = 0, \qquad i = 1,2,\ldots,n \tag{2.41}$$

$$\frac{\partial L}{\partial \lambda_j} = g_j(\mathbf{X}) = 0, \qquad j = 1,2,\ldots,m \tag{2.42}$$

Equations (2.41) and (2.42) represent $n + m$ equations in terms of the $n + m$ unknowns, x_i and λ_j. The solution of Eqs. (2.41) and (2.42) gives

$$\mathbf{X}^* = \begin{Bmatrix} x_1^* \\ x_2^* \\ \vdots \\ x_n^* \end{Bmatrix} \quad \text{and} \quad \boldsymbol{\lambda}^* = \begin{Bmatrix} \lambda_1^* \\ \lambda_2^* \\ \vdots \\ \lambda_m^* \end{Bmatrix}$$

The vector $\mathbf{X}^*$ corresponds to the relative constrained minimum of $f(\mathbf{X})$ (sufficient conditions are to be verified) while the vector $\boldsymbol{\lambda}^*$ provides the sensitivity information, as discussed in the next subsection.

Sufficiency Conditions for a General Problem. A sufficient condition for $f(\mathbf{X})$ to have a constrained relative minimum at $\mathbf{X}^*$ is given by the following theorem.

Theorem 2.6: Sufficient Condition A sufficient condition for $f(\mathbf{X})$ to have a relative minimum at $\mathbf{X}^*$ is that the quadratic, Q, defined by

$$Q = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 L}{\partial x_i \, \partial x_j} \, dx_i \, dx_j \tag{2.43}$$

evaluated at $\mathbf{X} = \mathbf{X}^*$ must be positive definite for all values of $d\mathbf{X}$ for which the constraints are satisfied.

Proof: The proof is similar to that of Theorem 2.4.

Notes:

1. If

$$Q = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 L}{\partial x_i \, \partial x_j} (\mathbf{X}^*, \boldsymbol{\lambda}^*) \, dx_i \, dx_j$$

   is negative for all choices of the admissible variations dx_i, $\mathbf{X}^*$ will be a constrained maximum of $f(\mathbf{X})$.
2. It has been shown by Hancock [2.1] that a necessary condition for the quadratic form Q, defined by Eq. (2.43), to be positive (negative) definite for all admissible variations $d\mathbf{X}$ is that each root of the polynomial

z_i, defined by the following determinantal equation, be positive (negative):

$$\begin{vmatrix} L_{11} - z & L_{12} & L_{13} & \cdots & L_{1n} & g_{11} & g_{21} & \cdots & g_{m1} \\ L_{21} & L_{22} - z & L_{23} & \cdots & L_{2n} & g_{12} & g_{22} & \cdots & g_{m2} \\ \vdots & & & & & & & & \\ L_{n1} & L_{n2} & L_{n3} & \cdots & L_{nn} - z & g_{1n} & g_{2n} & \cdots & g_{mn} \\ g_{11} & g_{12} & g_{13} & \cdots & g_{1n} & 0 & 0 & \cdots & 0 \\ g_{21} & g_{22} & g_{23} & \cdots & g_{2n} & 0 & 0 & \cdots & 0 \\ \vdots & & & & & & & & \\ g_{m1} & g_{m2} & g_{m3} & \cdots & g_{mn} & 0 & 0 & \cdots & 0 \end{vmatrix} = 0 \tag{2.44}$$

where

$$L_{ij} = \frac{\partial^2 L}{\partial x_i \, \partial x_j} (\mathbf{X}^*, \lambda^*) \tag{2.45}$$

$$g_{ij} = \frac{\partial g_i}{\partial x_j} (\mathbf{X}^*) \tag{2.46}$$

3. Equation (2.44), on expansion, leads to an $(n - m)$th-order polynomial in z. If some of the roots of this polynomial are positive while the others are negative, the point $\mathbf{X}^*$ is not an extreme point.

The application of the necessary and sufficient conditions in the Lagrange multiplier method is illustrated with the help of the following example.

Example 2.10 Find the dimensions of a cylindrical tin (with top and bottom) made up of sheet metal to maximize its volume such that the total surface area is equal to $A_0 = 24\pi$.

SOLUTION If x_1 and x_2 denote the radius of the base and length of the tin, respectively, the problem can be stated as:

$$\text{Maximize } f(x_1,x_2) = \pi x_1^2 x_2$$

subject to

$$2\pi x_1^2 + 2\pi x_1 x_2 = A_0 = 24\pi$$

The Lagrange function is

$$L(x_1,x_2,\lambda) = \pi x_1^2 x_2 + \lambda(2\pi x_1^2 + 2\pi x_1 x_2 - A_0)$$

and the necessary conditions for the maximum of f give

$$\frac{\partial L}{\partial x_1} = 2\pi x_1 x_2 + 4\pi\lambda x_1 + 2\pi\lambda x_2 = 0 \tag{E$_1$}$$

$$\frac{\partial L}{\partial x_2} = \pi x_1^2 + 2\pi\lambda x_1 = 0 \tag{E$_2$}$$

$$\frac{\partial L}{\partial \lambda} = 2\pi x_1^2 + 2\pi x_1 x_2 - A_0 = 0 \tag{E$_3$}$$

Equations (E_1) and (E_2) lead to

$$\lambda = -\frac{x_1 x_2}{2x_1 + x_2} = -\frac{1}{2} x_1$$

that is,

$$x_1 = \tfrac{1}{2} x_2 \tag{E$_4$}$$

and Eqs. (E_3) and (E_4) give the desired solution as

$$x_1^* = \left(\frac{A_0}{6\pi}\right)^{1/2}, \quad x_2^* = \left(\frac{2A_0}{3\pi}\right)^{1/2}, \quad \text{and} \quad \lambda^* = -\left(\frac{A_0}{24\pi}\right)^{1/2}$$

This gives the maximum value of f as

$$f^* = \left(\frac{A_0^3}{54\pi}\right)^{1/2}$$

If $A_0 = 24\pi$, the optimum solution becomes

$$x_1^* = 2, \quad x_2^* = 4, \quad \lambda^* = -1, \quad \text{and} \quad f^* = 16\pi$$

To see that this solution really corresponds to the maximum of f, we apply the sufficiency condition of Eq. (2.44). In this case

$$L_{11} = \left.\frac{\partial^2 L}{\partial x_1^2}\right|_{(\mathbf{X}^*,\lambda^*)} = 2\pi x_2^* + 4\pi\lambda^* = 4\pi$$

$$L_{12} = \left.\frac{\partial^2 L}{\partial x_1 \, \partial x_2}\right|_{(\mathbf{X}^*, \lambda^*)} = L_{21} = 2\pi x_1^* + 2\pi\lambda^* = 2\pi$$

$$L_{22} = \left.\frac{\partial^2 L}{\partial x_2^2}\right|_{(\mathbf{X}^*, \lambda^*)} = 0$$

$$g_{11} = \left.\frac{\partial g_1}{\partial x_2}\right|_{(\mathbf{X}^*, \lambda^*)} = 4\pi x_1^* + 2\pi x_2^* = 16\pi$$

$$g_{12} = \left.\frac{\partial g_1}{\partial x_2}\right|_{(\mathbf{X}^*, \lambda^*)} = 2\pi x_1^* = 4\pi$$

Thus Eq. (2.44) becomes

$$\begin{vmatrix} 4\pi - z & 2\pi & 16\pi \\ 2\pi & 0 - z & 4\pi \\ 16\pi & 4\pi & 0 \end{vmatrix} = 0$$

that is,

$$272\pi^2 z + 192\pi^3 = 0$$

This gives

$$z = -\tfrac{12}{17}\pi$$

Since the value of z is negative, the point (x_1^*, x_2^*) corresponds to the maximum of f.

Interpretation of the Lagrange Multipliers. To find the physical meaning of the Lagrange multipliers, consider the following optimization problem involving only a single equality constraint:

$$\text{Minimize } f(\mathbf{X}) \tag{2.47}$$

subject to

$$\underset{\sim}{g}(\mathbf{X}) = b \quad \text{or} \quad g(\mathbf{X}) = b - \underset{\sim}{g}(\mathbf{X}) = 0 \tag{2.48}$$

where b is a constant. The necessary conditions to be satisfied for the solution of the problem are

$$\frac{\partial f}{\partial x_i} + \lambda \frac{\partial g}{\partial x_i} = 0, \qquad i = 1,2,\ldots,n \tag{2.49}$$

$$g = 0 \tag{2.50}$$

Let the solution of Eqs. (2.49) and (2.50) be given by $\mathbf{X}^*$, λ^*, and $f^* = f(\mathbf{X}^*)$. Suppose that we want to find the effect of a small relaxation or tightening of the constraint on the optimum value of the objective function (i.e., we want to find the effect of a small change in b on f^*). For this we differentiate Eq. (2.48) to obtain

$$db - d\tilde{g} = 0$$

or

$$db = d\tilde{g} = \sum_{i=1}^{n} \frac{\partial \tilde{g}}{\partial x_i} dx_i \tag{2.51}$$

Equation (2.49) can be rewritten as

$$\frac{\partial f}{\partial x_i} + \lambda \frac{\partial g}{\partial x_i} = \frac{\partial f}{\partial x_i} - \lambda \frac{\partial \tilde{g}}{\partial x_i} = 0 \tag{2.52}$$

or

$$\frac{\partial \tilde{g}}{\partial x_i} = \frac{\partial f/\partial x_i}{\lambda}, \qquad i = 1,2,\ldots,n \tag{2.53}$$

Substituting Eq. (2.53) into Eq. (2.51), we obtain

$$db = \sum_{i=1}^{n} \frac{1}{\lambda} \frac{\partial f}{\partial x_i} dx_i = \frac{df}{\lambda} \tag{2.54}$$

since

$$df = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} dx_i \tag{2.55}$$

Equation (2.54) gives

$$\lambda = \frac{df}{db} \quad \text{or} \quad \lambda^* = \frac{df^*}{db} \tag{2.56}$$

or

$$df^* = \lambda^* \, db \tag{2.57}$$

Thus λ^* denotes the sensitivity (or rate of change) of f with respect to b or the marginal or incremental change in f^* with respect to b at x^*. In other words, λ^* indicates how tightly the constraint is binding at the optimum point. Depending on the value of λ^* (positive, negative, or zero), the following physical meaning can be attributed to λ^*:

1. $\lambda^* > 0$. In this case, a unit decrease in b is positively valued since one gets a smaller minimum value of the objective function f. In fact, the decrease in f^* will be exactly equal to λ^* since $df = \lambda^*(-1) = -\lambda^* < 0$. Hence λ^* may be interpreted as the marginal gain (further reduction) in f^* due to the tightening of the constraint. On the other hand, if b is increased by 1 unit, f will also increase to a new optimum level, with the amount of increase in f^* being determined by the magnitude of λ^* since $df = \lambda^*(+1) > 0$. In this case, λ^* may be thought of as the marginal cost (increase) in f^* due to the relaxation of the constraint.
2. $\lambda^* < 0$. Here a unit increase in b is positively valued. This means that it decreases the optimum value of f. In this case the marginal gain (reduction) in f^* due to a relaxation of the constraint by 1 unit is determined by the value of λ^* as $df^* = \lambda^*(+1) < 0$. If b is decreased by 1 unit, the marginal cost (increase) in f^* by the tightening of the constraint is $df^* = \lambda^*(-1) > 0$ since, in this case, the minimum value of the objective function increases.
3. $\lambda^* = 0$. In this case, any incremental change in b has absolutely no effect on the optimum value of f and hence the constraint will not be binding. This means that the optimization of f subject to $g = 0$ leads to the same optimum point $\mathbf{X}^*$ as with the unconstrained optimization of f.

In economics and operations research, Lagrange multipliers are known as *shadow prices* of the constraints since they indicate the changes in optimal value of the objective function per unit change in the right-hand side of the equality constraints.

Example 2.11 Find the maximum of the function $f(\mathbf{X}) = 2x_1 + x_2 + 10$ subject to $g(\mathbf{X}) = x_1 + 2x_2^2 = 3$ using the Lagrange multiplier method. Also find the effect of changing the right-hand side of the constraint on the optimum value of f.

SOLUTION The Lagrange function is given by

$$L(\mathbf{X},\lambda) = 2x_1 + x_2 + 10 + \lambda(3 - x_1 - 2x_2^2) \tag{E$_1$}$$

The necessary conditions for the solution of the problem are

$$\begin{aligned} \frac{\partial L}{\partial x_1} &= 2 - \lambda = 0 \\ \frac{\partial L}{\partial x_2} &= 1 - 4\lambda x_2 = 0 \\ \frac{\partial L}{\partial \lambda} &= 3 - x_1 - 2x_2^2 = 0 \end{aligned} \tag{E$_2$}$$

The solution of Eqs. (E_2) is

$$\mathbf{X}^* = \begin{Bmatrix} x_1^* \\ x_2^* \end{Bmatrix} = \begin{Bmatrix} 2.97 \\ 0.13 \end{Bmatrix}$$
$$\lambda^* = 2.0 \tag{E$_3$}$$

The application of the sufficiency condition of Eq. (2.52) yields

$$\begin{vmatrix} L_{11} - z & L_{12} & g_{11} \\ L_{21} & L_{22} - z & g_{12} \\ g_{11} & g_{12} & 0 \end{vmatrix} = 0$$

$$\begin{vmatrix} -z & 0 & -1 \\ 0 & -4\lambda - z & -4x_2 \\ -1 & -4x_2 & 0 \end{vmatrix} = \begin{vmatrix} -z & 0 & -1 \\ 0 & -8 - z & -0.52 \\ -1 & -0.52 & 0 \end{vmatrix} = 0$$

$$0.2704z + 8 + z = 0$$
$$z = -6.2972$$

Hence $\mathbf{X}^*$ will be a maximum of f with $f^* = f(\mathbf{X}^*) = 16.07$.

One procedure for finding the effect on f^* of changes in the value of b (right-hand side of the constraint) would be to solve the problem all over with the new value of b. Another procedure would involve the use of the value of λ^*. When the original constraint is tightened by 1 unit (i.e., $db = -1$), Eq. (2.57) gives

$$df^* = \lambda^* \, db = 2(-1) = -2$$

Thus the new value of f^* is $f^* + df^* = 14.07$. On the other hand, if we relax the original constraint by 2 units (i.e., $db = 2$), we obtain

$$df^* = \lambda^* \, db = 2(+2) = 4$$

and hence the new value of f^* is $f^* + df^* = 20.07$.

2.5 MULTIVARIABLE OPTIMIZATION WITH INEQUALITY CONSTRAINTS

This section is concerned with the solution of the following problem:

$$\text{Minimize } f(\mathbf{X})$$

subject to

$$g_j(\mathbf{X}) \le 0, \qquad j = 1,2,\ldots,m \tag{2.58}$$

The inequality constraints in Eq. (2.58) can be transformed to equality constraints by adding nonnegative slack variables, y_j^2, as

$$g_j(\mathbf{X}) + y_j^2 = 0, \qquad j = 1,2,\ldots,m \tag{2.59}$$

where the values of the slack variables are yet unknown. The problem now becomes

$$\text{Minimize } f(\mathbf{X})$$

subject to

$$G_j(\mathbf{X},\mathbf{Y}) = g_j(\mathbf{X}) + y_j^2 = 0, \qquad j = 1,2,\ldots,m \tag{2.60}$$

where $\mathbf{Y} = \begin{Bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{Bmatrix}$ is the vector of slack variables.

This problem can be solved conveniently by the method of Lagrange multipliers. For this, we construct the Lagrange function L as

$$L(\mathbf{X},\mathbf{Y},\boldsymbol{\lambda}) = f(\mathbf{X}) + \sum_{j=1}^{m} \lambda_j G_j(\mathbf{X},\mathbf{Y}) \tag{2.61}$$

where $\boldsymbol{\lambda} = \begin{Bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_m \end{Bmatrix}$ is the vector of Lagrange multipliers. The stationary points of the Lagrange function can be found by solving the following equations (necessary conditions):

$$\frac{\partial L}{\partial x_i}(\mathbf{X},\mathbf{Y},\boldsymbol{\lambda}) = \frac{\partial f}{\partial x_i}(\mathbf{X}) + \sum_{j=1}^{m} \lambda_j \frac{\partial g_j}{\partial x_i}(\mathbf{X}) = 0, \qquad i = 1,2,\ldots,n \tag{2.62}$$

$$\frac{\partial L}{\partial \lambda_j}(\mathbf{X},\mathbf{Y},\boldsymbol{\lambda}) = G_j(\mathbf{X},\mathbf{Y}) = g_j(\mathbf{X}) + y_j^2 = 0, \qquad j = 1,2,\ldots,m \tag{2.63}$$

$$\frac{\partial L}{\partial y_j}(\mathbf{X},\mathbf{Y},\boldsymbol{\lambda}) = 2\lambda_j y_j = 0, \qquad j = 1,2,\ldots,m \tag{2.64}$$

It can be seen that Eqs. (2.62) to (2.64) represent $(n + 2m)$ equations in the $(n + 2m)$ unknowns, $\mathbf{X}$, $\boldsymbol{\lambda}$, and $\mathbf{Y}$. The solution of Eqs. (2.62) to (2.64) thus gives the optimum solution vector $\mathbf{X}^*$, the Lagrange multiplier vector, $\boldsymbol{\lambda}^*$, and the slack variable vector, $\mathbf{Y}^*$.

Equations (2.63) ensure that the constraints $g_j(\mathbf{X}) \leq 0, j = 1,2,\ldots,m$, are satisfied, while Eqs. (2.64) imply that either $\lambda_j = 0$ or $y_j = 0$. If $\lambda_j = 0$, it means that the jth constraint is inactive† and hence can be ignored. On the other hand, if $y_j = 0$, it means that the constraint is active ($g_j = 0$) at the optimum point. Consider the division of the constraints into two subsets, J_1 and J_2, where $J_1 + J_2$ represent the total set of constraints. Let the set J_1 indicate the indices of those constraints that are active at the optimum point and J_2 include the indices of all the inactive constraints.

Thus for $j \in J_1$,‡ $y_j = 0$ (constraints are active), for $j \in J_2$, $\lambda_j = 0$ (constraints are inactive), and Eqs. (2.62) can be simplified as

$$\frac{\partial f}{\partial x_i} + \sum_{j\in J_1} \lambda_j \frac{\partial g_j}{\partial x_i} = 0, \qquad i = 1,2,\ldots,n \tag{2.65}$$

Similarly, Eqs. (2.63) can be written as

$$g_j(\mathbf{X}) = 0, \qquad j \in J_1 \tag{2.66}$$

$$g_j(\mathbf{X}) + y_j^2 = 0, \qquad j \in J_2 \tag{2.67}$$

Equations (2.65) to (2.67) represent $n + p + (m - p) = n + m$ equations in the $n + m$ unknowns x_i $(i = 1,2,\ldots,n)$, λ_j $(j \in J_1)$, and y_j $(j \in J_2)$, where p denotes the number of active constraints.

Assuming that the first p constraints are active, Eqs. (2.65) can be expressed as

$$-\frac{\partial f}{\partial x_i} = \lambda_1 \frac{\partial g_1}{\partial x_i} + \lambda_2 \frac{\partial g_2}{\partial x_i} + \cdots + \lambda_p \frac{\partial g_p}{\partial x_i}, \qquad i = 1,2,\ldots,n \tag{2.68}$$

These equations can be written collectively as

$$-\nabla f = \lambda_1 \nabla g_1 + \lambda_2 \nabla g_2 + \cdots + \lambda_p \nabla g_p \tag{2.69}$$

†Those constraints that are satisfied with an equality sign, $g_j = 0$, at the optimum point are called the *active constraints*, while those that are satisfied with a strict inequality sign, $g_j < 0$, are termed *inactive constraints*.

‡The symbol $\in$ is used to denote the meaning "belongs to" or "element of."

where ∇f and ∇g_j are the gradients of the objective function and the jth constraint, respectively:

$$\nabla f = \begin{Bmatrix} \partial f/\partial x_1 \\ \partial f/\partial x_2 \\ \vdots \\ \partial f/\partial x_n \end{Bmatrix} \quad \text{and} \quad \nabla g_j = \begin{Bmatrix} \partial g_j/\partial x_1 \\ \partial g_j/\partial x_2 \\ \vdots \\ \partial g_j/\partial x_n \end{Bmatrix}$$

Equation (2.69) indicates that the negative of the gradient of the objective function can be expressed as a linear combination of the gradients of the active constraints at the optimum point.

Further, we can show that in the case of a minimization problem, the λ_j values ($j \in J_1$) have to be positive. For simplicity of illustration, suppose that only two constraints are active ($p = 2$) at the optimum point. Then Eq. (2.69) reduces to

$$-\nabla f = \lambda_1 \nabla g_1 + \lambda_2 \nabla g_2 \tag{2.70}$$

Let **S** be a feasible direction[†] at the optimum point. By premultiplying both sides of Eq. (2.70) by $\mathbf{S}^T$, we obtain

$$-\mathbf{S}^T \nabla f = \lambda_1 \mathbf{S}^T \nabla g_1 + \lambda_2 \mathbf{S}^T \nabla g_2 \tag{2.71}$$

where the superscript T denotes the transpose. Since **S** is a feasible direction, it should satisfy the relations

$$\begin{aligned} \mathbf{S}^T \nabla g_1 &< 0 \\ \mathbf{S}^T \nabla g_2 &< 0 \end{aligned} \tag{2.72}$$

[†]A vector **S** is called a *feasible direction* from a point **X** if at least a small step can be taken along **S** that does not immediately leave the feasible region. Thus for problems with sufficiently smooth constraint surfaces, vector **S** satisfying the relation

$$\mathbf{S}^T \nabla g_j < 0$$

can be called a feasible direction. On the other hand, if the constraint is either linear or concave, as shown in Fig. 2.8*b* and *c*, any vector satisfying the relation

$$\mathbf{S}^T \nabla g_j \le 0$$

can be called a feasible direction. The geometric interpretation of a feasible direction is that the vector **S** makes an obtuse angle with all the constraint normals, except that for the linear or outward-curving (concave) constraints, the angle may go to as low as 90°.

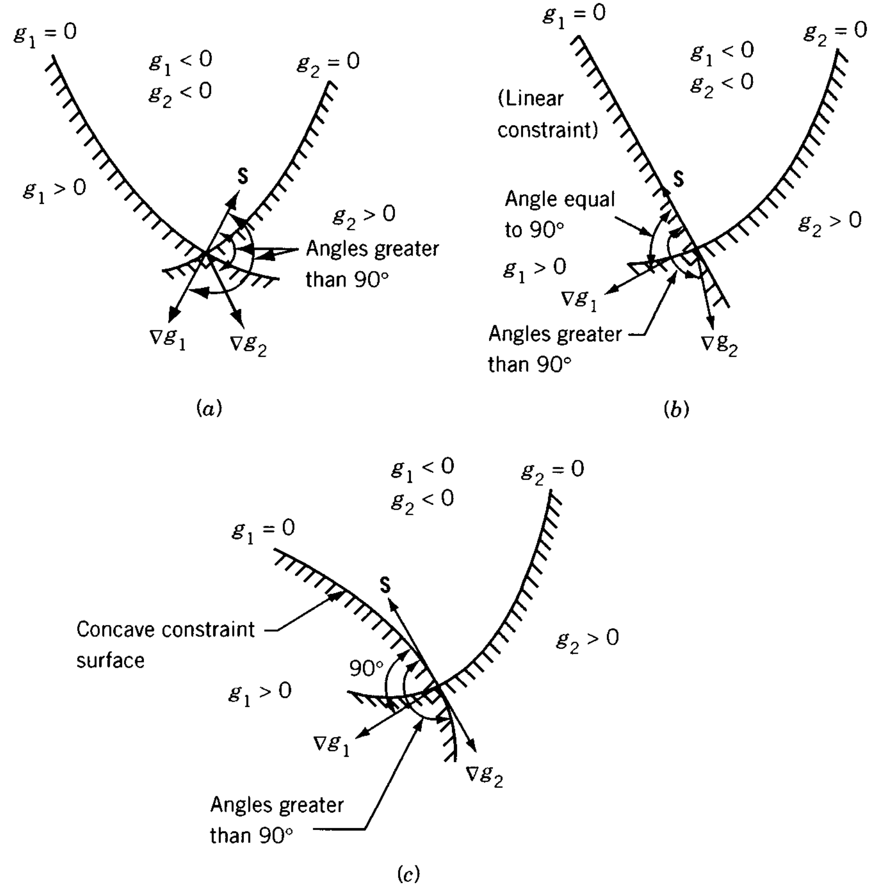

Figure 2.8 Feasible direction **S**.

Thus if $\lambda_1 > 0$ and $\lambda_2 > 0$, the quantity $\mathbf{S}^T\nabla f$ can be seen always to be positive. As ∇f indicates the gradient direction, along which the value of the function increases at the maximum rate,† $\mathbf{S}^T\nabla f$ represents the component of the increment of f along the direction **S**. If $\mathbf{S}^T\nabla f > 0$, the function value increases as we move along the direction **S**. Hence if λ_1 and λ_2 are positive, we will not be able to find any direction in the feasible domain along which the function value can be decreased further. Since the point at which Eq. (2.72) is valid is assumed to be optimum, λ_1 and λ_2 have to be positive. This reasoning can be extended to cases where there are more than two constraints active. By proceeding in a similar manner, one can show that the λ_j values have to be negative for a maximization problem.

†See Section 6.10.2 for a proof of this statement.

2.5.1 Kuhn–Tucker Conditions

As shown above, the conditions to be satisfied at a constrained minimum point, $\mathbf{X}^*$, of the problem stated in Eq. (2.58) can be expressed as

$$\frac{\partial f}{\partial x_i} + \sum_{j \in J_1} \lambda_j \frac{\partial g_j}{\partial x_i} = 0, \qquad i = 1,2,\ldots,n \tag{2.73}$$

$$\lambda_j > 0, \qquad j \in J_1 \tag{2.74}$$

These are called *Kuhn–Tucker conditions* after the mathematicians who derived them as the necessary conditions to be satisfied at a relative minimum of $f(\mathbf{X})$ [2.8]. These conditions are, in general, not sufficient to ensure a relative minimum. However, there is a class of problems, called *convex programming problems*,† for which the Kuhn–Tucker conditions are necessary and sufficient for a global minimum.

If the set of active constraints is not known, the Kuhn–Tucker conditions can be stated as follows:

$$\begin{aligned} \frac{\partial f}{\partial x_i} + \sum_{j=1}^{m} \lambda_j \frac{\partial g_j}{\partial x_i} &= 0, && i = 1,2,\ldots,n \\ \lambda_j g_j &= 0,^{\ddagger} && j = 1,2,\ldots,m \\ g_j &\le 0, && j = 1,2,\ldots,m \\ \lambda_j &\ge 0, && j = 1,2,\ldots,m \end{aligned} \tag{2.75}$$

Note that if the problem is one of maximization or if the constraints are of the type $g_j \ge 0$, the λ_j have to be nonpositive in Eqs. (2.75). On the other hand, if the problem is one of maximization with constraints in the form $g_j \ge 0$, the λ_j have to be nonnegative in Eqs. (2.75).

2.5.2 Constraint Qualification

When the optimization problem is stated as:

$$\text{Minimize } f(\mathbf{X})$$

subject to

$$\begin{aligned} g_j(\mathbf{X}) &\le 0, && j = 1,2,\ldots,m \\ h_k(\mathbf{X}) &= 0, && k = 1,2,\ldots,p \end{aligned} \tag{2.76}$$

†See Sections 2.6 and 7.14 for a detailed discussion of convex programming problems.
‡This condition is the same as Eq. (2.64).

the Kuhn–Tucker conditions become

$$\begin{aligned}
&\nabla f + \sum_{j=1}^{m} \lambda_j \nabla g_j - \sum_{k=1}^{p} \beta_k \nabla h_k = \mathbf{0} \\
&\lambda_j g_j = 0, \quad j = 1,2,\ldots,m \\
&g_j \le 0, \quad j = 1,2,\ldots,m \\
&h_k = 0, \quad k = 1,2,\ldots,p \\
&\lambda_j \ge 0, \quad j = 1,2,\ldots,m
\end{aligned} \tag{2.77}$$

where λ_j and β_k denote the Lagrange multipliers associated with the constraints $g_j \le 0$ and $h_k = 0$, respectively. Although we found qualitatively that the Kuhn–Tucker conditions represent the necessary conditions of optimality, the following theorem gives the precise conditions of optimality.

Theorem 2.7 Let $\mathbf{X}^*$ be a feasible solution to the problem of Eqs. (2.76). If $\nabla g_j(\mathbf{X}^*)$, $j \in J_1$ and $\nabla h_k(\mathbf{X}^*)$, $k = 1,2,\ldots,p$, are linearly independent, there exist $\boldsymbol{\lambda}^*$ and $\boldsymbol{\beta}^*$ such that $(\mathbf{X}^*, \boldsymbol{\lambda}^*, \boldsymbol{\beta}^*)$ satisfy Eqs. (2.77).

Proof: See Ref. [2.11].

The requirement that $\nabla g_j(\mathbf{X}^*)$, $j \in J_1$ and $\nabla h_k(\mathbf{X}^*)$, $k = 1,2,\ldots,p$, be linearly independent is called the *constraint qualification*. If the constraint qualification is violated at the optimum point, Eqs. (2.77) may or may not have a solution. It is difficult to verify the constraint qualification without knowing $\mathbf{X}^*$ beforehand. However, the constraint qualification is always satisfied for problems having any of the following characteristics:

1. All the inequality and equality constraint functions are linear.
2. All the inequality constraint functions are convex, all the equality constraint functions are linear, and at least one feasible vector $\tilde{\mathbf{X}}$ exists that lies strictly inside the feasible region, so that

$$g_j(\tilde{\mathbf{X}}) < 0, \quad j = 1,2,\ldots,m \quad \text{and} \quad h_k(\tilde{\mathbf{X}}) = 0, \quad k = 1,2,\ldots,p$$

Example 2.12 Consider the problem:

$$\text{Minimize } f(x_1, x_2) = (x_1 - 1)^2 + x_2^2 \tag{E$_1$}$$

subject to

$$g_1(x_1, x_2) = x_1^3 - 2x_2 \le 0 \tag{E$_2$}$$

$$g_2(x_1, x_2) = x_1^3 + 2x_2 \le 0 \tag{E$_3$}$$

Determine whether the constraint qualification and the Kuhn–Tucker conditions are satisfied at the optimum point.

SOLUTION The feasible region and the contours of the objective function are shown in Fig. 2.9. It can be seen that the optimum solution is (0, 0). Since g_1 and g_2 are both active at the optimum point (0, 0), their gradients can be computed as

$$\nabla g_1(\mathbf{X}^*) = \begin{Bmatrix} 3x_1^2 \\ -2 \end{Bmatrix}_{(0,0)} = \begin{Bmatrix} 0 \\ -2 \end{Bmatrix} \quad \text{and} \quad \nabla g_2(\mathbf{X}^*) = \begin{Bmatrix} 3x_1^2 \\ 2 \end{Bmatrix}_{(0,0)} = \begin{Bmatrix} 0 \\ 2 \end{Bmatrix}$$

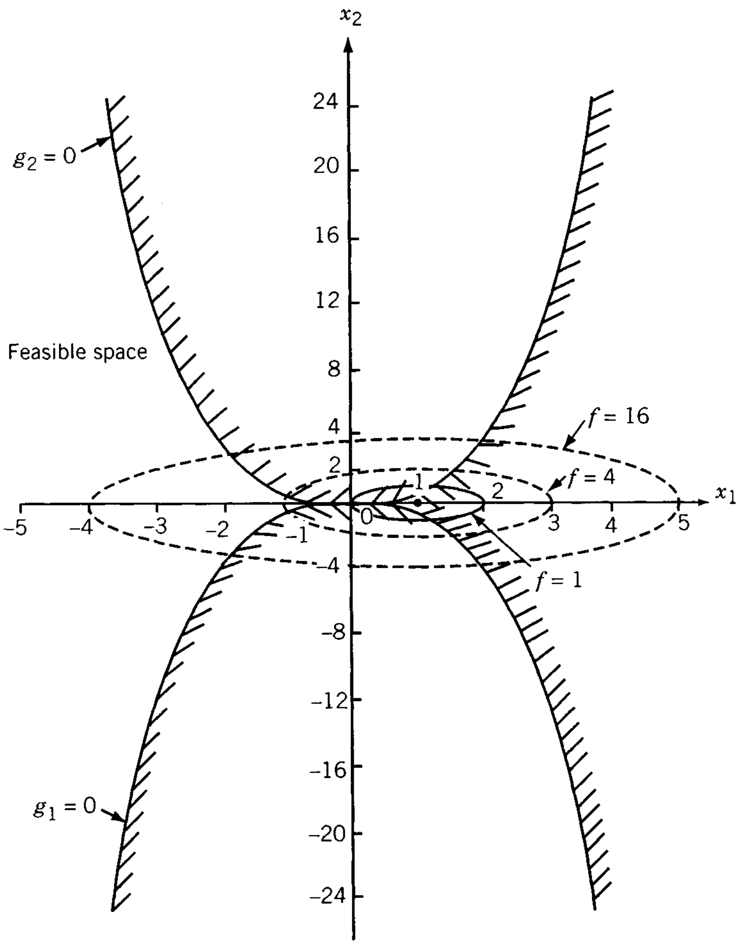

Figure 2.9 Feasible region and contours of the objective function.

It is clear that $\nabla g_1(\mathbf{X}^*)$ and $\nabla g_2(\mathbf{X}^*)$ are not linearly independent. Hence the constraint qualification is not satisfied at the optimum point.

Noting that

$$\nabla f(\mathbf{X}^*) = \begin{Bmatrix} 2(x_1 - 1) \\ 2x_2 \end{Bmatrix}_{(0,0)} = \begin{Bmatrix} -2 \\ 0 \end{Bmatrix}$$

the Kuhn–Tucker conditions can be written, using Eqs. (2.73) and (2.74), as

$$-2 + \lambda_1(0) + \lambda_2(0) = 0 \tag{E$_4$}$$

$$0 + \lambda_1(-2) + \lambda_2(2) = 0 \tag{E$_5$}$$

$$\lambda_1 > 0 \tag{E$_6$}$$

$$\lambda_2 > 0 \tag{E$_7$}$$

Since Eq. (E_4) is not satisfied and Eq. (E_5) can be satisfied for negative values of $\lambda_1 = \lambda_2$ also, the Kuhn–Tucker conditions are not satisfied at the optimum point.

Example 2.13 A manufacturing firm producing small refrigerators has entered into a contract to supply 50 refrigerators at the end of the first month, 50 at the end of the second month, and 50 at the end of the third. The cost of producing x refrigerators in any month is given by $\$(x^2 + 1000)$. The firm can produce more refrigerators in any month and carry them to a subsequent month. However, it costs \$20 per unit for any refrigerator carried over from one month to the next. Assuming that there is no initial inventory, determine the number of refrigerators to be produced in each month to minimize the total cost.

SOLUTION Let x_1, x_2, and x_3 represent the number of refrigerators produced in the first, second, and third month, respectively. The total cost to be minimized is given by

$$\text{total cost} = \text{production cost} + \text{holding cost}$$

or

$$\begin{aligned} f(x_1, x_2, x_3) &= (x_1^2 + 1000) + (x_2^2 + 1000) + (x_3^2 + 1000) + 20(x_1 - 50) \\ &\quad + 20(x_1 + x_2 - 100) \\ &= x_1^2 + x_2^2 + x_3^2 + 40x_1 + 20x_2 \end{aligned}$$

The constraints can be stated as

$$g_1(x_1, x_2, x_3) = x_1 - 50 \geq 0$$

$$g_2(x_1,x_2,x_3) = x_1 + x_2 - 100 \geq 0$$

$$g_3(x_1,x_2,x_3) = x_1 + x_2 + x_3 - 150 \geq 0$$

The Kuhn–Tucker conditions are given by

$$\frac{\partial f}{\partial x_i} + \lambda_1 \frac{\partial g_1}{\partial x_i} + \lambda_2 \frac{\partial g_2}{\partial x_i} + \lambda_3 \frac{\partial g_3}{\partial x_i} = 0, \qquad i = 1,2,3$$

that is,

$$2x_1 + 40 + \lambda_1 + \lambda_2 + \lambda_3 = 0 \tag{E$_1$}$$

$$2x_2 + 20 + \lambda_2 + \lambda_3 = 0 \tag{E$_2$}$$

$$2x_3 + \lambda_3 = 0 \tag{E$_3$}$$

$$\lambda_j g_j = 0, \qquad j = 1, 2, 3$$

that is,

$$\lambda_1(x_1 - 50) = 0 \tag{E$_4$}$$

$$\lambda_2(x_1 + x_2 - 100) = 0 \tag{E$_5$}$$

$$\lambda_3(x_1 + x_2 + x_3 - 150) = 0 \tag{E$_6$}$$

$$g_j \geq 0, \qquad j = 1, 2, 3$$

that is,

$$x_1 - 50 \geq 0 \tag{E$_7$}$$

$$x_1 + x_2 - 100 \geq 0 \tag{E$_8$}$$

$$x_1 + x_2 + x_3 - 150 \geq 0 \tag{E$_9$}$$

$$\lambda_j \leq 0, \qquad j = 1, 2, 3$$

that is,

$$\lambda_1 \leq 0 \tag{E$_{10}$}$$

$$\lambda_2 \leq 0 \tag{E$_{11}$}$$

$$\lambda_3 \leq 0 \tag{E$_{12}$}$$

The solution of Eqs. (E_1) to (E_{12}) can be found in several ways. We proceed to solve these equations by first noting that either $\lambda_1 = 0$ or $x_1 = 50$ according to Eq. (E_4). Using this information, we investigate the following cases to identify the optimum solution of the problem.

Case 1: $\lambda_1 = 0$. Equations (E_1) to (E_3) give

$$x_3 = -\frac{\lambda_3}{2}$$

$$x_2 = -10 - \frac{\lambda_2}{2} - \frac{\lambda_3}{2} \qquad (E_{13})$$

$$x_1 = -20 - \frac{\lambda_2}{2} - \frac{\lambda_3}{2}$$

Substituting Eqs. (E_{13}) in Eqs. (E_5) and (E_6), we obtain

$$\lambda_2(-130 - \lambda_2 - \lambda_3) = 0$$

$$\lambda_3(-180 - \lambda_2 - \tfrac{3}{2}\lambda_3) = 0 \qquad (E_{14})$$

The four possible solutions of Eqs. (E_{14}) are:

1. $\lambda_2 = 0,\ -180 - \lambda_2 - \frac{3}{2}\lambda_3 = 0$. These equations, along with Eqs. (E_{13}), yield the solution

   $$\lambda_2 = 0, \quad \lambda_3 = -120, \quad x_1 = 40, \quad x_2 = 50, \quad x_3 = 60$$

   This solution satisfies Eqs. (E_{10}) to (E_{12}) but violates Eqs. (E_7) and (E_8) and hence cannot be optimum.
2. $\lambda_3 = 0,\ -130 - \lambda_2 - \lambda_3 = 0$. The solution of these equations leads to

   $$\lambda_2 = -130, \quad \lambda_3 = 0, \quad x_1 = 45, \quad x_2 = 55, \quad x_3 = 0$$

   This solution can be seen to satisfy Eqs. (E_{10}) to (E_{12}) but violate Eqs. (E_7) and (E_9).
3. $\lambda_2 = 0, \lambda_3 = 0$. Equations ($E_{13}$) give

   $$x_1 = -20, \quad x_2 = -10, \quad x_3 = 0$$

   This solution satisfies Eqs. (E_{10}) to (E_{12}) but violates the constraints, Eqs. (E_7) to (E_9).
4. $-130 - \lambda_2 - \lambda_3 = 0, \quad -180 - \lambda_2 - \frac{3}{2}\lambda_3 = 0$. The solution of these equations and Eqs. (E_{13}) yields

   $$\lambda_2 = -30, \quad \lambda_3 = -100, \quad x_1 = 45, \quad x_2 = 55, \quad x_3 = 50$$

This solution satisfies Eqs. (E_{10}) to (E_{12}) but violates the constraint, Eq. (E_7).

Case 2: $x_1 = 50$. In this case, Eqs. (E_1) to (E_3) give

$$\begin{aligned} \lambda_3 &= -2x_3 \\ \lambda_2 &= -20 - 2x_2 - \lambda_3 = -20 - 2x_2 + 2x_3 \\ \lambda_1 &= -40 - 2x_1 - \lambda_2 - \lambda_3 = -120 + 2x_2 \end{aligned} \tag{E_{15}}$$

Substitution of Eqs. (E_{15}) in Eqs. (E_5) and (E_6) leads to

$$\begin{aligned} (-20 - 2x_2 + 2x_3)(x_1 + x_2 - 100) &= 0 \\ (-2x_3)(x_1 + x_2 + x_3 - 150) &= 0 \end{aligned} \tag{E_{16}}$$

Once again, it can be seen that there are four possible solutions to Eqs. (E_{16}), as indicated below.

1. $-20 - 2x_2 + 2x_3 = 0$, $x_1 + x_2 + x_3 - 150 = 0$: The solution of these equations yields

$$x_1 = 50, \qquad x_2 = 45, \qquad x_3 = 55$$

   This solution can be seen to violate Eq. (E_8).
2. $-20 - 2x_2 + 2x_3 = 0$, $-2x_3 = 0$: These equations lead to the solution

$$x_1 = 50, \qquad x_2 = -10, \qquad x_3 = 0$$

   This solution can be seen to violate Eqs. (E_8) and (E_9).
3. $x_1 + x_2 - 100 = 0$, $-2x_3 = 0$: These equations give

$$x_1 = 50, \qquad x_2 = 50, \qquad x_3 = 0$$

   This solution violates the constraint Eq. (E_9).
4. $x_1 + x_2 - 100 = 0$, $x_1 + x_2 + x_3 - 150 = 0$: The solution of these equations yields

$$x_1 = 50, \qquad x_2 = 50, \qquad x_3 = 50$$

   This solution can be seen to satisfy all the constraint Eqs. (E_7) to (E_9). The values of λ_1, λ_2, and λ_3 corresponding to this solution can be obtained from Eqs. (E_{15}) as

$$\lambda_1 = -20, \qquad \lambda_2 = -20, \qquad \lambda_3 = -100$$

Since these values of λ_i satisfy the requirements [Eqs. (E_{10}) to (E_{12})], this solution can be identified as the optimum solution. Thus

$$x_1^* = 50, \quad x_2^* = 50, \quad x_3^* = 50$$

2.6 CONVEX PROGRAMMING PROBLEM

The optimization problem stated in Eq. (2.58) is called a *convex programming problem* if the objective function $f(\mathbf{X})$, and the constraint functions, $g_j(\mathbf{X})$, are convex. The definition and properties of a convex function are given in Appendix A. Suppose that $f(\mathbf{X})$ and $g_j(\mathbf{X})$, $j = 1,2,\ldots,m$, are convex functions. The Lagrange function of Eq. (2.61) can be written as

$$L(\mathbf{X},\mathbf{Y},\lambda) = f(\mathbf{X}) + \sum_{j=1}^{m} \lambda_j [g_j(\mathbf{X}) + y_j^2] \tag{2.78}$$

If $\lambda_j \geq 0$, then $\lambda_j g_j(\mathbf{X})$ is convex, and since $\lambda_j y_j = 0$ from Eq. (2.64), $L(\mathbf{X}, \mathbf{Y},\lambda)$ will be a convex function. As shown earlier, a necessary condition for $f(\mathbf{X})$ to be a relative minimum at $\mathbf{X}^*$ is that $L(\mathbf{X},\mathbf{Y}, \lambda)$ have a stationary point at $\mathbf{X}^*$. However, if $\mathbf{L}(\mathbf{X},\mathbf{Y},\lambda)$ is a convex function, its derivative vanishes only at one point, which must be an absolute minimum of the function $f(\mathbf{X})$. Thus the Kuhn–Tucker conditions are both necessary and sufficient for an absolute minimum of $f(\mathbf{X})$ at $\mathbf{X}^*$.

Notes:

1. If the given optimization problem is known to be a convex programming problem, there will be no relative minima or saddle points, and hence the extreme point found by applying the Kuhn–Tucker conditions is guaranteed to be an absolute minimum of $f(\mathbf{X})$. However, it is often very difficult to ascertain whether the objective and constraint functions involved in a practical engineering problem are convex.
2. The derivation of the Kuhn–Tucker conditions was based on the development given for equality constraints in Section 2.4. One of the requirements for these conditions was that at least one of the Jacobians composed of the m constraints and m of the $n + m$ variables ($x_1,x_2,\ldots,x_n$; $y_1,y_2,\ldots,y_m$) be nonzero. This requirement is implied in the derivation of the Kuhn–Tucker conditions.

REFERENCES AND BIBLIOGRAPHY

2.1 H. Hancock, *Theory of Maxima and Minima*, Dover, New York, 1960.
2.2 M. E. Levenson, *Maxima and Minima*, Macmillan, New York, 1967.

2.3 G. B. Thomas, Jr., *Calculus and Analytic Geometry*, Addison-Wesley, Reading. Mass., 1967.

2.4 A. E. Richmond, *Calculus for Electronics*, McGraw-Hill, New York, 1972.

2.5 B. Kolman and W. F. Trench, *Elementary Multivariable Calculus*, Academic Press, New York, 1971.

2.6 G. S. G. Beveridge and R. S. Schechter, *Optimization: Theory and Practice*, McGraw-Hill, New York, 1970.

2.7 R. Gue and M. E. Thomas, *Mathematical Methods of Operations Research*, Macmillan, New York, 1968.

2.8 H. W. Kuhn and A. Tucker, Nonlinear Programming, in *Proceedings of the 2nd Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press, Berkeley, 1951.

2.9 F. Ayres, Jr., *Theory and Problems of Matrices*, Schaum's Outline Series, Schaum, New York, 1962.

2.10 M. J. Panik, *Classical Optimization: Foundations and Extensions*, North-Holland, Amsterdam, 1976.

2.11 M. S. Bazaraa and C. M. Shetty, *Nonlinear Programming: Theory and Algorithms*, Wiley, New York, 1979.

2.12 D. M. Simmons, *Nonlinear Programming for Operations Research*, Prentice Hall, Englewood Cliffs, N.J., 1975.

2.13 J. R. Howell and R. O. Buckius, *Fundamentals of Engineering Thermodynamics*, 2nd ed., McGraw-Hill, New York, 1992.

REVIEW QUESTIONS

2.1 State the necessary and sufficient conditions for the minimum of a function $f(x)$.

2.2 Under what circumstances can the condition $df(x)/dx = 0$ not be used to find the minimum of the function $f(x)$?

2.3 Define the rth differential, $d^r f(\mathbf{X})$, of a multivariable function $f(\mathbf{X})$.

2.4 Write the Taylor's series expansion of a function $f(\mathbf{X})$.

2.5 State the necessary and sufficient conditions for the maximum of a multivariable function $f(\mathbf{X})$.

2.6 What is a quadratic form?

2.7 How do you test the positive, negative, or indefiniteness of a square matrix $[A]$?

2.8 Define a saddle point and indicate its significance.

2.9 State the various methods available for solving a multivariable optimization problem with equality constraints.

2.10 State the principle behind the method of constrained variation.

2.11 What is the Lagrange multiplier method?

2.12 What is the significance of Lagrange multipliers?

2.13 Convert an inequality constrained problem into an equivalent unconstrained problem.

2.14 State the Kuhn–Tucker conditions.

2.15 What is an active constraint?

2.16 Define a usable feasible direction.

2.17 What is a convex programming problem? What is its significance?

2.18 Answer whether each of the following quadratic forms is positive definite, negative definite, or neither.
(a) $f = x_1^2 - x_2^2$
(b) $f = 4x_1x_2$
(c) $f = x_1^2 + 2x_2^2$
(d) $f = -x_1^2 + 4x_1x_2 + 4x_2^2$
(e) $f = -x_1^2 + 4x_1x_2 - 9x_2^2 + 2x_1x_3 + 8x_2x_3 - 4x_3^2$

2.19 State whether each of the following functions is convex, concave, or neither.
(a) $f = -2x^2 + 8x + 4$
(b) $f = x^2 + 10x + 1$
(c) $f = x_1^2 - x_2^2$
(d) $f = -x_1^2 + 4x_1x_2$
(e) $f = e^{-x}, x > 0$
(f) $f = \sqrt{x}, x > 0$
(g) $f = x_1x_2$
(h) $f = (x_1 - 1)^2 + 10(x_2 - 2)^2$

2.20 Match the following equations and their characteristics.

(a) $f = 4x_1 - 3x_2 + 2$	Relative maximum at (1, 2)
(b) $f = (2x_1 - 2)^2 + (x_1 - 2)^2$	Saddle point at origin
(c) $f = -(x_1 - 1)^2 - (x_2 - 2)^2$	No minimum
(d) $f = x_1x_2$	Inflection point at origin
(e) $f = x^3$	Relative minimum at (1, 2)

PROBLEMS

2.1 A dc generator has an internal resistance R ohms and develops an open-circuit voltage of V volts (Fig. 2.10). Find the value of the load resis-

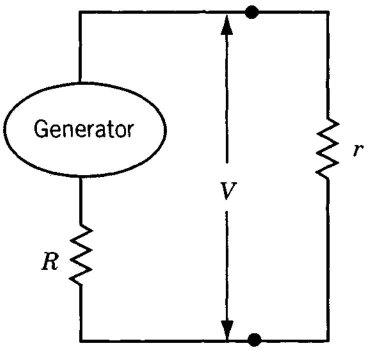

Figure 2.10 Electric generator with load.

tance r for which the power delivered by the generator will be a maximum.

2.2 Find the maxima and minima, if any, of the function

$$f(x) = \frac{x^4}{(x-1)(x-3)^3}$$

2.3 Find the maxima and minima, if any, of the function

$$f(x) = 4x^3 - 18x^2 + 27x - 7$$

2.4 The efficiency of a screw jack is given by

$$\eta = \frac{\tan \alpha}{\tan (\alpha + \phi)}$$

where α is the lead angle and ϕ is a constant. Prove that the efficiency of the screw jack will be maximum when $\alpha = 45° - \phi/2$ with $\eta_{\max} = (1 - \sin \phi)/(1 + \sin \phi)$.

2.5 Find the minimum of the function

$$f(x) = 10x^6 - 48x^5 + 15x^4 + 200x^3 - 120x^2 - 480x + 100$$

2.6 Find the angular orientation of a cannon to maximize the range of the projectile.

2.7 In a submarine telegraph cable the speed of signalling varies as $x^2 \log(1/x)$, where x is the ratio of the radius of the core to that of the covering. Show that the greatest speed is attained when this ratio is $1 : \sqrt{e}$.

2.8 The horsepower generated by a Pelton wheel is proportional to $u(V - u)$, where u is the velocity of the wheel, which is variable, and V is the velocity of the jet, which is fixed. Show that the efficiency of the Pelton wheel will be maximum when $u = V/2$.

2.9 A pipe of length l and diameter D has at one end a nozzle of diameter d through which water is discharged from a reservoir. The level of water in the reservoir is maintained at a constant value h above the center of nozzle. Find the diameter of the nozzle so that the kinetic energy of the jet is a maximum. The kinetic energy of the jet can be expressed as

$$\frac{1}{4} \pi \rho d^2 \left(\frac{2gD^5 h}{D^5 + 4fld^4} \right)^{3/2}$$

where ρ is the density of water, f the friction coefficient and g the gravitational constant.

2.10 An electric light is placed directly over the center of a circular plot of lawn 100 m in diameter. Assuming that the intensity of light varies directly as the sine of the angle at which it strikes an illuminated surface, and inversely as the square of its distance from the surface, how high should the light be hung in order that the intensity may be as great as possible at the circumference of the plot?

2.11 If a crank is at an angle θ from dead center with $\theta = \omega t$, where ω is the angular velocity and t is time, the distance of the piston from the end of its stroke (x) is given by

$$x = r(1 - \cos\theta) + \frac{r^2}{4l}(1 - \cos 2\theta)$$

where r is the length of the crank and l is the length of the connecting rod. For $r = 1$ and $l = 5$, find **(a)** the angular position of the crank at which the piston moves with maximum velocity, and **(b)** the distance of the piston from the end of its stroke at that instant.

Determine whether each of the following matrices is positive definite, negative definite, or indefinite by finding its eigenvalues.

2.12 $[A] = \begin{bmatrix} 3 & 1 & -1 \\ 1 & 3 & -1 \\ -1 & -1 & 5 \end{bmatrix}$

2.13 $[B] = \begin{bmatrix} 4 & 2 & -4 \\ 2 & 4 & -2 \\ -4 & -2 & 4 \end{bmatrix}$

2.14 $[C] = \begin{bmatrix} -1 & -1 & -1 \\ -1 & -2 & -2 \\ -1 & -2 & -3 \end{bmatrix}$

Determine whether each of the following matrices is positive definite, negative definite, or indefinite by evaluating the signs of its submatrices.

2.15 $[A] = \begin{bmatrix} 3 & 1 & -1 \\ 1 & 3 & -1 \\ -1 & -1 & 5 \end{bmatrix}$

2.16 $[B] = \begin{bmatrix} 4 & 2 & -4 \\ 2 & 4 & -2 \\ -4 & -2 & 4 \end{bmatrix}$

2.17 $[C] = \begin{bmatrix} -1 & -1 & -1 \\ -1 & -2 & -2 \\ -1 & -2 & -3 \end{bmatrix}$

2.18 Express the function

$$f(x_1,x_2,x_3) = -x_1^2 - x_2^2 + 2x_1x_2 - x_3^2 + 6x_1x_3 + 4x_1 - 5x_3 + 2$$

in matrix form as

$$f(\mathbf{X}) = \tfrac{1}{2}\,\mathbf{X}^T\,[A]\,\mathbf{X} + \mathbf{B}^T\,\mathbf{X} + C$$

and determine whether the matrix $[A]$ is positive definite, negative definite, or indefinite.

2.19 Determine whether the following matrix is positive or negative definite.

$$[A] = \begin{bmatrix} 4 & -3 & 0 \\ -3 & 0 & 4 \\ 0 & 4 & 2 \end{bmatrix}$$

2.20 Determine whether the following matrix is positive definite.

$$[A] = \begin{bmatrix} -14 & 3 & 0 \\ 3 & -1 & 4 \\ 0 & 4 & 2 \end{bmatrix}$$

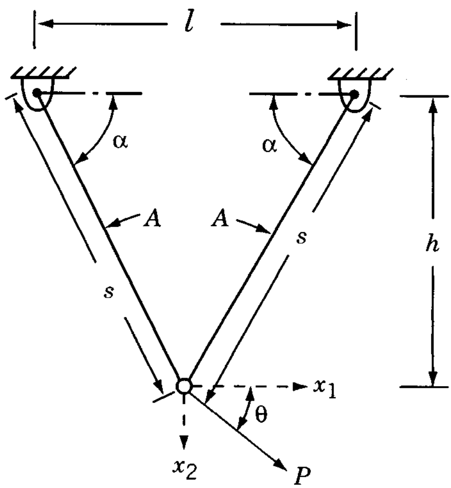

Figure 2.11 Two-bar truss.

2.21 The potential energy of the two-bar truss shown in Fig. 2.11 is given by

$$f(x_1, x_2) = \frac{EA}{s}\left(\frac{1}{2s}\right)^2 x_1^2 + \frac{EA}{s}\left(\frac{h}{s}\right)^2 x_2^2 - Px_1 \cos\theta - Px_2 \sin\theta$$

where E is Young's modulus, A the cross-sectional area of each member, l the span of the truss, s the length of each member, h the height of the truss, P the applied load, θ the angle at which the load is applied, and x_1 and x_2 are, respectively, the horizontal and vertical displacements of the free node. Find the values of x_1 and x_2 that minimize the potential energy when $E = 207 \times 10^9$ Pa, $A = 10^{-5}$ m^2, $l = 1.5$ m, $h = 4.0$ m, $P = 10^4$ N, and $\theta = 30°$.

2.22 The profit per acre of a farm is given by

$$20x_1 + 26x_2 + 4x_1x_2 - 4x_1^2 - 3x_2^2$$

where x_1 and x_2 denote, respectively, the labor cost and the fertilizer cost. Find the values of x_1 and x_2 to maximize the profit.

2.23 The temperatures measured at various points inside a heated wall are as follows:

Distance from the heated surface as a percentage of wall thickness, d	0	25	50	75	100
Temperature, t (°C)	380	200	100	20	0

It is decided to approximate this table by a linear equation (graph) of the form $t = a + bd$, where a and b are constants. Find the values of the constants a and b that minimize the sum of the squares of all differences between the graph values and the tabulated values.

2.24 Find the second-order Taylor's series approximation of the function

$$f(x_1,x_2) = (x_1 - 1)^2 e^{x_2} + x_1$$

at the points **(a)** (0,0) and **(b)** (1,1).

2.25 Find the third-order Taylor's series approximation of the function

$$f(x_1,x_2,x_3) = x_2^2 x_3 + x_1 e^{x_3}$$

at point (1,0,−2).

2.26 The volume of sales (f) of a product is found to be a function of the number of newspaper advertisements (x) and the number of minutes of television time (y) as

$$f = 12xy - x^2 - 3y^2$$

Each newspaper advertisement or each minute on television costs \$1000. How should the firm allocate \$48,000 between the two advertising media for maximizing its sales?

2.27 Find the value of x^* at which the following function attains its maximum:

$$f(x) = \frac{1}{10\sqrt{2\pi}} e^{-(1/2)[(x-100)/10]^2}$$

2.28 It is possible to establish the nature of stationary points of an objective function based on its quadratic approximation. For this, consider the quadratic approximation of a two-variable function as

$$f(\mathbf{X}) \approx a + \mathbf{b}^T\mathbf{X} + \tfrac{1}{2}\mathbf{X}^T[c]\,\mathbf{X}$$

where

$$\mathbf{X} = \begin{Bmatrix} x_1 \\ x_2 \end{Bmatrix}, \quad \mathbf{b} = \begin{Bmatrix} b_1 \\ b_2 \end{Bmatrix}, \quad \text{and} \quad [c] = \begin{bmatrix} c_{11} & c_{12} \\ c_{12} & c_{22} \end{bmatrix}$$

If the eigenvalues of the Hessian matrix, $[c]$, are denoted as β_1 and β_2, identify the nature of the contours of the objective function and the type of stationary point in each of the following situations.

(a) $\beta_1 = \beta_2$; both positive
(b) $\beta_1 > \beta_2$; both positive
(c) $|\beta_1| = |\beta_2|$; β_1 and β_2 have opposite signs
(d) $\beta_1 > 0$, $\beta_2 = 0$

Plot the contours of each of the following functions and identify the nature of its stationary point.

2.29 $f = 2 - x^2 - y^2 + 4xy$

2.30 $f = 2 + x^2 - y^2$

2.31 $f = xy$

2.32 $f = x^3 - 3xy^2$

2.33 Find the admissible and constrained variations at the point $\mathbf{X} = \begin{Bmatrix} 0 \\ 4 \end{Bmatrix}$ for the following problem:

$$\text{Minimize } f = x_1^2 + (x_2 - 1)^2$$

subject to

$$-2x_1^2 + x_2 = 4$$

2.34 Find the diameter of an open cylindrical can that will have the maximum volume for a given surface area, S.

2.35 A rectangular beam is to be cut from a circular log of radius r. Find the cross-sectional dimensions of the beam to **(a)** maximize the cross-sectional area of the beam, and **(b)** maximize the perimeter of the beam section.

2.36 Find the dimensions of a straight beam of circular cross section that can be cut from a conical log of height h and base radius r to maximize the volume of the beam.

2.37 The deflection of a rectangular beam is inversely proportional to the width and the cube of depth. Find the cross-sectional dimensions of a beam, which corresponds to minimum deflection, that can be cut from a cylindrical log of radius r.

2.38 A rectangular box of height a and width b is placed adjacent to a wall (Fig. 2.12). Find the length of the shortest ladder that can be made to lean against the wall.

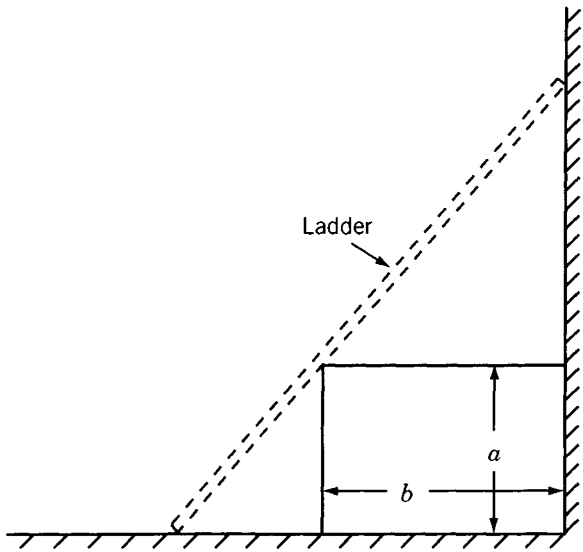

Figure 2.12 Ladder against a wall.

2.39 Show that the right circular cylinder of given surface (including the ends) and maximum volume is such that its height is equal to the diameter of the base.

2.40 Find the dimensions of a closed cylindrical soft drink can that can hold soft drink of volume V for which the surface area (including the top and bottom) is a minimum.

2.41 An open rectangular box is to be manufactured from a given amount of sheet metal (area S). Find the dimensions of the box to maximize the volume.

2.42 Find the dimensions of an open rectangular box of volume V for which the amount of material required for manufacture (surface area) is a minimum.

2.43 A rectangular sheet of metal with sides a and b has four equal square portions (of side d) removed at the corners, and the sides are then turned up so as to form an open rectangular box. Find the depth of the box that maximizes the volume.

2.44 Show that the cone of the greatest volume which can be inscribed in a given sphere has an altitude equal to two-thirds of the diameter of the sphere. Also prove that the curved surface of the cone is a maximum for the same value of the altitude.

2.45 Prove Theorem 2.6.

2.46 A log of length l is in the form of a frustum of a cone whose ends have radii a and b ($a > b$). It is required to cut from it a beam of uniform square section. Prove that the beam of greatest volume that can be cut has a length of $al/[3(a - b)]$.

2.47 It has been decided to leave a margin of 30 mm at the top and 20 mm each at the left side, right side, and the bottom on the printed page of a book. If the area of the page is specified as 5×10^4 mm^2, determine the dimensions of a page that provide the largest printed area.

2.48

$$\text{Minimize } f = 9 - 8x_1 - 6x_2 - 4x_3 + 2x_1^2 + 2x_2^2 + x_3^2 + 2x_1x_2 + 2x_1x_3$$

subject to

$$x_1 + x_2 + 2x_3 = 3$$

by **(a)** direct substitution, **(b)** constrained variation, and **(c)** Lagrange multiplier method.

2.49

$$\text{Minimize } f(\mathbf{X}) = \tfrac{1}{2}(x_1^2 + x_2^2 + x_3^2)$$

subject to

$$g_1(\mathbf{X}) = x_1 - x_2 = 0$$

$$g_2(\mathbf{X}) = x_1 + x_2 + x_3 - 1 = 0$$

by **(a)** direct substitution, **(b)** constrained variation, and **(c)** Lagrange multiplier method.

2.50 Find the values of x, y, and z that maximize the function

$$f(x,y,z) = \frac{6xyz}{x + 2y + 2z}$$

when x, y, and z are restricted by the relation $xyz = 16$.

2.51 A tent on a square base of side $2a$ consists of four vertical sides of height b surmounted by a regular pyramid of height h. If the volume enclosed by the tent is V, show that the area of canvas in the tent can be expressed as

$$\frac{2V}{a} - \frac{8ah}{3} + 4a\sqrt{h^2 + a^2}$$

Also show that the least area of the canvas corresponding to a given volume V, if a and h can both vary, is given by

$$a = \frac{\sqrt{5}\, h}{2} \quad \text{and} \quad h = 2b$$

2.52 A departmental store plans to construct a one-story building with a rectangular planform. The building is required to have a floor area of 22,500 ft^2 and a height of 18 ft. It is proposed to use brick walls on three sides and a glass wall on the fourth side. Find the dimensions of the building to minimize the cost of construction of the walls and the roof assuming that the glass wall costs twice as much as that of the brick wall and the roof costs three times as much as that of the brick wall per unit area.

2.53 Find the dimensions of the rectangular building described in Problem 2.52 to minimize the heat loss assuming that the relative heat losses per unit surface area for the roof, brick wall, glass wall, and floor are in the proportion 4 : 2 : 5 : 1.

2.54 A funnel, in the form of a right circular cone, is to be constructed from a sheet metal. Find the dimensions of the funnel for minimum lateral surface area when the volume of the funnel is specified as 200 in^3.

2.55 Find the effect on f^* when the value of A_0 is changed to **(a)** 25π and **(b)** 22π in Example 2.10 using the property of the Lagrange multiplier.

2.56 **(a)** Find the dimensions of a rectangular box of volume $V = 1000$ in^3 for which the total length of the 12 edges is a minimum using the Lagrange multiplier method.

(b) Find the change in the dimensions of the box when the volume is changed to 1200 in^3 by using the value of λ^* found in part (a).

(c) Compare the solution found in part (b) with the exact solution.

2.57 Find the effect on f^* of changing the constraint to **(a)** $x + x_2 + 2x_3 = 4$ and **(b)** $x + x_2 + 2x_3 = 2$ in Problem 2.48. Use the physical meaning of Lagrange multiplier in finding the solution.

2.58 A real estate company wants to construct a multistory apartment building on a 500 ft $\times$ 500 ft lot. It has been decided to have a total floor space of 8×10^5 ft^2. The height of each story is required to be 12 ft, the maximum height of the building is to be restricted to 75 ft, and the parking area is required to be at least 10% of the total floor area according to the city zoning rules. If the cost of the building is estimated at \$$(500{,}000h + 2000F + 500P)$, where h is the height in feet, F is the floor area in square feet, and P is the parking area in square feet. Find the minimum cost design of the building.

2.59 Identify the optimum point among the given design vectors, $\mathbf{X}_1$, $\mathbf{X}_2$, and $\mathbf{X}_3$, by applying the Kuhn–Tlucker conditions to the following problem:

$$\text{Minimize } f(\mathbf{X}) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$$

subject to

$$x_2^2 - x_1 \geq 0$$

$$x_1^2 - x_2 \geq 0$$

$$-\tfrac{1}{2} \leq x_1 \leq \tfrac{1}{2}, \qquad x_2 \leq 1$$

$$\mathbf{X}_1 = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix}, \quad \mathbf{X}_2 = \begin{Bmatrix} 0 \\ -1 \end{Bmatrix}, \quad \mathbf{X}_3 = \begin{Bmatrix} -\frac{1}{2} \\ \frac{1}{4} \end{Bmatrix}$$

2.60 Consider the following optimization problem:

$$\text{Maximize } f = -x_1^2 - x_2^2 + x_1x_2 + 7x_1 + 4x_2$$

subject to

$$2x_1 + 3x_2 \leq 24$$

$$-5x_1 + 12x_2 \leq 24$$

$$x_1 \geq 0, \qquad x_2 \geq 0, \qquad x_2 \leq 4$$

Find a usable feasible direction at each of the following design vectors:

$$\mathbf{X}_1 = \begin{Bmatrix} 1 \\ 1 \end{Bmatrix}, \quad \mathbf{X}_2 = \begin{Bmatrix} 6 \\ 4 \end{Bmatrix}$$

2.61 Consider the following problem:

$$\text{Minimize } f = (x_1 - 2)^2 + (x_2 - 1)^2$$

subject to

$$2 \geq x_1 + x_2$$

$$x_2 \geq x_1^2$$

Using Kuhn–Tucker conditions, find which of the following vectors are local minima:

$$\mathbf{X}_1 = \begin{Bmatrix} 1.5 \\ 0.5 \end{Bmatrix}, \quad \mathbf{X}_2 = \begin{Bmatrix} 1 \\ 1 \end{Bmatrix}, \quad \mathbf{X}_3 = \begin{Bmatrix} 2 \\ 0 \end{Bmatrix}$$

2.62 Using Kuhn–Tucker conditions, find the value(s) of β for which the point $x_1^* = 1, x_2^* = 2$ will be optimal to the problem:

$$\text{Maximize } f(x_1,x_2) = 2x_1 + \beta x_2$$

subject to

$$g_1(x_1,x_2) = x_1^2 + x_2^2 - 5 \le 0$$

$$g_2(x_1, x_2) = x_1 - x_2 - 2 \le 0$$

Verify your result using a graphical procedure.

2.63 Consider the following optimization problem:

$$\text{Maximize } f = -x_1 - x_2$$

subject to

$$x_1^2 + x_2 \ge 2$$

$$4 \le x_1 + 3x_2$$

$$x_1 + x_2^4 \le 30$$

(a) Find whether the design vector $\mathbf{X} = \begin{Bmatrix} 1 \\ 1 \end{Bmatrix}$ satisfies the Kuhn–Tucker conditions for a constrained optimum.

(b) What are the values of the Lagrange multipliers at the given design vector?

2.64 Consider the following problem:

$$\text{Minimize } f(\mathbf{X}) = x_1^2 + x_2^2 + x_3^2$$

subject to

$$x_1 + x_2 + x_3 \ge 5$$

$$2 - x_2 x_3 \le 0$$

$$x_1 \ge 0, \qquad x_2 \ge 0, \qquad x_3 \ge 2$$

Determine whether the Kuhn–Tucker conditions are satisfied at the following points:

$$\mathbf{X}_1 = \begin{Bmatrix} \frac{3}{2} \\ \frac{3}{2} \\ 2 \end{Bmatrix}, \quad \mathbf{X}_2 = \begin{Bmatrix} \frac{4}{3} \\ \frac{2}{3} \\ 3 \end{Bmatrix}, \quad \mathbf{X}_3 = \begin{Bmatrix} 2 \\ 1 \\ 2 \end{Bmatrix}$$

2.65 Find a usable and feasible direction **S** at (**a**) $\mathbf{X}_1 = \begin{Bmatrix} -1 \\ 5 \end{Bmatrix}$ and (**b**) $\mathbf{X}_2 = \begin{Bmatrix} 2 \\ 3 \end{Bmatrix}$ for the following problem:

$$\text{Minimize } f(\mathbf{X}) = (x_1 - 1)^2 + (x_2 - 5)^2$$

subject to

$$g_1(\mathbf{X}) = -x_1^2 + x_2 - 4 \leq 0$$

$$g_2(\mathbf{X}) = -(x_1 - 2)^2 + x_2 - 3 \leq 0$$

2.66 Consider the following problem:

$$\text{Minimize } f = x_1^2 - x_2$$

subject to

$$26 \geq x_1^2 + x_2^2$$

$$x_1 + x_2 \geq 6$$

$$x_1 \geq 0$$

Determine whether the following search direction is usable, feasible, or both at the design vector $\mathbf{X} = \begin{Bmatrix} 5 \\ 1 \end{Bmatrix}$:

$$\mathbf{S} = \begin{Bmatrix} 0 \\ 1 \end{Bmatrix}, \quad \mathbf{S} = \begin{Bmatrix} -1 \\ 1 \end{Bmatrix}, \quad \mathbf{S} = \begin{Bmatrix} 1 \\ 0 \end{Bmatrix}, \quad \mathbf{S} = \begin{Bmatrix} -1 \\ 2 \end{Bmatrix}$$

2.67 Consider the following problem:

$$\text{Minimize } f = x_1^3 - 6x_1^2 + 11x_1 + x_3$$

subject to

$$x_1^2 + x_2^2 - x_3^2 \le 0$$

$$4 - x_1^2 - x_2^2 - x_3^2 \le 0$$

$$x_i \ge 0, \qquad i = 1,2,3, \qquad x_3 \le 5$$

Determine whether the following vector represents an optimum solution:

$$\mathbf{X} = \begin{Bmatrix} 0 \\ \sqrt{2} \\ \sqrt{2} \end{Bmatrix}$$

2.68
$$\text{Minimize } f = x_1^2 + 2x_2^2 + 3x_3^2$$

subject to the constraints

$$g_1 = x_1 - x_2 - 2x_3 \le 12$$

$$g_2 = x_1 + 2x_2 - 3x_3 \le 8$$

using Kuhn–Tucker conditions.

2.69
$$\text{Minimize } f(x_1,x_2) = (x_1 - 1)^2 + (x_2 - 5)^2$$

subject to

$$-x_1^2 + x_2 \le 4$$
$$-(x_1 - 2)^2 + x_2 \le 3$$

by (**a**) the graphical method and (**b**) Kuhn–Tucker conditions.

2.70
$$\text{Maximize } f = 8x_1 + 4x_2 + x_1x_2 - x_1^2 - x_2^2$$

subject to

$$2x_1 + 3x_2 \le 24$$

$$-5x_1 + 12x_2 \le 24$$

$$x_2 \le 5$$

by applying Kuhn–Tucker conditions.

2.71 Consider the following problem:

$$\text{Maximize } f(x) = (x - 1)^2$$

subject to

$$-2 \leq x \leq 4$$

Determine whether the constraint qualification and Kuhn–Tucker conditions are satisfied at the optimum point.

2.72 Consider the following problem:

$$\text{Minimize } f = (x_1 - 1)^2 + (x_2 - 1)^2$$

subject to

$$2x_2 - (1 - x_1)^3 \leq 0$$
$$x_1 \geq 0$$
$$x_2 \geq 0$$

Determine whether the constraint qualification and the Kuhn–Tucker conditions are satisfied at the optimum point.

2.73 Verify whether the following problem is convex:

$$\text{Minimize } f(\mathbf{X}) = -4x_1 + x_1^2 - 2x_1x_2 + 2x_2^2$$

subject to

$$2x_1 + x_2 \leq 6$$
$$x_1 - 4x_2 \leq 0$$
$$x_1 \geq 0, \quad x_2 \geq 0$$

2.74 Check the convexity of the following problems.

(a) $$\text{Minimize } f(\mathbf{X}) = 2x_1 + 3x_2 - x_1^3 - 2x_2^2$$

subject to

$$x_1 + 3x_2 \leq 6$$
$$5x_1 + 2x_2 \leq 10$$
$$x_1 \geq 0, \quad x_2 \geq 0$$

(b) $$\text{Minimize } f(\mathbf{X}) = 9x_1^2 - 18x_1x_2 + 13x_1 - 4$$

subject to

$$x_1^2 + x_2^2 + 2x_1 \geq 16$$