رقم المقرر ورمزه	اسم المقرر	توزيع ساعات الدراسة اسبوعياً			الاوراق الامتحانية		الدرجة
		ن	ع	ت	ن	ع	
454ر	بحوث عمليات (ii)	4	-	-	1	-	200

طرق الأمثلية التقليدية للمسائل ذات المتغير الواحد ومتعددة المتغيرات بقيود وبدون قيود- مسائل الامثلية الغير خطية وتبويبها- طرق حل مسألة البرمجة الغير خطية (للمتغير الواحد- متعددة المغيرات بدون قيود –متعددة المغيرات فى وجود قيود)-مقدمة في البرمجة الديناميكية

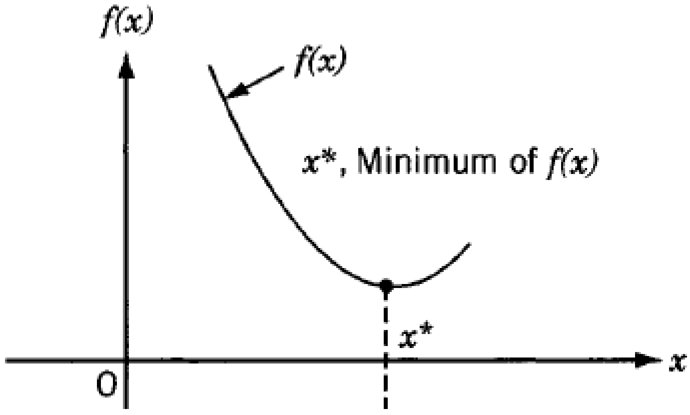
ch1_Classical Optimization Techniques
ch2_Unconstrained Multivariable Optimization
ch3 constrained optimization
ch5_Dynamic Programming

WENYU SUN and YA-XIAN G YUAN, OPTIMIZATION THEORY AND METHODS Nonlinear Programming, c 2006 Springer Science+Business Media, LLC

Ch1: Classical Optimization Techniques

SINGLE-VARIABLE OPTIMIZATION

A function of one variable $f(x)$ is said to have a relative or local minimum at $x = x^*$ if $f(x^*) < f(x^* + h)$ for all sufficiently small positive and negative values of h.



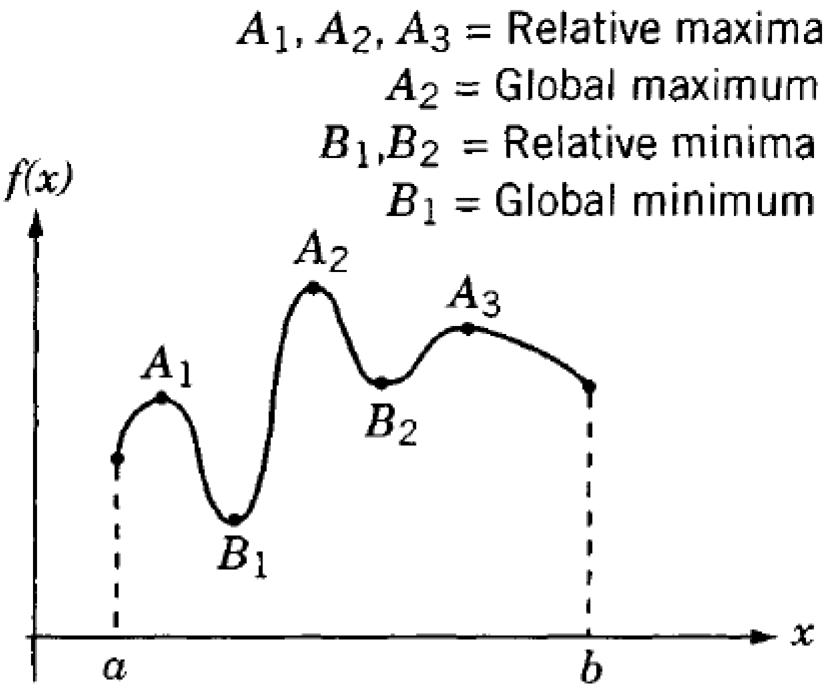


Figure 1

[Q1] (a) Write down the theorems that states the conditions for which the single variable optimization problem has its local optimum.

Theorem 1: Necessary Condition

If a function $f(x)$ is defined in the interval $a < x < b$ and have 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 $f'(x^*) = 0$.

Theorem 2: Sufficient Condition:

Let $f'(x^*) = f''(x^*) = \ldots = f^{(n-1)}(x^*) = 0$, 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.

(b) Use theorems in(a) to find the optimum values of

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

Answer:

$$f'(x) = 60x^4 - 3*60x^3 + 60*2*x^2$$
$$= 60x^2(x^2 - 3x + 2)$$
$$= 60x^2(x-1)(x-2) = 0$$

The extreme points are

$x = 0, x = 1$ and $x = 2$

$x = 0$	$x = 1$	$x = 2$
$f''(x) = 240x^3 - 540x^2 + 240x$ $f''(0) = 0$ We evaluate the next derivative $f'''(x) = 3*240x^2 - 2*540x + 240$ $f'''(0) = +240$, order of derivative is odd. So this point is neither maximum nor minimum	$f''(1) = -60$ this point is relative maximum $f_{Max} = 12(1) - 45\ (1) + 40(1) + 5$ $=12$	$f''(2) = 240$ this point is relative minimum $f_{Min} = -11$

[6]Detect which of the following Mathematical statements is true and which is false. Write the false one(s) in the correct case.

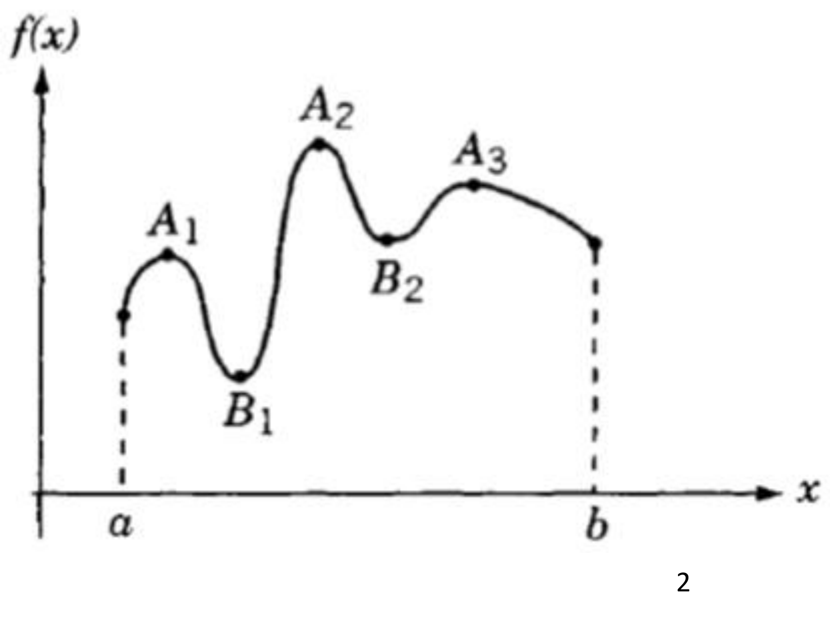


Figure 1

In Figure 1,

1	A_1 is relative minimum	×
2	A_2 is Global Maximum	√
3	A_3 is relative Maximum	×
4	B_1 is Global minimum	√
5	B_2 is Global minimum	×

1	The necessary condition for a function $f(x)$ to have a relative minimum at $x = x^*$, is $f'(x^*) = 0$.	√
2	The sufficient condition for a function $f(x)$ to have a relative minimum at $x = x^*$ depends on the order (even- or odd) of the first non zero derivative of $f(x)$	√
3		×
4		√
5		×

[7]Select the correct word

(1) A function of one variable $f(x)$ is said to have a relative or local minimum at $x = x^*$ if $f(x^*) \ldots f(x^* + h)$ for all sufficiently small positive and negative values of h.

(a)	$<$	(b)	$\geq$	(c)	$\leq$	(d)	

Consider using the necessary and sufficient condition to find the optimum values of $f(x) = 12x^5 - 45x^4 + 40x^3 + 5$. Answer the following questions:

$f'(x) = 60x^2(x-1)(x-2) = 0$

The extreme points are

$x = 0, x = 1$ and $x = 2$

$x = 0$	$x = 1$	$x = 2$
$f''(x)= 240x^3 - 540x^2 + 240x$ $f''(0) = 0$ We evaluate the next derivative $f'''(x)= 3 * 240x^2 - 2 * 540x + 240$ $f'''(0)= +240$, order of derivative is odd. So this point is neither maximum nor minimum	$f''(1) = -60$ this point is relative maximum $f_{Max} = 12(1) - 45 \ (1) + 40(1) + 5$ $=12$	$f''(2) = 240$ this point is relative minimum $f_{Min} = -11$

$$f'(x) = ax^2(x-b)(x-c)$$

(1) $a =$

(a)	60	(b)		(c)		(d)	

(1) $b =$

(a)		(b)	1	(c)		(d)	

(1) $c =$

(a)	2	(b)		(c)		(d)	

The extreme point $x = \cdots$ is neither maximum nor minimum

(a)	0	(b)		(c)		(d)	

The extreme point $x = \cdots$ is relative maximum

(a)	1	(b)		(c)		(d)	

The extreme point $x = \cdots$ is relative minimum

(a)		(b)		(c)	2	(d)	

<u>Excercises:</u>

(3)Find the maxima and minima, if any, of the functions

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

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

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

2.3 MULTIVARIABLE OPTIMIZATION WITH NO CONSTRAINTS

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 X*, the polynomial

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

is called the r th differential of f at X*.

For Example :

when $r = 1$ and $n = 3$, we have

$$df(X^*) = \sum_{i=1}^{3} h_i \frac{\partial f}{\partial x_i} = h_1 \frac{\partial f}{\partial x_1} + h_2 \frac{\partial f}{\partial x_2} + h_3 \frac{\partial f}{\partial x_3}$$

Which corresponds $df = \frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2 + \frac{\partial f}{\partial x_3} dx_3$

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(X) near a point X* is given by

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

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

Example 3 *:* Find the second-order Taylor's series approximation of the function

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

near 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

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

where

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

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

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

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

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

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

$$+ 2h_1 h_3(e^{x_3})] \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} = -4h_2^2 + e^{-2}h_3^2 + 2h_1 h_3 e^{-2}$$

Thus the Taylor's series approximation is given by

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

[Q2] (a) State and prove the theorems for which the multivariable unconstrained optimization problem has its local optimum.

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

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

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,

X=X*+h

$$f(\mathbf{X}^* + \mathbf{h}) = f(\mathbf{X}^*) + \sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(\mathbf{X}^*) + \frac{1}{2!} d^2 f(\mathbf{X}^* + \theta \mathbf{h}),$$

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

$$f(\mathbf{X}^* + \mathbf{h}) = f(\mathbf{X}^*) + \sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(\mathbf{X}^*) + \frac{1}{2!} \sum_{i=1}^{n} \sum_{j=1}^{n} h_i h_j \left. \frac{\partial^2 f}{\partial x_i\, \partial x_j} \right|_{\mathbf{X} = \mathbf{X}^* + \theta \mathbf{h}}, \qquad 0 < \theta < 1 \tag{2.10}$$

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

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

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

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

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

$$\begin{bmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_1 \partial x_3} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2 \partial x_2} & \frac{\partial^2 f}{\partial x_2 \partial x_3} \\ \frac{\partial^2 f}{\partial x_3 \partial x_1} & \frac{\partial^2 f}{\partial x_3 \partial x_2} & \frac{\partial^2 f}{\partial x_3 \partial x_3} \end{bmatrix}$$

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

==

Definition:

[Q3] (a) State 2 different definitions for the positiveness of a square matrix. Then show which of them is suitable in application for Hessian Matrix.

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

$$|A - \lambda I| = 0$$

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

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

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

$$A_3 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{32} \end{vmatrix}, \ldots,$$

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

The matrix **A** will be positive definite if and only if all the values $A_1, A_2, A_3, \ldots, A_n$ 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.

A matrix A will be positive definite if and only if all its determinants are positive;
A matrix A will be negative definite if and only if all its determinant A_k satisfies: $(-1)^k, k = 1,2,..$
A matrix A will be semidefinite definite if some of its determinant are positive, and the remaining are zeros

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*).

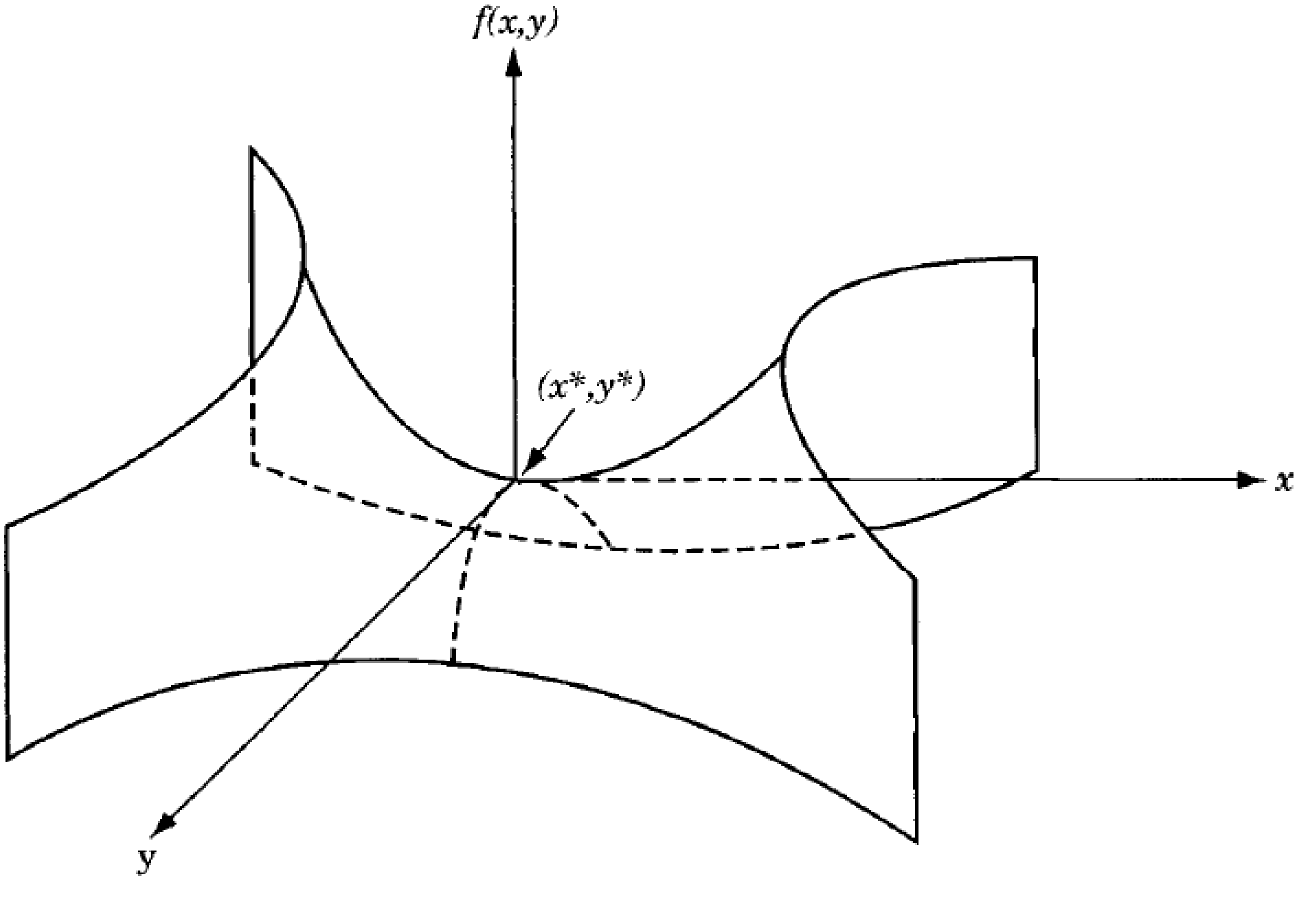


Figure 2.5 Saddle point of the function $f(x,y) = x^2 - y^2$.

1	A matrix A will be positive definite if all its eigenvalues are positive;	√
2	A matrix A will be positive definite if and only if all its determinants are positive;	√
3	A matrix A will be negative definite if and only if all its determinant A_k satisfies: $(-1)^k$	√
4	A matrix A will be semidefinite definite if some of its determinant are positive, and the remaining are zeros	√
5	A saddle point is corresponds to a relative minimum of f(x,*y)* with respect to one variable and a relative maximum with respect to the second variable	√

(1) A matrix A will be negative definite if and only if all its determinant A_k satisfies:

(a)	$(-1)^k$	(b)	$(1)^k$	(c)	$(k)^{-1}$	(d)	else

[Q3] (a) Define: positive definiteness of a square matrix- Semidefinite Case- Saddle Point

[Q3] (b) Example : Find the extreme points of the function

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

The necessary condition is

$\frac{\partial f}{\partial x_1} = 3x_1^2 + 4x_1 = x_1(3x_1 + 4) = 0$ $\frac{\partial f}{\partial x_2} = 3x_2^2 + 8x_2 = x_2(3x_2 + 8) = 0$	So $x_1(3x_1 + 4) = 0$ $x_2(3x_2 + 8) = 0$

$$x_1 = 0, \frac{-4}{3}, \; x_2 = 0, \frac{-8}{3}$$

Theses equations are satisfied at the points

$$(0,0), \left(0, \frac{-8}{3}\right), \left(\frac{-4}{3}, 0\right), \left(\frac{-4}{3}, \frac{-8}{3}\right)$$

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

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

$\frac{\partial f}{\partial x_1} = 3x_1^2 + 4x_1$ $\frac{\partial f}{\partial x_2} = 3x_2^2 + 8x_2$	$\frac{\partial^2 f}{\partial x_1^2} = 6x_1 + 4, \frac{\partial^2 f}{\partial x_2^2} = 6x_2 + 8, \frac{\partial^2 f}{\partial x_1 \partial x_2} = 0$
$\begin{bmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2 \partial x_2} \end{bmatrix}$	The Hessian matrix of f is given by $\begin{bmatrix} 6x_1 + 4 & 0 \\ 0 & 6x_2 + 8 \end{bmatrix}$

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.

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

Point **X**	Value of J_1	Value of J_2	Nature of **J**	Nature of **X**	f(**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

Consider finding the extreme points of the function $f(x_1, x_2) = x_1^3 + x_2^3 + 2x_1^2 + 4x_2^2 + 6$. Answer the following

(1) The necessary condition yields

(a)	$x_1(3x_1 + 4) = 0, x_2(3x_2 + 8) = 0$	(b)	$x_1(3x_1 - 4) = 0, x_2(3x_2 + 8) = 0$
(c)	$x_1(3x_1 + 4) = 0, x_2(3x_2 - 8) = 0$	(d)	Else

(2) The solutions of the necessary condition equations are

(a)	$x_1 = 0, \frac{4}{3},\ x_2 = 0, \frac{-8}{3}$	(b)	$x_1 = 0, \frac{-4}{3},\ x_2 = 0, \frac{-8}{3}$
(c)	$x_1 = 0, \frac{-4}{3},\ x_2 = 0, \frac{8}{3}$	(d)	Else

(3) The necessary condition equations are satisfied at the points

(a)	$(0,0), \left(\frac{-8}{3}, 0\right), \left(0, \frac{-4}{3}\right), \left(\frac{-4}{3}, \frac{-8}{3}\right)$	(b)	$(0,0), \left(0, \frac{-8}{3}\right), \left(\frac{-4}{3}, 0\right), \left(\frac{-4}{3}, \frac{-8}{3}\right)$
(c)	$(0,0), \left(\frac{8}{3}, 0\right), \left(0, \frac{4}{3}\right), \left(\frac{4}{3}, \frac{8}{3}\right)$	(d)	Else

(3) The following point satisfies the necessary condition

(a)	$\left(0, \frac{-4}{3}\right)$	(b)	$\left(\frac{-4}{3}, \frac{-8}{3}\right)$
(c)	$\left(\frac{-8}{3}, 0\right)$	(d)	$\left(\frac{4}{3}, \frac{8}{3}\right)$

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

(a)	$\frac{\partial^2 f}{\partial {x_1}^2} = 0, \frac{\partial^2 f}{\partial {x_2}^2} = 6x_2 + 8, \frac{\partial^2 f}{\partial x_1 \partial x_2} = 6x_1 + 4$	(b)	$\frac{\partial^2 f}{\partial {x_1}^2} = 6x_1 + 4, \frac{\partial^2 f}{\partial {x_2}^2} = 0, \frac{\partial^2 f}{\partial x_1 \partial x_2} = 6x_2 + 8$
(c)	$\frac{\partial^2 f}{\partial {x_1}^2} = 6x_1 + 4, \frac{\partial^2 f}{\partial {x_2}^2} = 6x_2 + 8, \frac{\partial^2 f}{\partial x_1 \partial x_2} = 0$	(d)	Else

(5) The Hessian matrix of f is given by $\begin{bmatrix} 6x_1 + 4 & 0 \\ 0 & 6x_2 + 8 \end{bmatrix}$

(a)	$\begin{bmatrix} 0 & 6x_1 + 4 \\ 0 & 6x_2 + 8 \end{bmatrix}$	(b)	$\begin{bmatrix} 6x_1 + 4 & 0 \\ 0 & 6x_2 + 8 \end{bmatrix}$
(c)	$\begin{bmatrix} 6x_1 + 4 & 0 \\ 6x_2 + 8 & 0 \end{bmatrix}$	(d)	Else

(6) The nature of the extreme point $(0,0)$ is

(a)	Relative minimum	(b)	Saddle point	(c)	Relative maximum	(d)	else

(7) The nature of the extreme point $\left(0, \frac{-8}{3}\right)$ is

(a)	Relative minimum	(b)	Saddle point	(c)	Relative maximum	(d)	else

(8) The nature of the extreme point $\left(\frac{-4}{3}, 0\right)$ is

(a)	Relative minimum	(b)	Saddle point	(c)	Relative maximum	(d)	else

(9) The nature of the extreme point $\left(\frac{-4}{3}, \frac{-8}{3}\right)$ is

(a)	Relative minimum	(b)	Saddle point	(c)	Relative maximum	(d)	else

(10) The Relative maximum of the function is

(a)	194/27	(b)	6	(c)	50/3	(d)	else

[1] 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) 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)

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

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

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

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

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

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

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

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

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

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

(6) 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

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

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

(7) 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.

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.

Multivariable Optimization With Equality Constraints

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

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

$$\text{subject to} \qquad (2.16)$$

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

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.

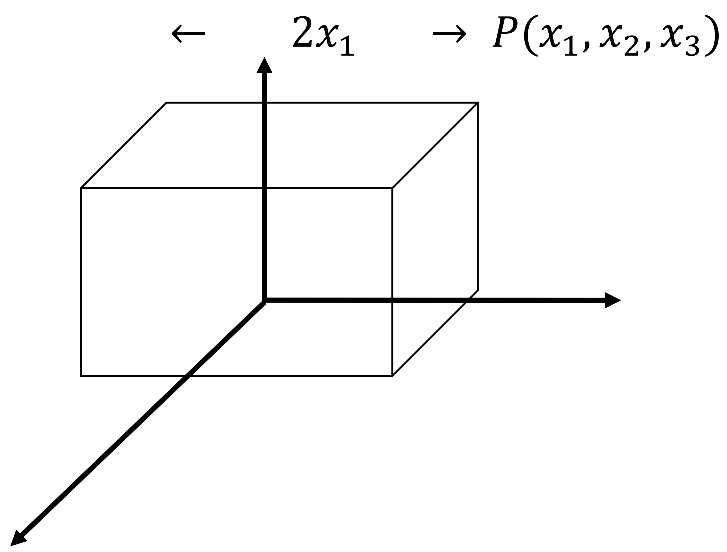
Solution by Direct Substitution

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.

[Q4] 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

أوجد أبعاد صندوق بحيث يكون له أكبر حجم يمكن احتواؤه في كرة نصف قطرها الوحدة.



volume of the box is given by $f(x_1,x_2,x_3) = 8x_1x_2x_3$

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

$$x_1^2 + x_2^2 + x_3^2 = 1$$

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

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

$$f(x_1,x_2) = 8x_1x_2\sqrt{1 - x_1^2 - x_2^2}$$

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

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

For the sufficient condition, it is clear that the Hessian matrix is negative definite. Hence the point X_1 is maximum for the given function.

1	In the equality constraints optimization problem, the number of constraints must be less than or equal to the number of variable.	√
2	If the number of constraints is greater than the number of variable in the equality constraints optimization problem, the problem becomes overdefined	√
3	Max. $f(x_{1,}x_2,x_3) = 8x_1x_2x_3$ subject to $x_1^2 + x_2^2 + x_3^2 = 1$ is equivalent to Max. $f(x_{1,}x_2) = 8x_1x_2\sqrt{1 - x_1^2 - x_2^2}$	√
4		√
5		√

===

2.48 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,

[Q2] Consider the problem

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

Subject to

$$g_1(X) = x_1 - x_2 = 0$$
$$g_2(X) = x_1 + x_2 + x_3 = 1$$

By

(a) Direct substitution

(b) [Q3] find the value of x, y, and z that maximize the function

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

(d) When x, y, and z are restricted by the relation xyz = 16.

Solution by the method of constrained variation

[] (a)Discuss the basic idea of method of constrained variation for solving multivariable optimization with equality constraints. Then show when the variation dx1,dx2 form admissible variation.

(b) Derive the necessary condition in order to have X^* as an extreme point (minimum or maximum).

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

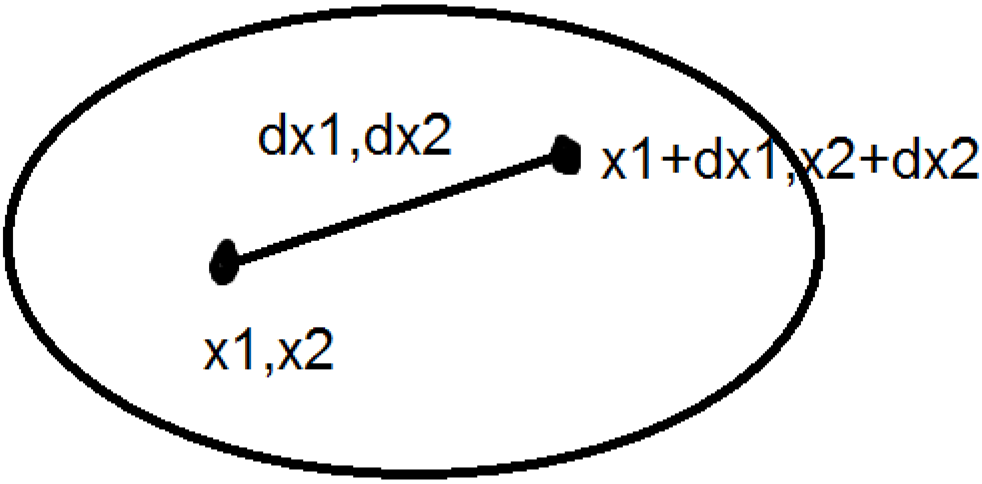
$$g(x_1,x_2) = 0 \tag{2.18}$$

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:

Define: admissible variation.



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

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

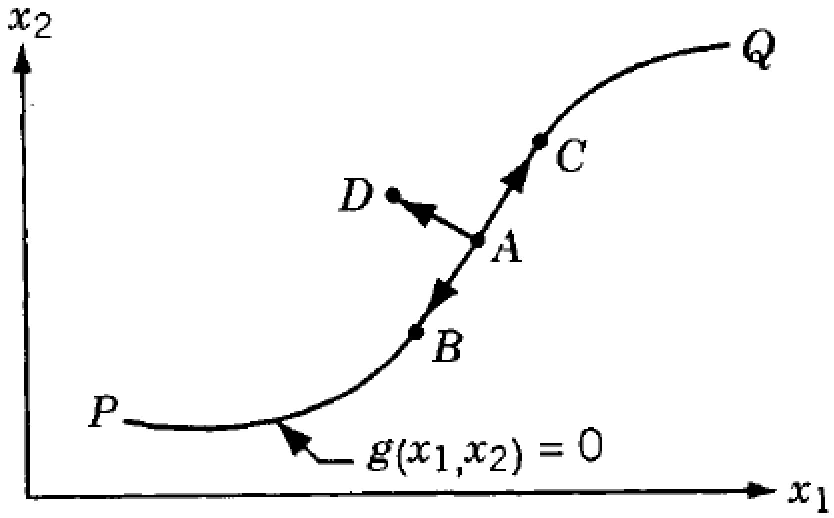
$$g(x_1^* + dx_1, x_2^* + dx_2)$$

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

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.



(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).

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

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

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

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

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).

==========================

[](c) **Use** the necessary condition derived in (b) to **find the solution of the optimization problem**

$$\text{Minimize } f = k/xy^2$$
$$\text{Subject to } x^2 + y^2 = a^2$$

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

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

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

that is,

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

Thus The optimum values of x and y can be obtained from Eqs. (E_3) and (E_2) as

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

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

==================================

Solution by the Method of Lagrange Multipliers

[](a) Derive the necessary condition for optimality in Lagrange multiplier method for a problem of two variables and one constraint.

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:

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

subject to

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

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

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

Dividing by $\frac{\partial g}{\partial x_2} \neq 0$

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

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

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

Equation (2.32) can be expressed as

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

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

and Eq. (2.33) can be written as

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

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

[q6](b) Example 2.9
Find the solution of

$$\text{Minimize } f = k/xy^2$$

$$\text{Subject to } x^2 + y^2 = a^2$$

using the necessary condition of Lagrange multiplier method
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

$$\frac{\partial L}{\partial x} = -kx^{-2}y^{-2} + 2x\lambda = 0 \qquad (E_1)$$

$$\frac{\partial L}{\partial y} = -2kx^{-1}y^{-3} + 2y\lambda = 0 \qquad (E_2)$$

$$\frac{\partial L}{\partial \lambda} = x^2 + y^2 - a^2 = 0 \qquad (E_3)$$

Equations (E1) and (E2) yield

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

$$\frac{1}{x^2} = \frac{2}{y^2}$$

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

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

==============================

[Q5] State the necessary and sufficiency conditions for optimality of a General Lagrange multiplier method.

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)

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

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

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

$$L(X,\lambda) = f(X) + \sum_{j=1}^{m} \lambda_j g_j(X)$$

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

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

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

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

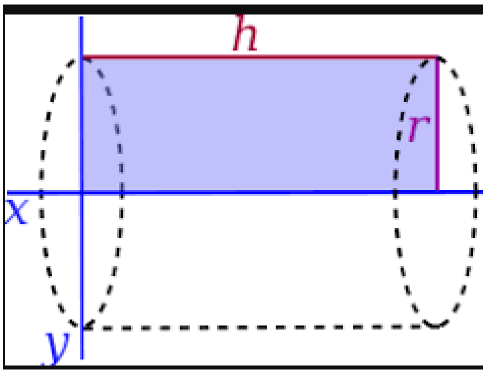
Theorem 2.6: Sufficient Condition A sufficient condition for $f(\mathbf{X})$ to have a relative minimum at $\mathbf{X}^*$ is that the quadratic, Q, defined by

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

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

[Q5](b)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 $A_0 = 24\pi$.



Let the radius of the tin is $r = x_1$ and the length is $h = x_2$. 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

$$L(X, \lambda) = f(X) + \sum_{j=1}^{m} \lambda_j g_j(X)$$

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

and the necessary conditions for the maximum of f give

$$\frac{\partial L}{\partial x_1} = 2\pi x_1 x_2 + 4\pi\lambda x_1 + 2\pi\lambda x_2 = 0 \qquad (E_1)$$

$$\frac{\partial L}{\partial x_2} = \pi x_1^2 + 2\pi\lambda x_1 = 0 \qquad (E_2)$$

$$\frac{\partial L}{\partial \lambda} = 2\pi x_1^2 + 2\pi x_1 x_2 - A_0 = 0 \qquad (E_3)$$

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

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

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

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

Now since

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

$$\frac{\partial^2 L}{\partial x_1 \partial \lambda}\bigg|_{(\mathbf{X}^*, \lambda^*)} = 4\pi x_1^* + 2\pi x_2^* = 16\pi$$

And since

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

$$\frac{\partial^2 L}{\partial x_2 \partial \lambda}\bigg|_{(\mathbf{X}^*, \lambda^*)} = 2\pi x_1^* = 4\pi$$

And since

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

$$\frac{\partial^2 L}{\partial \lambda \partial \lambda} = 0$$

$$\begin{bmatrix} \frac{\partial^2 L}{\partial x_1 \partial x_1} & \frac{\partial^2 L}{\partial x_1 \partial x_2} & \frac{\partial^2 L}{\partial x_1 \partial \lambda} \\ \frac{\partial^2 L}{\partial x_2 \partial x_1} & \frac{\partial^2 L}{\partial x_2 \partial x_2} & \frac{\partial^2 L}{\partial x_2 \partial \lambda} \\ \frac{\partial^2 L}{\partial \lambda \partial x_1} & \frac{\partial^2 L}{\partial \lambda \partial x_2} & \frac{\partial^2 L}{\partial \lambda \partial \lambda} \end{bmatrix}$$

$$H = \begin{bmatrix} 4\pi & 2\pi & 16\pi \\ 2\pi & 0 & 4\pi \\ 16\pi & 4\pi & 0 \end{bmatrix}$$

$$J_1 = [4\pi] = +$$

$$J_2 = \begin{bmatrix} 4\pi & 2\pi \\ 2\pi & 0 \end{bmatrix} = -4\pi^2$$

J is Indefinit

So the point is not max. nor min.

Another method for test the positiveness of Hessian matrix:

$$H = \begin{bmatrix} 4\pi & 2\pi & 16\pi \\ 2\pi & 0 & 4\pi \\ 16\pi & 4\pi & 0 \end{bmatrix}$$

$$|H - \lambda I| = 0$$

$$\begin{bmatrix} 4\pi - \lambda & 2\pi & 16\pi \\ 2\pi & 0 - \lambda & 4\pi \\ 16\pi & 4\pi & 0 - \lambda \end{bmatrix} = 0$$

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

$$\lambda = -\frac{12}{17}\pi$$

Since the value of λ is negative, the point $({x_1}^*, {x_2}^*)$corresponds to the maximum Of f.

****

==========================

[Q5](c) 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.

SOLUTION The Lagrange function is given by

$$L(\mathbf{X},\lambda) = 2x_1 + x_2 + 10 + \lambda(3 - x_1 - 2x_2^2) \qquad (E_1)$$

The necessary conditions for the solution of the problem are

$$\frac{\partial L}{\partial x_1} = 2 - \lambda = 0$$

$$\frac{\partial L}{\partial x_2} = 1 - 4\lambda x_2 = 0 \qquad (E_2)$$

$$\frac{\partial L}{\partial \lambda} = 3 - x_1 - 2x_2^2 = 0$$

The solution of Eqs. (E_2) is

$$\mathbf{X}^* = \begin{Bmatrix} x_1^* \\ x_2^* \end{Bmatrix} = \begin{Bmatrix} 2.97 \\ 0.13 \end{Bmatrix}$$

$$\lambda^* = 2.0 \qquad (E_3)$$

Sufficient condition is HomeWork

======================

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

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

2.48

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

subject to

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

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

[Q6] Consider the problem

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

Subject to

$$g_1(X) = x_1 - x_2 = 0$$
$$g_2(X) = x_1 + x_2 + x_3 = 1$$

By
(e) Constrained variation, and
(f) Lagrange multipliers method.

==

[Q7] (b) Minimize $f(X) = \frac{x_1^2 + x_2^2 + x_3^2}{2}$ (1)

Subject to $g_1(X) = x_1 - x_2 = 0,$ (2)

$g_2(X) = x_1 + x_2 + x_3 = 1$ (3)

By Lagrange multipliers method.
Answer:

Substituting from 2 in 1 and 3, we obtain

$f(X) = \frac{2x_1^2+x_3^2}{2}$, , $g(X) = 2x_1 + x_3 - 1 = 0$

the lagrange function is

$$L(x_1, x_3, \lambda) = \frac{1}{2}(2x_1^2 + x_3^2) + \lambda(2x_1 + x_3 - 1)$$

$$\frac{\partial L}{\partial x_1} = 2x_1 + 2\lambda = 0$$

$$\frac{\partial L}{\partial x_3} = x_3 + \lambda = 0$$

$$\frac{\partial L}{\partial \lambda} = 2x_1 + x_3 - 1$$

$$\therefore \lambda^* = -x_3^* = -x_1^* \quad \therefore x_1^* = x_3^*$$

$$g(X) = 2x_1^* + x_3^* - 1 = 0 \quad \therefore 3x_1^* - 1 = 0$$

$$x_1^* = \frac{1}{3}, \quad x_3^* = \frac{1}{3}, \quad x_2^* = \frac{1}{3}$$

==

[Q8] find the value 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.

Ch 2
Unconstrained Multivariable Optimization Techniques

This chapter deals with the various methods of solving the unconstrained minimization problem:

$$\text{Find } \mathbf{X} = \begin{Bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{Bmatrix} \text{ which minimizes } f(\mathbf{X}) \tag{6.1}$$

It is true that rarely a practical design problem would be unconstrained; still, a study of this class of problems is important for the following reasons:

1. The constraints do not have significant influence in certain design problems.
2. Some of the powerful and robust methods of solving constrained minimization problems require the use of unconstrained minimization techniques.
3. The study of unconstrained minimization techniques provide the basic understanding necessary for the study of constrained minimization methods.
4. The unconstrained minimization methods can be used to solve certain complex engineering analysis problems. For example, the displacement response (linear or nonlinear) of any structure under any specified load condition can be found by minimizing its potential energy. Similarly, the eigenvalues and eigenvectors of any discrete system can be found by minimizing the Rayleigh quotient.

As discussed in Chapter 1, a point X* will be a relative minimum of f (X) if the necessary conditions

$$\frac{\partial f}{\partial x_i}(\mathbf{X} = \mathbf{X}^*) = 0, \qquad i = 1,2,\ldots,n \tag{6.2}$$

are satisfied. The point $\mathbf{X}^*$ is guaranteed to be a relative minimum if the Hessian matrix is positive definite, that is,

$$\mathbf{J}_{\mathbf{X}^*} = [J]_{\mathbf{X}^*} = \left[\frac{\partial^2 f}{\partial x_i\, \partial x_j}(\mathbf{X}^*)\right] = \text{positive definite} \tag{6.3}$$

Equations (6.2) and (6.3) can be used to identify the optimum point during numerical computations. However, if the function is not differentiate, Eqs. (6.2) and (6.3) cannot be applied to identify the optimum point.

6.1.1 Classification of Unconstrained Minimization Methods

Several methods are available for solving an unconstrained minimization problem. These methods can be classified into two broad categories as direct search methods and descent methods as indicated in Table 6.1.

TABLE 6.1 Unconstrained Minimization Methods

Direct Search Methods [a]	Descent Methods [b]
Random search method	Steepest descent (Cauchy) method
Grid search method	Fletcher–Reeves method
Univariate method	Newton's method
Pattern search methods	Marquardt method
Powell's method	Quasi-Newton methods
Hooke–Jeeves method	Davidon–Fletcher–Powell method
Rosenbrock's method	Broyden–Fletcher–Goldfarb–Shanno method
Simplex method	

[a] Do not require the derivatives of the function.
[b] Require the derivatives of the function.

6.1.2 General Approach

All the unconstrained minimization methods are iterative in nature and hence they start from an initial trial solution and proceed toward the minimum point in a sequential manner. The general iterative scheme is shown in Fig. 6.3 as a flow diagram. It is important to note that all the unconstrained minimization methods (1) require an initial point $\mathbf{X}_1$ to start the iterative procedure, and (2) differ from one another only in the method of generating the new point $\mathbf{X}_{i+1}$ (from $\mathbf{X}_i$) and in testing the point $\mathbf{X}_{i+1}$ for optimality.

[Q1]Draw the flowchart of general iterative scheme of unconstrained multivariable optimization

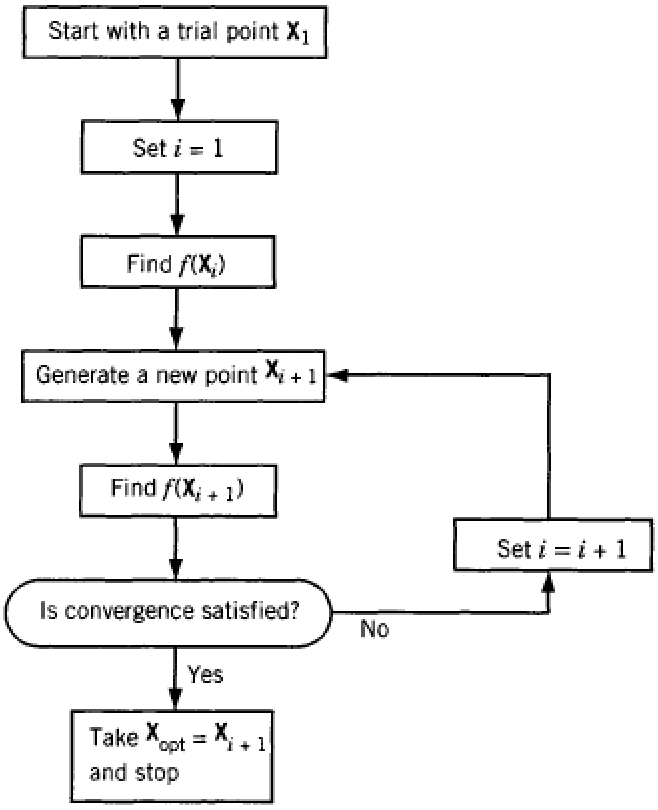


Figure 6.3 General iterative scheme of optimization.

6.1.3 Rate of Convergence

Define

Different iterative optimization methods have different rates of convergence. In general, an optimization method is said to have convergence of order p if

$$\frac{\|\mathbf{X}_{i+1} - \mathbf{X}^*\|}{\|\mathbf{X}_i - \mathbf{X}^*\|^p} \le k, \qquad k \ge 0, \quad p \ge 1 \tag{6.4}$$

where $\mathbf{X}_i$ and $\mathbf{X}_{i+1}$ denote the points obtained at the end of iterations i and $i + 1$, respectively, $\mathbf{X}^*$ represents the optimum point, and $\|\mathbf{X}\|$ denotes the length or norm of the vector $\mathbf{X}$:

$$\|\mathbf{X}\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2} \tag{6.5}$$

If $p = 1$ and $0 \le k \le 1$, the method is said to be linearly convergent (corresponds to slow convergence), If $p = 2$, the method is said to be quadratically convergent (corresponds to fast convergence). An optimization method is said to have superlinear convergence (corresponds to fast convergence) if

$$\lim_{i \to \infty} \frac{\|\mathbf{X}_{i+1} - \mathbf{X}^*\|}{\|\mathbf{X}_i - \mathbf{X}^*\|} \to 0 \tag{6.6}$$

The definitions of rates of convergence given in Eqs. (6.4) and (6.6) are applicable to single-variable as well as multivariable optimization problems. In the case of single-variable problems, the vector, $\mathbf{X}_i$, for example, degenerates to a scalar, x_i.

MCQ

1	Some of the methods for solving constrained minimization problems require the use of unconstrained minimization techniques.	√
2	The study of unconstrained minimization techniques provides the basic understanding necessary for the study of constrained minimization methods.	√
3	All the unconstrained minimization methods are iterative in nature.	√
4	Different iterative optimization methods have the same rates of convergence.	x
5	If we move along the gradient direction from any point in n-dimensional space, the function value increases at the fastest rate.	√
6	the gradient vector represents the direction of steepest descent.	x
7	the maximum rate of change of ***f*** at any point ***X*** is equal to the magnitude of the gradient vector at the same point.	√
8	Unconstrained Minimization Methods can be classified into two categories as direct search methods and descent methods.	√
9	Direct Search Methods require the derivatives of the function.	x
10	Descent Methods Require the derivatives of the function.	√

4 Scaling of Design Variables

The rate of convergence of most unconstrained minimization methods can be improved by scaling the design variables. For a quadratic objective function, the scaling of the design variables changes the **condition number** of the Hessian matrix. When the condition number of the Hessian matrix is 1, the steepest descent method, for example, finds the minimum of a quadratic objective function in one iteration.

If

$$f = \frac{1}{2}\mathbf{X}^T[A]\mathbf{X}$$

denotes a quadratic term, a transformation of the form

$$\mathbf{X} = [R]\mathbf{Y} \quad \text{or} \quad \begin{Bmatrix} x_1 \\ x_2 \end{Bmatrix} = \begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix} \begin{Bmatrix} y_1 \\ y_2 \end{Bmatrix} \tag{6.7}$$

can be used to obtain a new quadratic term a

$$\frac{1}{2}\mathbf{Y}^T[\tilde{A}]\mathbf{Y} = \frac{1}{2}\mathbf{Y}^T[R]^T[A][R]\mathbf{Y} \tag{6.8}$$

The matrix $[R]$ can be selected to make $[\tilde{A}] = [R]^T[A][R]$ diagonal (i.e., to eliminate the mixed quadratic terms). For this, the columns of the matrix $[R]$ are to be chosen as the eigenvectors of the matrix $[A]$. Next the diagonal elements of the matrix $[\tilde{A}]$ can be reduced to 1 (so that the condition number of the resulting matrix will be 1) by using the transformation

$$\mathbf{Y} = [S]\mathbf{Z} \quad \text{or} \quad \begin{Bmatrix} y_1 \\ y_2 \end{Bmatrix} = \begin{bmatrix} s_{11} & 0 \\ 0 & s_{22} \end{bmatrix} \begin{Bmatrix} z_1 \\ z_2 \end{Bmatrix} \tag{6.9}$$

where the matrix [S] is given by

$$[S] = \begin{bmatrix} s_{11} = \frac{1}{\sqrt{\tilde{a}_{11}}} & 0 \\ 0 & s_{22} = \frac{1}{\sqrt{\tilde{a}_{22}}} \end{bmatrix} \tag{6.10}$$

Thus the complete transformation that reduces the Hessian matrix of f to an identity matrix is given by

$$\mathbf{X} = [R][S]\mathbf{Z} \equiv [T]\mathbf{Z} \tag{6.11}$$

so that the quadratic term $\frac{1}{2}\mathbf{X}'[A]\mathbf{X}$ reduces to $\frac{1}{2}\mathbf{Z}'[I]\mathbf{Z}$.

If the objective function is not a quadratic, the Hessian matrix and hence the transformations vary with the design vector from iteration to iteration. For example, the second-order Taylor's series approximation of a general nonlinear function at the design vector $\mathbf{X}_i$ can be expressed as

$$f(\mathbf{X}) = c + \mathbf{B}^T\mathbf{X} + \frac{1}{2}\mathbf{X}^T[A]\mathbf{X} \tag{6.12}$$

where

$$c = f(\mathbf{X}_i) \tag{6.13}$$

$$\mathbf{B} = \begin{Bmatrix} \left.\dfrac{\partial f}{\partial x_1}\right|_{\mathbf{X}_i} \\ \vdots \\ \left.\dfrac{\partial f}{\partial x_n}\right|_{\mathbf{X}_i} \end{Bmatrix} \tag{6.14}$$

$$[A] = \begin{bmatrix} \left.\dfrac{\partial^2 f}{\partial x_1^2}\right|_{\mathbf{X}_i} & \cdots & \left.\dfrac{\partial^2 f}{\partial x_1\, \partial x_n}\right|_{\mathbf{X}_i} \\ \vdots & & \vdots \\ \left.\dfrac{\partial^2 f}{\partial x_n\, \partial x_1}\right|_{\mathbf{X}_i} & \cdots & \left.\dfrac{\partial^2 f}{\partial x_n^2}\right|_{\mathbf{X}_i} \end{bmatrix} \tag{6.15}$$

The transformations indicated by Eqs. (6.7) and (6.9) can be applied to the matrix $[A]$ given by Eq. (6.15). The procedure of scaling the design variables is illustrated with the following example.

The condition number of an *n Xn* matrix, *[A]*, is defined as

$$\text{cond}([A]) = \|[A]\|\; \|[A]^{-1}\| \geq 1$$

where $\|[A]\|$ denotes a norm of the matrix $[A]$. For example, the infinite norm of $[A]$ is defined as the maximum row sum given by

$$\|[A]\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^{n} |a_{ij}|$$

If the condition number is close to 1, the round-off errors are expected to be small in dealing with the matrix $[A]$. For example, if cond$[A]$ is large, the solution vector $\mathbf{X}$ of the system of equations $[A]\mathbf{X} = \mathbf{B}$ is expected to be very sensitive to small variations in $[A]$ and $\mathbf{B}$. If cond$[A]$ is close to 1, the matrix $[A]$ is said to be *well behaved* or *well conditioned*. On the other hand, if cond$[A]$ is significantly greater than 1, the matrix $[A]$ is said to be *not well behaved* or *ill conditioned*.

If the condition number is close to 1, the round-off errors are expected to be small in dealing with the matrix H. For example, if condH is large, the solution vector X of the system of equations HX = B is expected to be very sensitive to small variations in *[A]* and B. If cond H is close to 1, the matrix *[A]* is said to be *well behaved* or *well conditioned*. On the other hand, if cond H is significantly greater than 1, the matrix *[A]* is said to be *not well behaved* or *ill conditioned.*

MCQ

Example 6.2

Find a suitable scaling (or transformation) of variables to reduce the condition number of the Hessian matrix of the following function to 1:

$$f(x_1, x_2) = 6x_1^2 - 6x_1x_2 + 2x_2^2 - x_1 - 2x_2 \qquad (E_1)$$

SOLUTION

The quadratic function can be expressed as

$$f(\mathbf{X}) = \mathbf{B}^T\mathbf{X} + \tfrac{1}{2}\mathbf{X}^T[A]\mathbf{X} \qquad (E_2)$$

where

$$\mathbf{X} = \begin{Bmatrix} x_1 \\ x_2 \end{Bmatrix}, \quad \mathbf{B} = \begin{Bmatrix} -1 \\ -2 \end{Bmatrix}, \quad \text{and} \quad [A] = \begin{bmatrix} 12 & -6 \\ -6 & 4 \end{bmatrix}$$

As indicated above, the desired scaling of variables can be accomplished in two stages.

Stage 1: Reducing $[A]$ to a Diagonal Form, $[\tilde{A}]$

The eigenvectors of the matrix $[A]$ can be found by solving the eigenvalue problem

$$[[A] - \lambda_i[I]]\,\mathbf{u}_i = \mathbf{0} \qquad (E_3)$$

where λ_i is the ith eigenvalue and $\mathbf{u}_i$ is the corresponding eigenvector. In the present case, the eigenvalues, λ_i, are given by

$$\begin{vmatrix} 12 - \lambda_i & -6 \\ -6 & 4 - \lambda_i \end{vmatrix} = \lambda_i^2 - 16\lambda_i + 12 = 0 \qquad (E_4)$$

which yield $\lambda_1 = 8 + \sqrt{52} = 15.2111$ and $\lambda_2 = 8 - \sqrt{52} = 0.7889$. The eigenvector $\mathbf{u}_i$ corresponding to λ_i can be found by solving Eq. (E_3):

$$[[A] - \lambda_i[I]]\, \mathbf{u}_i = \mathbf{0}$$

س

$$\begin{bmatrix} 12 - \lambda_1 & -6 \\ -6 & 4 - \lambda_1 \end{bmatrix} \begin{Bmatrix} u_{11} \\ u_{21} \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix} \quad \text{or} \quad (12 - \lambda_1)u_{11} - 6u_{21} = 0$$

$$\text{or} \quad u_{21} = -0.5332u_{11}$$

$$\mathbf{u}_1 = \begin{Bmatrix} u_{11} \\ u_{21} \end{Bmatrix} = \begin{Bmatrix} 1.0 \\ -0.5332 \end{Bmatrix}$$

and

$$\begin{bmatrix} 12 - \lambda_2 & -6 \\ -6 & 4 - \lambda_2 \end{bmatrix} \begin{Bmatrix} u_{12} \\ u_{22} \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix} \quad \text{or} \quad (12 - \lambda_2)u_{12} - 6u_{22} = 0$$

$$\text{or} \quad u_{22} = 1.8685u_{12}$$

that is,

$$\mathbf{u}_2 = \begin{Bmatrix} u_{12} \\ u_{22} \end{Bmatrix} = \begin{Bmatrix} 1.0 \\ 1.8685 \end{Bmatrix}$$

Thus the transformation that reduces $[A]$ to a diagonal form is given by

$$\mathbf{X} = [R]\mathbf{Y} = [\mathbf{u}_1 \quad \mathbf{u}_2]\mathbf{Y} = \begin{bmatrix} 1 & 1 \\ -0.5352 & 1.8685 \end{bmatrix} \begin{Bmatrix} y_1 \\ y_2 \end{Bmatrix} \qquad (E_5)$$

that is,

$$x_1 = y_1 + y_2$$
$$x_2 = -0.5352y_1 + 1.8685y_2$$

This yields the new quadratic term as $\frac{1}{2}\mathbf{Y}^T[\tilde{A}]\mathbf{Y}$, where

$$[\tilde{A}] = [R]^T[A][R] = \begin{bmatrix} 19.5682 & 0.0 \\ 0.0 & 3.5432 \end{bmatrix}$$

and hence the quadratic function becomes

$$\begin{aligned} f(y_1, y_2) &= \mathbf{B}^T[R]\mathbf{Y} + \tfrac{1}{2}\mathbf{Y}^T[\tilde{A}]\mathbf{Y} \\ &= 0.0704y_1 - 4.7370y_2 + \tfrac{1}{2}(19.5682)y_1^2 + \tfrac{1}{2}(3.5432)y_2^2 \end{aligned} \quad (E_6)$$

Stage 2: Reducing $[\tilde{A}]$ to a Unit Matrix
The transformation is given by $\mathbf{Y} = [S]\mathbf{Z}$, where

$$[S] = \begin{bmatrix} \frac{1}{\sqrt{19.5682}} & 0 \\ 0 & \frac{1}{\sqrt{3.5432}} \end{bmatrix} = \begin{bmatrix} 0.2262 & 0.0 \\ 0.0 & 0.5313 \end{bmatrix}$$

Stage 3: Complete Transformation
The total transformation is given by

$$\mathbf{X} = [R]\mathbf{Y} = [R][S]\mathbf{Z} = [T]\mathbf{Z} \quad (E_7)$$

where

$$\begin{aligned} [T] = [R][S] &= \begin{bmatrix} 1 & 1 \\ -0.5352 & 1.8685 \end{bmatrix} \begin{bmatrix} 0.2262 & 0 \\ 0 & 0.5313 \end{bmatrix} \\ &= \begin{bmatrix} 0.2262 & 0.5313 \\ -0.1211 & 0.9927 \end{bmatrix} \end{aligned} \quad (E_8)$$

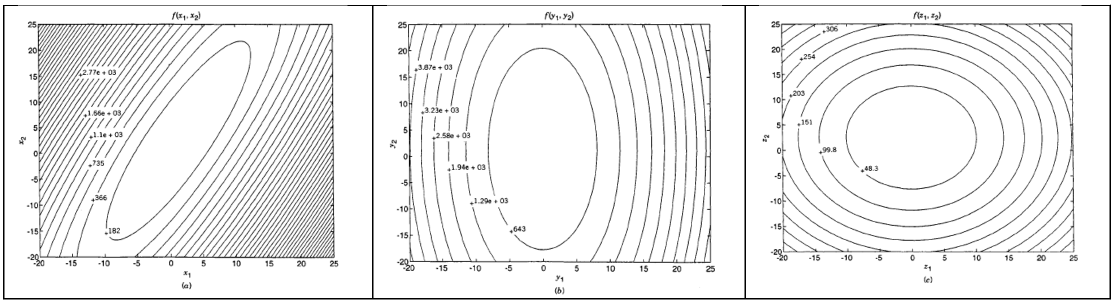
or

$$x_1 = 0.2262z_1 + 0.5313z_2$$
$$x_2 = -0.1211z_1 + 0.9927z_2$$

With this transformation, the quadratic function of Eq. (E_1) becomes

$$\begin{aligned} f(z_1, z_2) &= \mathbf{B}^T[T]\mathbf{Z} + \tfrac{1}{2}\mathbf{Z}^T[T]^T[A][T]\mathbf{Z} \\ &= 0.0160z_1 - 2.5167z_2 + \tfrac{1}{2}z_1^2 + \tfrac{1}{2}z_2^2 \qquad (E_9) \end{aligned}$$

The contours of the quadratic functions given by Eqs. (E_1), (E_6), and (E_9) are shown in Fig. 6.4*a*, *b*, and *c*, respectively.



Part 2

Indirect search (descent) methods

Gradient of a function

Define

The gradient of a function is an n-component vector given by

$$\underset{n\times 1}{\nabla f} = \left[\frac{\partial f}{\partial x_1}\ \frac{\partial f}{\partial x_2}\ \frac{\partial f}{\partial x_3} \cdots \frac{\partial f}{\partial x_n}\right]^T$$

The gradient has a very important property. If we move along the gradient direction from any point in n-dimensional space, the function value increases at the fastest rate. Hence the gradient direction is called the *direction of steepest ascent*. Unfortunately, the direction of steepest ascent is a local property and not a global one. This is illustrated in Fig. 6.15, where the gradient vectors ∇f evaluated at points 1, 2, 3, and 4 lie along the directions 11′, 22′, 33′, and 44′, respectively. Thus the function value increases at the fastest rate in the direction 11′ at point 1, but not at point 2. Similarly, the function value increases at the fastest rate in direction 22′ (33′) at point 2(3), but not at point 3 (4). In other words, the direction of steepest ascent generally varies from point to point, and if we make infinitely small moves along the direction of steepest ascent, the path will be a curved line like the curve 1–2–3–4 in Fig. 6.15.

Since the gradient vector represents the direction of steepest ascent, the negative of the gradient vector denotes the direction of steepest descent. Thus any method that makes use of the gradient vector can be expected to give the minimum point faster than one that does not make use of the gradient vector. All the descent methods make use of the gradient vector, either directly or

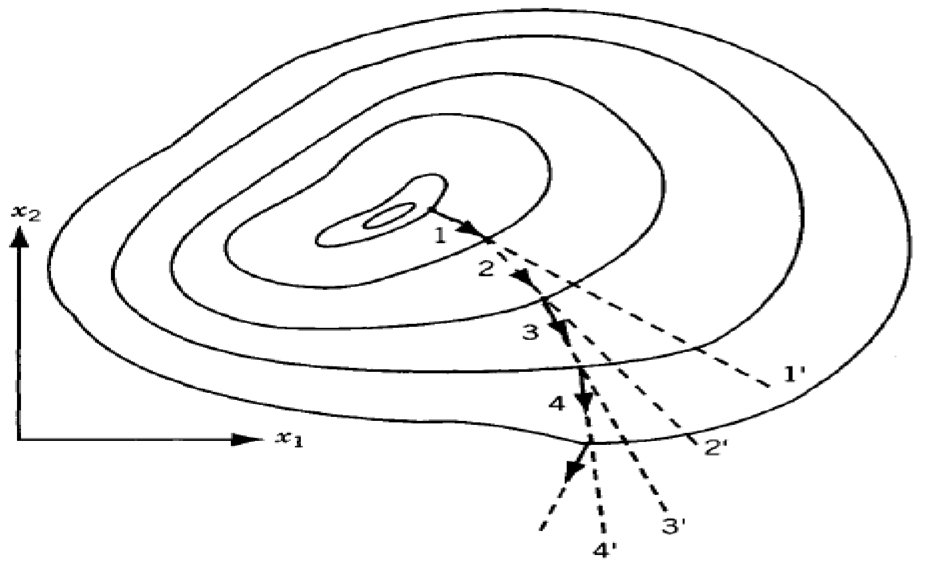


Figure 6.15 Steepest ascent directions.

indirectly, in finding the search directions. Before considering the descent methods of minimization, we prove that the gradient vector represents the direction of steepest ascent.

[Q1] Prove that the gradient vector represents the direction of steepest ascent.

Theorem 6.3 The gradient vector represents the direction of steepest ascent.

Proof: Consider an arbitrary point $\mathbf{X}$ in the n-dimensional space. Let f denote the value of the objective function at the point $\mathbf{X}$. Consider a neighboring point $\mathbf{X} + d\mathbf{X}$ with

$$d\mathbf{X} = \begin{Bmatrix} dx_1 \\ dx_2 \\ \vdots \\ dx_n \end{Bmatrix} \tag{6.57}$$

where $dx_1, dx_2, \ldots, dx_n$ represent the components of the vector $d\mathbf{X}$. The magnitude of the vector $d\mathbf{X}$, ds, is given by

$$d\mathbf{X}^T d\mathbf{X} = (ds)^2 = \sum_{i=1}^{n} (dx_i)^2 \tag{6.58}$$

If $f + df$ denotes the value of the objective function at $\mathbf{X} + d\mathbf{X}$, the change in f, df, associated with $d\mathbf{X}$ can be expressed as

$$df = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} dx_i = \nabla f^T d\mathbf{X} \tag{6.59}$$

If $\mathbf{u}$ denotes the unit vector along the direction $d\mathbf{X}$ and ds the length of $d\mathbf{X}$, we can write

$$d\mathbf{X} = \mathbf{u}\, ds \tag{6.60}$$

The rate of change of the function with respect to the step length ds is given by Eq. (6.59) as

$$\frac{df}{ds} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}\frac{dx_i}{ds} = \nabla f^T \frac{d\mathbf{X}}{ds} = \nabla f^T \mathbf{u} \tag{6.61}$$

The value of df/ds will be different for different directions and we are interested in finding the particular step $d\mathbf{X}$ along which the value of df/ds will be maximum. This will give the direction of steepest ascent.† By using the definition of the dot product, Eq. (6.61) can be rewritten as

$$\frac{df}{ds} = \|\nabla f\| \; \|\mathbf{u}\| \cos \theta \tag{6.62}$$

†In general, if $df/ds = \nabla f^T \mathbf{u} > 0$ along a vector $d\mathbf{X}$, it is called a direction of *ascent*, and if $df/ds < 0$, it is called a direction of *descent*.

where $\|\nabla f\|$ and $\|\mathbf{u}\|$ denote the lengths of the vectors ∇f and $\mathbf{u}$, respectively, and θ indicates the angle between the vectors ∇f and $\mathbf{u}$. It can be seen that df/ds will be maximum when $\theta = 0°$ and minimum when $\theta = 180°$. This indicates that the function value increases at a maximum rate in the direction of the gradient (i.e., when $\mathbf{u}$ is along ∇f).

[Q2] Prove that the maximum rate of change of/at any point X is equal to the magnitude of the gradient vector at the same point. Then show what we can do if the Evaluation of the Gradient poses certain problem

Theorem 6.4 The maximum rate of change of f at any point $\mathbf{X}$ is equal to the magnitude of the gradient vector at the same point.

Proof: The rate of change of the function f with respect to the step length s along a direction $\mathbf{u}$ is given by Eq. (6.62). Since df/ds is maximum when $\theta = 0°$ and $\mathbf{u}$ is a unit vector, Eq. (6.62) gives

$$\left.\left(\frac{df}{ds}\right)\right|_{\max} = \|\nabla f\|$$

which proves the theorem.

6.10.1 Evaluation of the Gradient

[Q3]"The evaluation of the gradient poses certain problems". Discuss this sentence.

The evaluation of the gradient requires the computation of the partial derivatives $\partial f/\partial x_i$, i = 1,2,. . .,ft. There are three situations where the evaluation of the gradient poses certain problems:

1. The function is differentiable at all the points, but the calculation of the components of the gradient, $\partial f/\partial x_i$, is either impractical or impossible.
2. The expressions for the partial derivatives $\partial f/\partial x_i$ can be derived, but they require large computational time for evaluation.
3. The gradient ∇f is not defined at all the points.

In the first case, we can use the forward finite-difference formula

$$\left.\frac{\partial f}{\partial x_i}\right|_{\mathbf{X}_m} \simeq \frac{f(\mathbf{X}_m + \Delta x_i \mathbf{u}_i) - f(\mathbf{X}_m)}{\Delta x_i}, \qquad i = 1,2,\ldots,n \tag{6.63}$$

6.11 steepest descent (Cauchy) method

The use of the negative of the gradient vector as a direction for minimization was first made by Cauchy in 1847 [6.12]. In this method we start from an initial trial point X_1 and iteratively move along the steepest descent directions until the optimum point is found. The steepest descent method can be summarized by the following steps:

[Q4](a)Summarize the steps of steepest descent method for Multivariable Unconstrained Minimization problem.

1. Start with an arbitrary initial point $\mathbf{X}_1$. Set the iteration number as i = 1.
2. Find the search direction $\mathbf{S}_i$ as

$$\mathbf{S}_i = -\nabla f_i = -\nabla f(\mathbf{X}_i) \tag{6.69}$$

3. Determine the optimal step length λ_i^* in the direction $\mathbf{S}_i$ and set

$$\mathbf{X}_{i+1} = \mathbf{X}_i + \lambda_i^* \mathbf{S}_i = \mathbf{X}_i - \lambda_i^* \nabla f_i \tag{6.70}$$

4. Test the new point, $\mathbf{X}_{i+1}$, for optimality. If $\mathbf{X}_{i+1}$ is optimum, stop the process. Otherwise, go to step 5.
5. Set the new iteration number $i = i + 1$ and go to step 2.

The method of steepest descent may appear to be the *best unconstrained minimization* technique since each one-dimensional search starts in the "best" direction. However, owing to the fact that the steepest descent direction is a local property, the method is not really effective in most problems.

[Q4](b)Use steepest descent method to Minimize the following Multivariable Unconstrained Minimization problem starting from X= {0 0)^T

$$f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1 x_2 + x_2^2$$

SOLUTION

Iteration 1

The gradient of f is given by

$$\nabla f = \begin{Bmatrix} \partial f/\partial x_1 \\ \partial f/\partial x_2 \end{Bmatrix} = \begin{Bmatrix} 1 + 4x_1 + 2x_2 \\ -1 + 2x_1 + 2x_2 \end{Bmatrix}$$

$$\nabla f_1 = \nabla f(\mathbf{X}_1) = \begin{Bmatrix} 1 \\ -1 \end{Bmatrix}$$

Therefore,

$$\mathbf{S}_1 = -\nabla f_1 = \begin{Bmatrix} -1 \\ 1 \end{Bmatrix}$$

To find $\mathbf{X}_2$, we need to find the optimal step length λ_1^*. For this, we minimize $f(\mathbf{X}_1 + \lambda_1\mathbf{S}_1) = f(-\lambda_1, \lambda_1) = \lambda_1^2 - 2\lambda_1$ with respect to λ_1. Since $df/d\lambda_1 = 0$ at $\lambda_1^* = 1$, we obtain

$$\mathbf{X}_2 = \mathbf{X}_1 + \lambda_1^*\mathbf{S}_1 = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix} + 1\begin{Bmatrix} -1 \\ 1 \end{Bmatrix} = \begin{Bmatrix} -1 \\ 1 \end{Bmatrix}$$

As $\nabla f_2 = \nabla f(\mathbf{X}_2) = \begin{Bmatrix} -1 \\ -1 \end{Bmatrix} \neq \begin{Bmatrix} 0 \\ 0 \end{Bmatrix}$, $\mathbf{X}_2$ is not optimum.

Iteration 2

$$\mathbf{S}_2 = -\nabla f_2 = \begin{Bmatrix} 1 \\ 1 \end{Bmatrix}$$

To minimize

$$f(\mathbf{X}_2 + \lambda_2\mathbf{S}_2) = f(-1 + \lambda_2, 1 + \lambda_2)$$
$$= 5\lambda_2^2 - 2\lambda_2 - 1$$

we set $df/d\lambda_2 = 0$. This gives $\lambda_2^* = \frac{1}{5}$, and hence

$$\mathbf{X}_3 = \mathbf{X}_2 + \lambda_2^*\mathbf{S}_2 = \begin{Bmatrix} -1 \\ 1 \end{Bmatrix} + \frac{1}{5}\begin{Bmatrix} 1 \\ 1 \end{Bmatrix} = \begin{Bmatrix} -0.8 \\ 1.2 \end{Bmatrix}$$

Since the components of the gradient at $\mathbf{X}_3$, $\nabla f_3 = \begin{Bmatrix} 0.2 \\ -0.2 \end{Bmatrix}$, are not zero, we proceed to the next iteration.

Iteration 3

$$\mathbf{S}_3 = -\nabla f_3 = \begin{Bmatrix} -0.2 \\ 0.2 \end{Bmatrix}$$

As

$$f(\mathbf{X}_3 + \lambda_3\mathbf{S}_3) = f(-0.8 - 0.2\lambda_3,\ 1.2 + 0.2\lambda_3)$$

$$= 0.04\lambda_3^2 - 0.08\lambda_3 - 1.20, \qquad \frac{df}{d\lambda_3} = 0 \text{ at } \lambda_3^* = 1.0$$

Therefore,

$$\mathbf{X}_4 = \mathbf{X}_3 + \lambda_3^*\mathbf{S}_3 = \begin{Bmatrix} -0.8 \\ 1.2 \end{Bmatrix} + 1.0\begin{Bmatrix} -0.2 \\ 0.2 \end{Bmatrix} = \begin{Bmatrix} -1.0 \\ 1.4 \end{Bmatrix}$$

The gradient at $\mathbf{X}_4$ is given by

$$\nabla f_4 = \begin{Bmatrix} -0.20 \\ -0.20 \end{Bmatrix}$$

Since $\nabla f_4 \neq \begin{Bmatrix} 0 \\ 0 \end{Bmatrix}$, $\mathbf{X}_4$ is not optimum and hence we have to proceed to the next iteration. This process has to be continued until the optimum point, $\mathbf{X}^* = \begin{Bmatrix} -1.0 \\ 1.5 \end{Bmatrix}$, is found.

Convergence Criteria. The following criteria can be used to terminate the iterative process:

1. When the change in function value in two consecutive iterations is small:

$$\left| \frac{f(\mathbf{X}_{i+1}) - f(\mathbf{X}_i)}{f(\mathbf{X}_i)} \right| \le \varepsilon_1 \tag{6.71}$$

6.12 Conjugate Gradient (Fletcher-Reeves) Method

The convergence characteristics of the steepest descent method can be improved greatly by modifying it into a conjugate gradient method (which can be considered as a conjugate directions method involving the use of the gradient of the function). We saw (in Section 6.7) that any minimization method that makes use of the conjugate directions is quadratically convergent. This property of quadratic convergence is very useful because it ensures that the method will minimize a quadratic function in *n* steps or less. Since any general function can be approximated reasonably well by a quadratic near the optimum point, any quadratically convergent method is expected to find the optimum point in a finite number of iterations.

We have seen that Powell's conjugate direction method requires *n* single variable minimizations per iteration and sets up a new conjugate direction at the end of each iteration. Thus it requires, in general, n^2 single-variable minimizations to find the minimum of a quadratic function. On the other hand, if we can evaluate the gradients of the objective function, we can set up a new conjugate direction after every one-dimensional minimization, and hence we can achieve faster convergence. The construction of conjugate directions and development of the Fletcher-Reeves method are discussed in this section.

6.12.1 Development of the Fletcher-Reeves Method

[Q5] Develop the Fletcher-Reeves Method

Consider the development of an algorithm by modifying the steepest descent method applied to a quadratic function $f(\mathbf{X}) = \frac{1}{2}\mathbf{X}^T\mathbf{A}\mathbf{X} + \mathbf{B}^T\mathbf{X} + C$ by imposing the condition that the successive directions be mutually conjugate. Let $\mathbf{X}_1$ be the starting point for the minimization and let the first search direction be the steepest descent direction:

$$\mathbf{S}_1 = -\nabla f_1 = -\mathbf{A}\mathbf{X}_1 - \mathbf{B} \tag{6.74}$$

$$\mathbf{X}_2 = \mathbf{X}_1 + \lambda_1^* \mathbf{S}_1 \tag{6.75}$$

or

$$\mathbf{S}_1 = \frac{\mathbf{X}_2 - \mathbf{X}_1}{\lambda_1^*} \tag{6.76}$$

where λ_1^* is the minimizing step length in the direction $\mathbf{S}_1$, so that

$$\mathbf{S}_1^T \nabla f|_{\mathbf{X}_2} = 0 \tag{6.77}$$

Equation (6.77) can be expanded as

$$\mathbf{S}_1^T[\mathbf{A}(\mathbf{X}_1 + \lambda_1^*\mathbf{S}_1) + \mathbf{B}] = 0 \tag{6.78}$$

from which the value of λ_1^* can be found as

$$\lambda_1^* = \frac{-\mathbf{S}_1^T(\mathbf{A}\mathbf{X}_1 + \mathbf{B})}{\mathbf{S}_1^T \mathbf{A}\mathbf{S}_1} = -\frac{\mathbf{S}_1^T \nabla f_1}{\mathbf{S}_1^T \mathbf{A}\mathbf{S}_1} \tag{6.79}$$

Now express the second search direction as a linear combination of $\mathbf{S}_1$ and $-\nabla f_2$:

$$\mathbf{S}_2 = -\nabla f_2 + \beta_2\mathbf{S}_1 \tag{6.80}$$

where β_2 is to be chosen so as to make $\mathbf{S}_1$ and $\mathbf{S}_2$ conjugate. This requires that

$$\mathbf{S}_1^T \mathbf{A}\mathbf{S}_2 = 0 \tag{6.81}$$

Substituting Eq. (6.80) into Eq. (6.81) leads to

$$\mathbf{S}_1^T\mathbf{A}\,(-\nabla f_2 + \beta_2\mathbf{S}_1) = 0 \tag{6.82}$$

Equations (6.76) and (6.82) yield

$$-\frac{(\mathbf{X}_2 - \mathbf{X}_1)^T}{\lambda_1^*}\mathbf{A}\,(\nabla f_2 - \beta_2\mathbf{S}_1) = 0 \tag{6.83}$$

The difference of the gradients $(\nabla f_2 - \nabla f_1)$ can be expressed as

$$(\nabla f_2 - \nabla f_1) = (\mathbf{A}\mathbf{X}_2 + \mathbf{B}) - (\mathbf{A}\mathbf{X}_1 + \mathbf{B}) = \mathbf{A}(\mathbf{X}_2 - \mathbf{X}_1) \tag{6.84}$$

With the help of Eq. (6.84), Eq. (6.83) can be written as

$$(\nabla f_2 - \nabla f_1)^T(\nabla f_2 - \beta_2\mathbf{S}_1) = 0 \tag{6.85}$$

where the symmetricity of the matrix $\mathbf{A}$ has been used. Equation (6.85) can be expanded as

$$\nabla f_2^T \nabla f_2 - \nabla f_1^T \nabla f_2 - \beta_2 \nabla f_2^T \mathbf{S}_1 + \beta_2 \nabla f_1^T \mathbf{S}_1 = 0 \tag{6.86}$$

Since $\nabla f_1^T \nabla f_2 = -\mathbf{S}_1^T \nabla f_2 = 0$ from Eq. (6.77), Eq. (6.86) gives

$$\beta_2 = -\frac{\nabla f_2^T \nabla f_2}{\nabla f_1^T \mathbf{S}_1} = \frac{\nabla f_2^T \nabla f_2}{\nabla f_1^T \nabla f_1} \tag{6.87}$$

Next we consider the third search direction as a linear combination of $\mathbf{S}_1$, $\mathbf{S}_2$, and $-\nabla f_3$ as

$$\mathbf{S}_3 = -\nabla f_3 + \beta_3 \mathbf{S}_2 + \delta_3 \mathbf{S}_1 \tag{6.88}$$

where the values of β_3 and δ_3 can be found by making $\mathbf{S}_3$ conjugate to $\mathbf{S}_1$ and $\mathbf{S}_2$. By using the condition $\mathbf{S}_1^T\mathbf{A}\mathbf{S}_3 = 0$, the value of δ_3 can be found to be zero (see Problem 6.40). When the condition $\mathbf{S}_2^T\mathbf{A}\mathbf{S}_3 = 0$ is used, the value of β_3 can be obtained as (see Problem 6.41)

$$\beta_3 = \frac{\nabla f_3^T \nabla f_3}{\nabla f_2^T \nabla f_2} \tag{6.89}$$

so that Eq. (6.88) becomes

$$\mathbf{S}_3 = -\nabla f_3 + \beta_3 \mathbf{S}_2 \tag{6.90}$$

where β_3 is given by Eq. (6.89). In fact, Eq. (6.90) can be generalized as

$$\mathbf{S}_i = -\nabla f_i + \beta_i \mathbf{S}_{i-1} \tag{6.91}$$

where

$$\beta_i = \frac{\nabla f_i^T \nabla f_i}{\nabla f_{i-1}^T \nabla f_{i-1}} \tag{6.92}$$

Equations (6.91) and (6.92) define the search directions used in the Fletcher–Reeves method [6.13].

Part 2

Indirect search (descent) methods

Gradient of a function

Define

The gradient of a function is an n-component vector given by

$$\underset{n\times 1}{\nabla f} = \left[\frac{\partial f}{\partial x_1}\ \frac{\partial f}{\partial x_2}\ \frac{\partial f}{\partial x_3} \cdots \frac{\partial f}{\partial x_n}\right]^T$$

The gradient has a very important property. If we move along the gradient direction from any point in n-dimensional space, the function value increases at the fastest rate. Hence the gradient direction is called the *direction of steepest ascent*. Unfortunately, the direction of steepest ascent is a local property and not a global one. This is illustrated in Fig. 6.15, where the gradient vectors ∇f evaluated at points 1, 2, 3, and 4 lie along the directions 11′, 22′, 33′, and 44′, respectively. Thus the function value increases at the fastest rate in the direction 11′ at point 1, but not at point 2. Similarly, the function value increases at the fastest rate in direction 22′ (33′) at point 2(3), but not at point 3 (4). In other words, the direction of steepest ascent generally varies from point to point, and if we make infinitely small moves along the direction of steepest ascent, the path will be a curved line like the curve 1–2–3–4 in Fig. 6.15.

Since the gradient vector represents the direction of steepest ascent, the negative of the gradient vector denotes the direction of steepest descent. Thus any method that makes use of the gradient vector can be expected to give the minimum point faster than one that does not make use of the gradient vector. All the descent methods make use of the gradient vector, either directly or

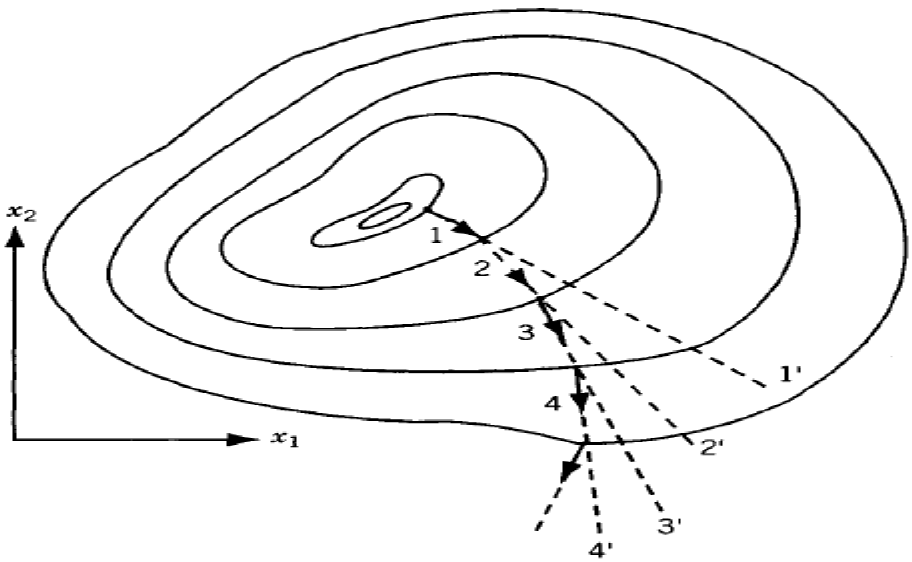


Figure 6.15 Steepest ascent directions.

indirectly, in finding the search directions. Before considering the descent methods of minimization, we prove that the gradient vector represents the direction of steepest ascent.

[Q1] Prove that the gradient vector represents the direction of steepest ascent.

Theorem 6.3 The gradient vector represents the direction of steepest ascent.

Proof: Consider an arbitrary point **X** in the n-dimensional space. Let f denote the value of the objective function at the point **X**. Consider a neighboring point $\mathbf{X} + d\mathbf{X}$ with

$$d\mathbf{X} = \begin{Bmatrix} dx_1 \\ dx_2 \\ \vdots \\ dx_n \end{Bmatrix} \tag{6.57}$$

where $dx_1, dx_2, \ldots, dx_n$ represent the components of the vector $d\mathbf{X}$. The magnitude of the vector $d\mathbf{X}$, ds, is given by

$$d\mathbf{X}^T d\mathbf{X} = (ds)^2 = \sum_{i=1}^{n} (dx_i)^2 \tag{6.58}$$

If $f + df$ denotes the value of the objective function at $\mathbf{X} + d\mathbf{X}$, the change in f, df, associated with $d\mathbf{X}$ can be expressed as

$$df = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} dx_i = \nabla f^T d\mathbf{X} \tag{6.59}$$

If **u** denotes the unit vector along the direction $d\mathbf{X}$ and ds the length of $d\mathbf{X}$, we can write

$$d\mathbf{X} = \mathbf{u}\, ds \tag{6.60}$$

The rate of change of the function with respect to the step length ds is given by Eq. (6.59) as

$$\frac{df}{ds} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \frac{dx_i}{ds} = \nabla f^T \frac{d\mathbf{X}}{ds} = \nabla f^T \mathbf{u} \tag{6.61}$$

The value of df/ds will be different for different directions and we are interested in finding the particular step $d\mathbf{X}$ along which the value of df/ds will be maximum. This will give the direction of steepest ascent.† By using the definition of the dot product, Eq. (6.61) can be rewritten as

$$\frac{df}{ds} = \|\nabla f\| \; \|\mathbf{u}\| \cos \theta \tag{6.62}$$

†In general, if $df/ds = \nabla f^T \mathbf{u} > 0$ along a vector $d\mathbf{X}$, it is called a direction of *ascent*, and if $df/ds < 0$, it is called a direction of *descent*.

where $\|\nabla f\|$ and $\|\mathbf{u}\|$ denote the lengths of the vectors ∇f and $\mathbf{u}$, respectively, and θ indicates the angle between the vectors ∇f and $\mathbf{u}$. It can be seen that df/ds will be maximum when $\theta = 0°$ and minimum when $\theta = 180°$. This indicates that the function value increases at a maximum rate in the direction of the gradient (i.e., when $\mathbf{u}$ is along ∇f).

[Q2] Prove that the maximum rate of change of/at any point X is equal to the magnitude of the gradient vector at the same point. Then show what we can do if the Evaluation of the Gradient poses certain problem

Theorem 6.4 The maximum rate of change of f at any point $\mathbf{X}$ is equal to the magnitude of the gradient vector at the same point.

Proof: The rate of change of the function f with respect to the step length s along a direction $\mathbf{u}$ is given by Eq. (6.62). Since df/ds is maximum when $\theta = 0°$ and $\mathbf{u}$ is a unit vector, Eq. (6.62) gives

$$\left.\left(\frac{df}{ds}\right)\right|_{\max} = \|\nabla f\|$$

which proves the theorem.

6.10.1 Evaluation of the Gradient

[Q3]"The evaluation of the gradient poses certain problems". Discuss this sentence.

The evaluation of the gradient requires the computation of the partial derivatives $\partial f/\partial x_i$, i = 1,2,. . .,ft. There are three situations where the evaluation of the gradient poses certain problems:

1. The function is differentiable at all the points, but the calculation of the components of the gradient, $\partial f/\partial x_i$, is either impractical or impossible.
2. The expressions for the partial derivatives $\partial f/\partial x_i$ can be derived, but they require large computational time for evaluation.
3. The gradient ∇f is not defined at all the points.

In the first case, we can use the forward finite-difference formula

$$\left.\frac{\partial f}{\partial x_i}\right|_{\mathbf{X}_m} \simeq \frac{f(\mathbf{X}_m + \Delta x_i \mathbf{u}_i) - f(\mathbf{X}_m)}{\Delta x_i}, \quad i = 1,2,\ldots,n \tag{6.63}$$

6.11 steepest descent (Cauchy) method

The use of the negative of the gradient vector as a direction for minimization was first made by Cauchy in 1847 [6.12]. In this method we start from an initial trial point X_1 and iteratively move along the steepest descent directions until the optimum point is found. The steepest descent method can be summarized by the following steps:

[Q4](a)Summarize the steps of steepest descent method for Multivariable Unconstrained Minimization problem.

1. Start with an arbitrary initial point $\mathbf{X}_1$. Set the iteration number as i = 1.
2. Find the search direction $\mathbf{S}_i$ as

$$\mathbf{S}_i = -\nabla f_i = -\nabla f(\mathbf{X}_i) \tag{6.69}$$

3. Determine the optimal step length λ_i^* in the direction $\mathbf{S}_i$ and set

$$\mathbf{X}_{i+1} = \mathbf{X}_i + \lambda_i^* \mathbf{S}_i = \mathbf{X}_i - \lambda_i^* \nabla f_i \tag{6.70}$$

4. Test the new point, $\mathbf{X}_{i+1}$, for optimality. If $\mathbf{X}_{i+1}$ is optimum, stop the process. Otherwise, go to step 5.
5. Set the new iteration number $i = i + 1$ and go to step 2.

The method of steepest descent may appear to be the *best unconstrained minimization* technique since each one-dimensional search starts in the "best" direction. However, owing to the fact that the steepest descent direction is a local property, the method is not really effective in most problems.

[Q4](b)Use steepest descent method to Minimize the following Multivariable Unconstrained Minimization problem starting from X= {0 0)^T

$$f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1 x_2 + x_2^2$$

SOLUTION

Iteration 1

The gradient of f is given by

$$\nabla f = \begin{Bmatrix} \partial f/\partial x_1 \\ \partial f/\partial x_2 \end{Bmatrix} = \begin{Bmatrix} 1 + 4x_1 + 2x_2 \\ -1 + 2x_1 + 2x_2 \end{Bmatrix}$$

$$\nabla f_1 = \nabla f(\mathbf{X}_1) = \begin{Bmatrix} 1 \\ -1 \end{Bmatrix}$$

Therefore,

$$\mathbf{S}_1 = -\nabla f_1 = \begin{Bmatrix} -1 \\ 1 \end{Bmatrix}$$

To find $\mathbf{X}_2$, we need to find the optimal step length λ_1^*. For this, we minimize $f(\mathbf{X}_1 + \lambda_1 \mathbf{S}_1) = f(-\lambda_1, \lambda_1) = \lambda_1^2 - 2\lambda_1$ with respect to λ_1. Since $df/d\lambda_1 = 0$ at $\lambda_1^* = 1$, we obtain

$$\mathbf{X}_2 = \mathbf{X}_1 + \lambda_1^* \mathbf{S}_1 = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix} + 1 \begin{Bmatrix} -1 \\ 1 \end{Bmatrix} = \begin{Bmatrix} -1 \\ 1 \end{Bmatrix}$$

As $\nabla f_2 = \nabla f(\mathbf{X}_2) = \begin{Bmatrix} -1 \\ -1 \end{Bmatrix} \neq \begin{Bmatrix} 0 \\ 0 \end{Bmatrix}$, $\mathbf{X}_2$ is not optimum.

Iteration 2

$$\mathbf{S}_2 = -\nabla f_2 = \begin{Bmatrix} 1 \\ 1 \end{Bmatrix}$$

To minimize

$$f(\mathbf{X}_2 + \lambda_2 \mathbf{S}_2) = f(-1 + \lambda_2, 1 + \lambda_2)$$
$$= 5\lambda_2^2 - 2\lambda_2 - 1$$

we set $df/d\lambda_2 = 0$. This gives $\lambda_2^* = \frac{1}{5}$, and hence

$$\mathbf{X}_3 = \mathbf{X}_2 + \lambda_2^* \mathbf{S}_2 = \begin{Bmatrix} -1 \\ 1 \end{Bmatrix} + \frac{1}{5} \begin{Bmatrix} 1 \\ 1 \end{Bmatrix} = \begin{Bmatrix} -0.8 \\ 1.2 \end{Bmatrix}$$

Since the components of the gradient at $\mathbf{X}_3$, $\nabla f_3 = \begin{Bmatrix} 0.2 \\ -0.2 \end{Bmatrix}$, are not zero, we proceed to the next iteration.

Iteration 3

$$\mathbf{S}_3 = -\nabla f_3 = \begin{Bmatrix} -0.2 \\ 0.2 \end{Bmatrix}$$

As

$$f(\mathbf{X}_3 + \lambda_3 \mathbf{S}_3) = f(-0.8 - 0.2\lambda_3,\ 1.2 + 0.2\lambda_3)$$

$$= 0.04\lambda_3^2 - 0.08\lambda_3 - 1.20, \qquad \frac{df}{d\lambda_3} = 0 \text{ at } \lambda_3^* = 1.0$$

Therefore,

$$\mathbf{X}_4 = \mathbf{X}_3 + \lambda_3^* \mathbf{S}_3 = \begin{Bmatrix} -0.8 \\ 1.2 \end{Bmatrix} + 1.0 \begin{Bmatrix} -0.2 \\ 0.2 \end{Bmatrix} = \begin{Bmatrix} -1.0 \\ 1.4 \end{Bmatrix}$$

The gradient at $\mathbf{X}_4$ is given by

$$\nabla f_4 = \begin{Bmatrix} -0.20 \\ -0.20 \end{Bmatrix}$$

Since $\nabla f_4 \neq \begin{Bmatrix} 0 \\ 0 \end{Bmatrix}$, $\mathbf{X}_4$ is not optimum and hence we have to proceed to the next iteration. This process has to be continued until the optimum point, $\mathbf{X}^* = \begin{Bmatrix} -1.0 \\ 1.5 \end{Bmatrix}$, is found.

Convergence Criteria. The following criteria can be used to terminate the iterative process:

1. When the change in function value in two consecutive iterations is small:

$$\left|\frac{f(\mathbf{X}_{i+1}) - f(\mathbf{X}_i)}{f(\mathbf{X}_i)}\right| \le \varepsilon_1 \tag{6.71}$$

6.12 Conjugate Gradient (Fletcher-Reeves) Method

The convergence characteristics of the steepest descent method can be improved greatly by modifying it into a conjugate gradient method (which can be considered as a conjugate directions method involving the use of the gradient of the function). We saw (in Section 6.7) that any minimization method that makes use of the conjugate directions is quadratically convergent. This property of quadratic convergence is very useful because it ensures that the method will minimize a quadratic function in *n* steps or less. Since any general function can be approximated reasonably well by a quadratic near the optimum point, any quadratically convergent method is expected to find the optimum point in a finite number of iterations.

We have seen that Powell's conjugate direction method requires *n* single variable minimizations per iteration and sets up a new conjugate direction at the end of each iteration. Thus it requires, in general, n^2 single-variable minimizations to find the minimum of a quadratic function. On the other hand, if we can evaluate the gradients of the objective function, we can set up a new conjugate direction after every one-dimensional minimization, and hence we can achieve faster convergence. The construction of conjugate directions and development of the Fletcher-Reeves method are discussed in this section.

6.12.1 Development of the Fletcher-Reeves Method

[Q5] Develop the Fletcher-Reeves Method

Consider the development of an algorithm by modifying the steepest descent method applied to a quadratic function $f(\mathbf{X}) = \frac{1}{2}\mathbf{X}^T\mathbf{A}\mathbf{X} + \mathbf{B}^T\mathbf{X} + C$ by imposing the condition that the successive directions be mutually conjugate. Let $\mathbf{X}_1$ be the starting point for the minimization and let the first search direction be the steepest descent direction:

$$\mathbf{S}_1 = -\nabla f_1 = -\mathbf{A}\mathbf{X}_1 - \mathbf{B} \tag{6.74}$$

$$\mathbf{X}_2 = \mathbf{X}_1 + \lambda_1^*\mathbf{S}_1 \tag{6.75}$$

or

$$\mathbf{S}_1 = \frac{\mathbf{X}_2 - \mathbf{X}_1}{\lambda_1^*} \tag{6.76}$$

where λ_1^* is the minimizing step length in the direction $\mathbf{S}_1$, so that

$$\mathbf{S}_1^T \nabla f|_{\mathbf{X}_2} = 0 \tag{6.77}$$

Equation (6.77) can be expanded as

$$\mathbf{S}_1^T[\mathbf{A}(\mathbf{X}_1 + \lambda_1^*\mathbf{S}_1) + \mathbf{B}] = 0 \tag{6.78}$$

from which the value of λ_1^* can be found as

$$\lambda_1^* = \frac{-\mathbf{S}_1^T(\mathbf{A}\mathbf{X}_1 + \mathbf{B})}{\mathbf{S}_1^T \mathbf{A}\mathbf{S}_1} = -\frac{\mathbf{S}_1^T \nabla f_1}{\mathbf{S}_1^T \mathbf{A}\mathbf{S}_1} \tag{6.79}$$

Now express the second search direction as a linear combination of $\mathbf{S}_1$ and $-\nabla f_2$:

$$\mathbf{S}_2 = -\nabla f_2 + \beta_2\mathbf{S}_1 \tag{6.80}$$

where β_2 is to be chosen so as to make $\mathbf{S}_1$ and $\mathbf{S}_2$ conjugate. This requires that

$$\mathbf{S}_1^T \mathbf{A}\mathbf{S}_2 = 0 \tag{6.81}$$

Substituting Eq. (6.80) into Eq. (6.81) leads to

$$\mathbf{S}_1^T\mathbf{A}\,(-\nabla f_2 + \beta_2\mathbf{S}_1) = 0 \tag{6.82}$$

Equations (6.76) and (6.82) yield

$$-\frac{(\mathbf{X}_2 - \mathbf{X}_1)^T}{\lambda_1^*}\,\mathbf{A}\,(\nabla f_2 - \beta_2\mathbf{S}_1) = 0 \tag{6.83}$$

The difference of the gradients $(\nabla f_2 - \nabla f_1)$ can be expressed as

$$(\nabla f_2 - \nabla f_1) = (\mathbf{A}\mathbf{X}_2 + \mathbf{B}) - (\mathbf{A}\mathbf{X}_1 + \mathbf{B}) = \mathbf{A}(\mathbf{X}_2 - \mathbf{X}_1) \tag{6.84}$$

With the help of Eq. (6.84), Eq. (6.83) can be written as

$$(\nabla f_2 - \nabla f_1)^T(\nabla f_2 - \beta_2\mathbf{S}_1) = 0 \tag{6.85}$$

where the symmetricity of the matrix $\mathbf{A}$ has been used. Equation (6.85) can be expanded as

$$\nabla f_2^T \nabla f_2 - \nabla f_1^T \nabla f_2 - \beta_2 \nabla f_2^T \mathbf{S}_1 + \beta_2 \nabla f_1^T \mathbf{S}_1 = 0 \tag{6.86}$$

Since $\nabla f_1^T \nabla f_2 = -\mathbf{S}_1^T \nabla f_2 = 0$ from Eq. (6.77), Eq. (6.86) gives

$$\beta_2 = -\frac{\nabla f_2^T \nabla f_2}{\nabla f_1^T \mathbf{S}_1} = \frac{\nabla f_2^T \nabla f_2}{\nabla f_1^T \nabla f_1} \tag{6.87}$$

Next we consider the third search direction as a linear combination of $\mathbf{S}_1$, $\mathbf{S}_2$, and $-\nabla f_3$ as

$$\mathbf{S}_3 = -\nabla f_3 + \beta_3 \mathbf{S}_2 + \delta_3 \mathbf{S}_1 \tag{6.88}$$

where the values of β_3 and δ_3 can be found by making $\mathbf{S}_3$ conjugate to $\mathbf{S}_1$ and $\mathbf{S}_2$. By using the condition $\mathbf{S}_1^T \mathbf{A} \mathbf{S}_3 = 0$, the value of δ_3 can be found to be zero (see Problem 6.40). When the condition $\mathbf{S}_2^T \mathbf{A} \mathbf{S}_3 = 0$ is used, the value of β_3 can be obtained as (see Problem 6.41)

$$\beta_3 = \frac{\nabla f_3^T \, \nabla f_3}{\nabla f_2^T \, \nabla f_2} \tag{6.89}$$

so that Eq. (6.88) becomes

$$\mathbf{S}_3 = -\nabla f_3 + \beta_3 \mathbf{S}_2 \tag{6.90}$$

where β_3 is given by Eq. (6.89). In fact, Eq. (6.90) can be generalized as

$$\mathbf{S}_i = -\nabla f_i + \beta_i \mathbf{S}_{i-1} \tag{6.91}$$

where

$$\beta_i = \frac{\nabla f_i^T \, \nabla f_i}{\nabla f_{i-1}^T \, \nabla f_{i-1}} \tag{6.92}$$

Equations (6.91) and (6.92) define the search directions used in the Fletcher–Reeves method [6.13].

6.12.2 Fletcher-Reeves Method

[Q14](a)Summarize the steps of iterative procedure of Fletcher-Reeves method for Multivariable Unconstrained Minimization problem.

The iterative procedure of Fletcher-Reeves method can be stated as follows:

1. Start with an arbitrary initial point $\mathbf{X}_1$.
2. Set the first search direction $\mathbf{S}_1 = -\nabla f(\mathbf{X}_1) = -\nabla f_1$.
3. Find the point $\mathbf{X}_2$ according to the relation

$$\mathbf{X}_2 = \mathbf{X}_1 + \lambda_1^* \mathbf{S}_1$$

where λ_1^* is the optimal step length in the direction $\mathbf{S}_1$. Set $i = 2$ and go to the next step.

4. Find $\nabla f_i = \nabla f(\mathbf{X}_i)$, and set

$$\mathbf{S}_i = -\nabla f_i + \frac{|\nabla f_i|^2}{|\nabla f_{i-1}|^2} \mathbf{S}_{i-1} \qquad (6.93)$$

5. Compute the optimum step length λ_i^* in the direction $\mathbf{S}_i$, and find the new point

$$\mathbf{X}_{i+1} = \mathbf{X}_i + \lambda_i^* \mathbf{S}_i \qquad (6.94)$$

6. Test for the optimality of the point $\mathbf{X}_{i+1}$. If $\mathbf{X}_{i+1}$ is optimum, stop the process. Otherwise, set the value of $i = i + 1$ and go to step 4.

[Q14](b) Use **Fletcher-Reeves Method** to minimize the following multivariable Unconstrained Minimization problem starting from X= {0 0)^T

Minimize $f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1 x_2 + x_2^2$

SOLUTION

Iteration 1

$$\nabla f = \begin{Bmatrix} \partial f/\partial x_1 \\ \partial f/\partial x_2 \end{Bmatrix} = \begin{Bmatrix} 1 + 4x_1 + 2x_2 \\ -1 + 2x_1 + 2x_2 \end{Bmatrix}$$

$$\nabla f_1 = \nabla f(\mathbf{X}_1) = \begin{Bmatrix} 1 \\ -1 \end{Bmatrix}$$

The search direction is taken as $\mathbf{S}_1 = -\nabla f_1 = \begin{Bmatrix} -1 \\ 1 \end{Bmatrix}$. To find the optimal step length λ_1^* along $\mathbf{S}_1$, we minimize $f(\mathbf{X}_1 + \lambda_1 \mathbf{S}_1)$ with respect to λ_1. Here

$$f(\mathbf{X}_1 + \lambda_1 \mathbf{S}_1) = f(-\lambda_1, +\lambda_1) = \lambda_1^2 - 2\lambda_1$$

$$\frac{df}{d\lambda_1} = 0 \quad \text{at} \quad \lambda_1^* = 1$$

Therefore,

$$\mathbf{X}_2 = \mathbf{X}_1 + \lambda_1^* \mathbf{S}_1 = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix} + 1 \begin{Bmatrix} -1 \\ 1 \end{Bmatrix} = \begin{Bmatrix} -1 \\ 1 \end{Bmatrix}$$

$$\nabla f = \begin{Bmatrix} \partial f/\partial x_1 \\ \partial f/\partial x_2 \end{Bmatrix} = \begin{Bmatrix} 1 + 4x_1 + 2x_2 \\ -1 + 2x_1 + 2x_2 \end{Bmatrix}$$

Iteration 2

Since $\nabla f_2 = \nabla f(\mathbf{X}_2) = \begin{Bmatrix} -1 \\ -1 \end{Bmatrix}$, Eq. (6.93) gives the next search direction as

$$\mathbf{S}_2 = -\nabla f_2 + \frac{|\nabla f_2|^2}{|\nabla f_1|^2} \mathbf{S}_1$$

where

$$|\nabla f_1|^2 = 2 \quad \text{and} \quad |\nabla f_2|^2 = 2$$

Therefore,

$$\mathbf{S}_2 = -\begin{Bmatrix} -1 \\ -1 \end{Bmatrix} + \left(\frac{2}{2}\right) \begin{Bmatrix} -1 \\ 1 \end{Bmatrix} = \begin{Bmatrix} 0 \\ +2 \end{Bmatrix}$$

$$\mathbf{X}_2 = \begin{Bmatrix} -1 \\ 1 \end{Bmatrix} \qquad \mathbf{S}_2 = \begin{Bmatrix} 0 \\ +2 \end{Bmatrix}$$

$$f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$$

To find λ_2^*, we minimize

$$\begin{aligned} f(\mathbf{X}_2 + \lambda_2 \mathbf{S}_2) &= f(-1,\ 1 + 2\lambda_2) \\ &= -1 - (1 + 2\lambda_2) + 2 - 2(1 + 2\lambda_2) + (1 + 2\lambda_2)^2 \\ &= 4\lambda_2^2 - 2\lambda_2 - 1 \end{aligned}$$

with respect to λ_2. As $df/d\lambda_2 = 8\lambda_2 - 2 = 0$ at $\lambda_2^* = \frac{1}{4}$, we obtain

$$\mathbf{X}_3 = \mathbf{X}_2 + \lambda_2^* \mathbf{S}_2 = \begin{Bmatrix} -1 \\ 1 \end{Bmatrix} + \frac{1}{4} \begin{Bmatrix} 0 \\ 2 \end{Bmatrix} = \begin{Bmatrix} -1 \\ 1.5 \end{Bmatrix}$$

$$\nabla f = \begin{Bmatrix} \partial f/\partial x_1 \\ \partial f/\partial x_2 \end{Bmatrix} = \begin{Bmatrix} 1 + 4x_1 + 2x_2 \\ -1 + 2x_1 + 2x_2 \end{Bmatrix} \qquad \nabla f_3 = \nabla f(\mathbf{X}_3) = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix},$$

$$\nabla f = \begin{Bmatrix} \partial f/\partial x_1 \\ \partial f/\partial x_2 \end{Bmatrix} = \begin{Bmatrix} 1 + 4x_1 + 2x_2 \\ -1 + 2x_1 + 2x_2 \end{Bmatrix} \qquad \mathbf{X}_3 = \begin{Bmatrix} -1 \\ 1.5 \end{Bmatrix} \qquad \nabla f_3 = \nabla f(\mathbf{X}_3) = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix},$$

Thus the optimum point is reached in two iterations

Ch3: Constrained optimization techniques

7.1 INTRODUCTION

This chapter deals with techniques that are applicable to the solution of the constrained optimization problem:

$$\text{Find } \mathbf{X} \text{ which minimizes } f(\mathbf{X})$$

subject to

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

$$h_k(\mathbf{X}) = 0, \qquad k = 1,2,\ldots,p \tag{7.1}$$

There are many techniques available for the solution of a constrained nonlinear programming problem. All the methods can be classified into two broad categories: direct methods and indirect methods, as shown in Table 7.1. In the *direct methods*, the constraints are handled in an explicit manner, whereas in most of the *indirect methods*, the constrained problem is solved as a sequence of unconstrained minimization problems. We discuss in this chapter all the methods indicated in Table 7.1.

TABLE 7.1 Constrained Optimization Techniques

Direct Methods	Indirect Methods
Random search methods	Transformation of variables technique
Heuristic search methods	Sequential unconstrained minimization techniques
Complex method	Interior penalty function method
Objective and constraint approximation methods	Exterior penalty function method
Sequential linear programming method	Augmented Lagrange multiplier method
Sequential quadratic programming method	
Methods of feasible directions	
Zoutendijk's method	
Rosen's gradient projection method	
Generalized reduced gradient method	

Indirect methods

3.1 Transformation techniques

[Q1] **discuss transformation technique as an indirect method for solving constrained optimization problem.**

If the constraints gj(X) are explicit functions of the variables x_i- and have certain simple forms, it may be possible to make a transformation of the independent variables such that the constraints are satisfied automatically . Thus it may be possible to convert a constrained optimization problem into an unconstrained one by making change of variables. Some typical transformations are indicated below:

If lower and upper bounds on x_i are specified as

$$a \leq x_i \leq b \qquad\qquad l_i \leq x_i \leq u_i$$

These can be satisfied by transformation of the variable x_i as

$$x_i = a + (b - a)\sin^2 y_i \qquad\qquad x_i = l_i + (u_i - l_i)\sin^2 y_i$$

where y_i is the new variable, which can take any value.

2. If a variable x_i is restricted to lie in the interval (0,1), we can use the transformation:

2. If a variable x_i is restricted to lie in the interval (0,1), we can use the transformation:

$$x_i = \sin^2 y_i \,, \qquad x_i = \cos^2 y_i \,,$$

$$x_i = \frac{e^{y_i}}{e^{y_i} + e^{-y_i}} \,, \text{ or } \qquad x_i = \frac{y_i^2}{1 + y_i^2}$$

3. If the variable x_i is constrained to take only positive values, the transformation can be

$$x_i = \text{abs}(y_i), \quad x_i = y_i^2 \quad \text{or} \quad x_i = e^{y_i} \qquad (7.151)$$

4. If the variable is restricted to take values lying only in between -1 and 1, the transformation can be

$$x_i = \sin y_i, \quad x_i = \cos y_i, \quad \text{or} \quad x_i = \frac{2y_i}{1 + y_i^2} \qquad (7.152)$$

Note the following aspects of transformation techniques:

1. The constraints gj(X) have to be very simple functions of x_1.

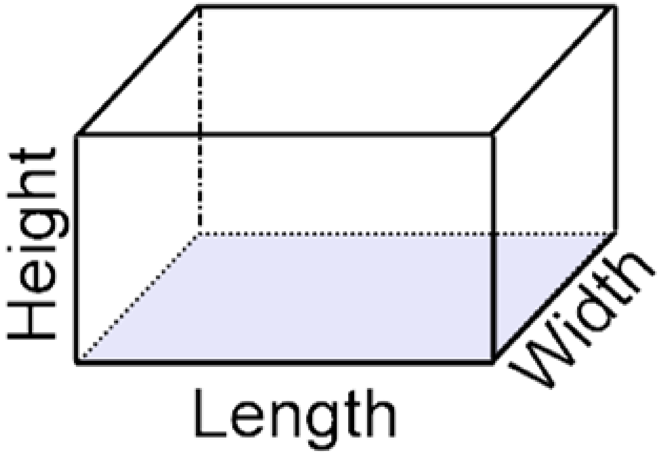
2. For certain constraints it may not be possible to find the necessary transformation.

3. If it is not possible to eliminate all the constraints by making change of variables, it may be better not to use the transformation at all. The partial transformation may sometimes produce a distorted objective function which might be more difficult to minimize than the original function.

To illustrate the method of transformation of variables, we consider the following problem.

[Q2] **Find the dimensions of a rectangular prism type box that has the largest volume when the sum of its length, width, and height is limited to a maximum value of 60 in. and its length is restricted to a maximum value of 36 in.**

SOLUTION Let x_1, x_2, and x_3 denote the length, width, and height of the box, respectively. The problem can be stated as follows:



$$\text{Maximize } f(x_1, x_2, x_3) = x_1 x_2 x_3$$

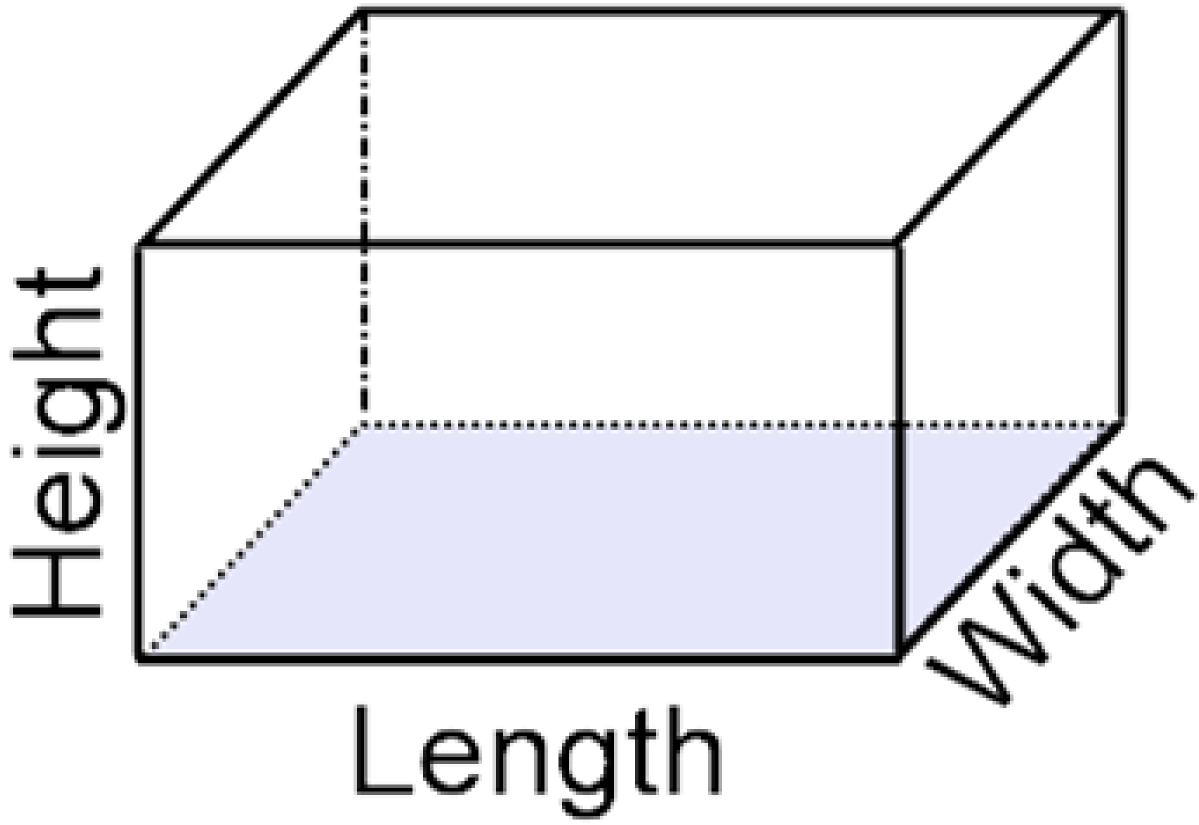
subject to

$$x_1 + x_2 + x_3 \leq 60$$

$$x_1 \leq 36 \quad (E_3)$$

$$x_i \geq 0, \quad i = 1,2,3 \quad (E_4)$$

By introducing new variables as


Height
Length
Width

$$x_1 + x_2 + x_3 \leq 60 \qquad (E_2)$$

$$x_1 \leq 36 \qquad (E_3)$$

$$x_i \geq 0, \quad i = 1,2,3 \qquad (E_4)$$

$$y_1 = x_1, \quad y_2 = x_2, \quad y_3 = x_1 + x_2 + x_3 \qquad (E_5)$$

or

$$x_1 = y_1, \quad x_2 = y_2, \quad x_3 = y_3 - y_1 - y_2 \qquad (E_6)$$

the constraints of Eqs. (E_2) to (E_4) can be restated as

$$0 \leq y_1 \leq 36, \quad 0 \leq y_2 \leq 60, \quad 0 \leq y_3 \leq 60 \qquad (E_7)$$

where the upper bound, for example, on y_2 is obtained by setting $x_1 = x_3 = 0$ in Eq. (E_2). The constraints of Eq. (E_7) will be satisfied automatically if we define new variables z_i, $i = 1,2,3$, as

$$y_1 = x_1, \quad y_2 = x_2, \quad y_3 = x_1 + x_2 + x_3 \tag{E$_5$}$$

or

$$x_1 = y_1, \quad x_2 = y_2, \quad x_3 = y_3 - y_1 - y_2 \tag{E$_6$}$$

the constraints of Eqs. (E_2) to (E_4) can be restated as

$$0 \le y_1 \le 36, \quad 0 \le y_2 \le 60, \quad 0 \le y_3 \le 60 \tag{E$_7$}$$

where the upper bound, for example, on y_2 is obtained by setting $x_1 = x_3 = 0$ in Eq. (E_2). The constraints of Eq. (E_7) will be satisfied automatically if we define new variables z_i, $i = 1,2,3$, as

$$y_1 = 36 \sin^2 z_1, \quad y_2 = 60 \sin^2 z_2, \quad y_3 = 60 \sin^2 z_3 \tag{E$_8$}$$

Thus the problem can be stated as an unconstrained problem as follows:

Maximize $f(x_1, x_2, x_3) = x_1 x_2 x_3$

$$x_1 = y_1, \quad x_2 = y_2, \quad x_3 = y_3 - y_1 - y_2$$

$$y_1 = 36 \sin^2 z_1, \quad y_2 = 60 \sin^2 z_2, \quad y_3 = 60 \sin^2 z_3$$

Maximize $f(z_1, z_2, z_3)$

$$= y_1 y_2 (y_3 - y_1 - y_2) \qquad (E_9)$$

$$= 2160 \sin^2 z_1 \sin^2 z_2 (60 \sin^2 z_3 - 36 \sin^2 z_1 - 60 \sin^2 z_2)$$

Maximize $f(z_1, z_2, z_3)$

$$= 2160 \sin^2 z_1 \sin^2 z_2 (60 \sin^2 z_3 - 36 \sin^2 z_1 - 60 \sin^2 z_2)$$

The necessary conditions of optimality yield the relations

$$\frac{\partial f}{\partial z_1} = 259{,}200 \sin z_1 \cos z_1 \sin^2 z_2 (\sin^2 z_3 - \tfrac{6}{5} \sin^2 z_1 - \sin^2 z_2) = 0 \qquad (E_{10})$$

$$\frac{\partial f}{\partial z_2} = 518{,}400 \sin^2 z_1 \sin z_2 \cos z_2 (\tfrac{1}{2} \sin^2 z_3 - \tfrac{3}{10} \sin^2 z_1 - \sin^2 z_2) = 0 \qquad (E_{11})$$

$$\frac{\partial f}{\partial z_1} = 259{,}200 \sin z_1 \cos z_1 \sin^2 z_2 \left(\sin^2 z_3 - \frac{6}{5}\sin^2 z_1 - \sin^2 z_2\right) = 0$$

$$\frac{\partial f}{\partial z_2} = 518{,}400 \sin^2 z_1 \sin z_2 \cos z_2 \left(\frac{1}{2}\sin^2 z_3 - \frac{3}{10}\sin^2 z_1 - \sin^2 z_2\right) = 0$$

$$\frac{\partial f}{\partial z_3} = 259{,}200 \sin^2 z_1 \sin^2 z_2 \sin z_3 \cos z_3 = 0 \qquad (E_{12})$$

$$\frac{\partial f}{\partial z_1} = 259{,}200 \sin z_1 \cos z_1 \sin^2 z_2 \left(\sin^2 z_3 - \frac{6}{5}\sin^2 z_1 - \sin^2 z_2\right) = 0$$

$$\frac{\partial f}{\partial z_2} = 518{,}400 \sin^2 z_1 \sin z_2 \cos z_2 \left(\frac{1}{2}\sin^2 z_3 - \frac{3}{10}\sin^2 z_1 - \sin^2 z_2\right) = 0$$

$$\frac{\partial f}{\partial z_1} = 0$$

$$259{,}200 \sin z_1 \cos z_1 \sin^2 z_2 = 0$$

$$\left(\sin^2 z_3 - \frac{6}{5}\sin^2 z_1 - \sin^2 z_2\right) = 0$$

$$\frac{\partial f}{\partial z_2} = 0$$

$$518{,}400 \sin^2 z_1 \sin z_2 \cos z_2 = 0$$

$$\left(\frac{1}{2}\sin^2 z_3 - \frac{3}{10}\sin^2 z_1 - \sin^2 z_2\right) = 0$$

$$\sin z_1 \cos z_1 \sin^2 z_2 = 0$$

$$\sin^2 z_3 - \frac{6}{5}\sin^2 z_1 - \sin^2 z_2 = 0$$

$$\sin^2 z_1 \sin z_2 \cos z_2 = 0$$

$$\frac{1}{2}\sin^2 z_3 - \frac{3}{10}\sin^2 z_1 - \sin^2 z_2 = 0$$

$$\sin z_1 \cos z_1 \sin^2 z_2 = 0$$
$$\sin^2 z_3 - \tfrac{6}{5}\sin^2 z_1 - \sin^2 z_2 = 0 \qquad (\mathrm{E}_{10})$$

$$\sin^2 z_1 \sin z_2 \cos z_2 = 0$$
$$\tfrac{1}{2}\sin^2 z_3 - \tfrac{3}{10}\sin^2 z_1 - \sin^2 z_2 = 0 \qquad (\mathrm{E}_{11})$$

$$\sin^2 z_1 \sin^2 z_2 \sin z_3 \cos z_3 = 0 \qquad (\mathrm{E}_{12})$$

Equation (E_{12}) gives the nontrivial solution as $\cos z_3 = 0$ or $\sin^2 z_3 = 1$. Hence Eqs. (E_{10}) and (E_{11}) yield $\sin^2 z_1 = \frac{5}{9}$ and $\sin^2 z_2 = \frac{1}{3}$. Thus the optimum solution is given by $x_1^* = 20$ in., $x_2^* = 20$ in., $x_3^* = 20$ in., and the maximum volume $= 8000\ \mathrm{in}^3$.

Ch 5 Dynamic Programming

Introduction

sequential decision problems- multistage decision problems

In most practical problems, decisions have to be made sequentially at different points in time, at different points in space, and at different levels, say, for a component, for a subsystem, and/or for a system. The problems in which the decisions are to be made sequentially are called *sequential decision problems.*
Since these decisions are to be made at a number of stages, they are also referred to as *multistage decision problems.* Dynamic programming is a mathematical technique well suited for the optimization of multistage decision problems.

The dynamic programming technique, when applicable, represents or decomposes a multistage decision problem as a sequence of single-stage decision problems.

Thus an *N*-variable problem is represented as a sequence of *N* single-variable problems that are solved successively.

In most cases, these N subproblems are easier to solve than the original problem.

The decomposition to N subproblems is done in such a manner that the optimal solution of the original N-variable problem can be obtained from the optimal solutions of the *N* one-dimensional problems.

Multistage decision problems can also be solved by direct application of the classical optimization techniques.(MCQ TF)

However, this requires the number of variables to be small, the functions involved to be continuous and continuously differentiable, and the optimum points not to lie at the boundary points. ***(MCQ TF)***

Further, the problem has to be relatively simple so that the set of resultant equations can be solved either analytically or numerically.

The nonlinear programming techniques can be used to solve slightly more complicated multistage decision problems.

But their application requires the variables to be continuous and prior knowledge about the region of the global minimum or maximum.

Multistage decision processes

Definition and Examples

As applied to dynamic programming, a multistage decision process is one in which a number of single-stage processes are connected in series so that the output of one stage is the input of the succeeding stage.

This type of process should be called a *serial multistage decision process*

Serial multistage decision problems arise in many types of practical problems. A few examples are given below

[Q1] Discuss *serial multistage decision process. Show how it can be represented schematically. Then represent the objective function as the composition of the individual stage returns.*

2 Representation of a Multistage Decision Process

A single-stage decision process (which is a component of the multistage problem) can be represented as a rectangular block(MCQ TF)

A decision process can be characterized by certain input parameters, S (or data), certain decision variables (X), and certain output parameters (T) representing the outcome obtained as a result of making the decision.

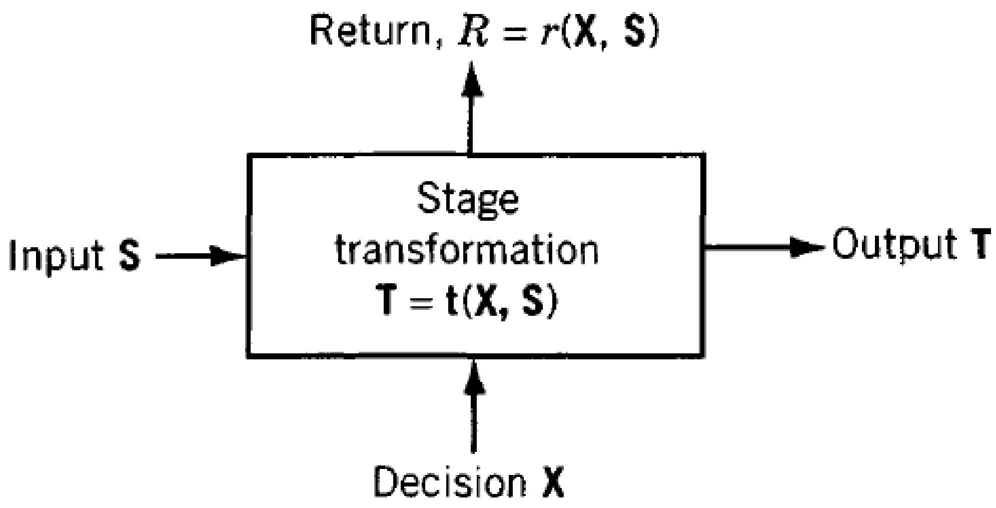


Figure 2 Single-stage decision problem.

The input parameters are called ***input state variables***, and the output parameters are called ***output state variables.***

Finally, there is a ***return or objective function*** *R,* which measures the effectiveness of the decisions made and the output that results from these decisions.

For a single-stage decision process shown in Fig. 2, the output is related to the input through a stage transformation function denoted by

$$\mathbf{T} = \mathbf{t}(\mathbf{X},\mathbf{S}) \qquad (9.1)$$

Since the input state of the system influences the decisions we make, the return function can be represented as

$$R = r(\mathbf{X},\mathbf{S}) \qquad (9.2)$$

A serial multistage decision process can be represented schematically as shown in Fig. 3.

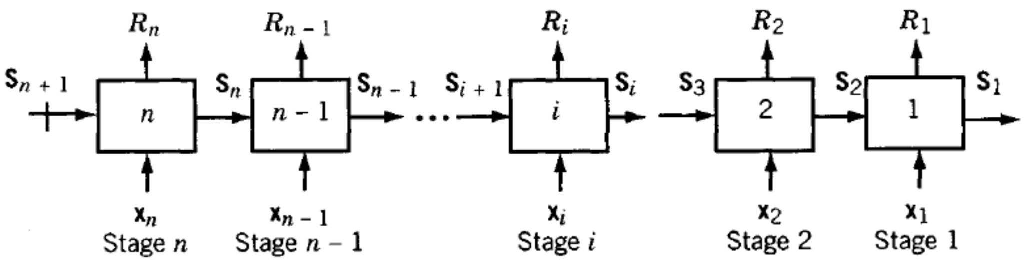


Fig. 3: Multistage decision problem (initial value problem).

the stages *n, n* — 1 , . . . , i , . . . , 2 , 1 are labeled in decreasing order. For the **i** *th* stage, the input state vector is denoted by **Si+1,** and the output state vector as Si Since the system is a serial one, the output from stage *i* + 1 must be equal to the input to stage *i.*

Hence the state transformation and return functions can be represented as

$$\mathbf{s}_i = \mathbf{t}_i(\mathbf{s}_{i+1},\mathbf{x}_i) \qquad (9.3)$$

$$R_i = r_i(\mathbf{s}_{i+1},\mathbf{x}_i) \qquad (9.4)$$

where Xi- denotes the vector of decision variables at stage i. The state transformation equations (3) are also called *design equations.*

The objective of a multistage decision problem is to find x1, x2 . . . , *xn* so as to optimize some function of the individual statge returns, say, $f(R_1,R_2,..,R_n)$ and satisfy Eqs. (3) and (4).

$$\mathbf{s}_i = \mathbf{t}_i(\mathbf{s}_{i+1},\mathbf{x}_i) \qquad (9.3)$$

$$R_i = r_i(\mathbf{s}_{i+1},\mathbf{x}_i) \qquad (9.4)$$

The nature of the n-stage return function, *f*, determines whether a given multistage problem can be solved by dynamic programming.

we must be able to represent the objective function as the composition of the individual stage returns. This requirement is satisfied for additive objective functions:

$$f = \sum_{i=1}^{n} R_i = \sum_{i=1}^{n} R_i(\mathbf{x}_i,\mathbf{s}_{i+1}) \qquad (9.5)$$

where Xi- are real, and for multiplicative objective functions,

$$f = \prod_{i=1}^{n} R_i = \prod_{i=1}^{n} R_i(\mathbf{x}_i,\mathbf{s}_{i+1}) \qquad (9.6)$$

where xi- are real and nonnegative. On the other hand, the following objective function is not separable:

$$f = [R_1(\mathbf{x}_1,\mathbf{s}_2) + R_2(\mathbf{x}_2,\mathbf{s}_3)][R_3(\mathbf{x}_3,\mathbf{s}_4) + R_4(\mathbf{x}_4,\mathbf{s}_5)] \qquad (9.7)$$

Fortunately, there are many practical problems that satisfy the separability condition. The objective function is said to be *monotonic* if for all values of **a** and **b** that make(MCQ TF)

$$R_i(\mathbf{x}_i = \mathbf{a}, \mathbf{s}_{i+1}) \geq R_i(\mathbf{x}_i = \mathbf{b}, \mathbf{s}_{i+1})$$

the following inequality is satisfied:

$$f(\mathbf{x}_n, \mathbf{x}_{n-1}, \ldots, \mathbf{x}_{i+1}, \mathbf{x}_i = \mathbf{a}, \mathbf{x}_{i-1}, \ldots, \mathbf{x}_1, \mathbf{s}_{n+1})$$
$$\geq f(\mathbf{x}_n, \mathbf{x}_{n-1}, \ldots, \mathbf{x}_{i+1}, \mathbf{x}_i = \mathbf{b}, \mathbf{x}_{i-1}, \ldots, \mathbf{x}_1, \mathbf{s}_{n+1}), \qquad i = 1,2,\ldots,n \qquad (9.8)$$

3 Conversion of a Nonserial System to a Serial System

According to the definition, a serial system is one whose components (stages) are connected in such a way that the output of any component is the input of the succeeding component.

As an example of a nonserial system, consider a steam power plant consisting of a pump, a feedwater heater, a boiler, a superheater, a steam turbine, and an electric generator, as shown in Fig. 4.

If we assume that some steam is taken from the turbine to heat the feedwater, a loop will be formed as shown in Fig. *4a*. This nonserial system can be converted to an equivalent serial system by regrouping the components so that a loop is redefined as a single element as shown in Fig. *4b* and *c*.

Thus the new serial multistage system consists of only three components: the pump, the boiler and turbine system, and the electric generator. This procedure can easily be extended to convert multistage systems with more than one loop to equivalent serial systems.

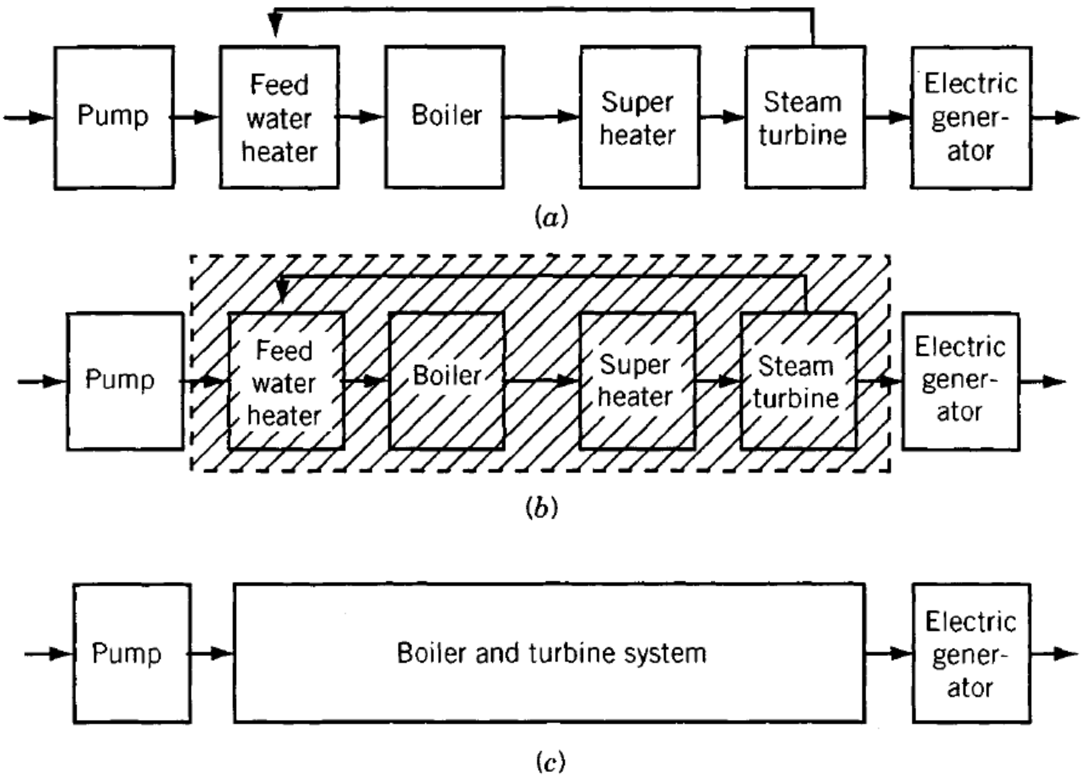


Figure 4 Serializing a nonserial system.

4 Types of Multistage Decision Problems

The serial multistage decision problems can be classified into three categories as follows.

1. *Initial Value Problem.* If the value of the initial state variable, *Sn+1*, is prescribed, the problem is called an *initial value problem*.*(MCQ MC)*

2. Final Value Problem. If the value of the final state variable S1 is prescribed, the problem is called *a final value problem.* *.(MCQ MC)*

Notice that a final value problem can be transformed into an initial value problem by reversing the directions of Si, *i* = 1, 2, . . . , *n* + 1. The details of this are given in Section 7.

3. *Boundary Value Problem.* If the values of both the input and output variables are specified, the problem is called a *boundary value problem.* The three types of problems are shown schematically in Fig. 5, where the symbol ⇸ is used to indicate a prescribed state variable.

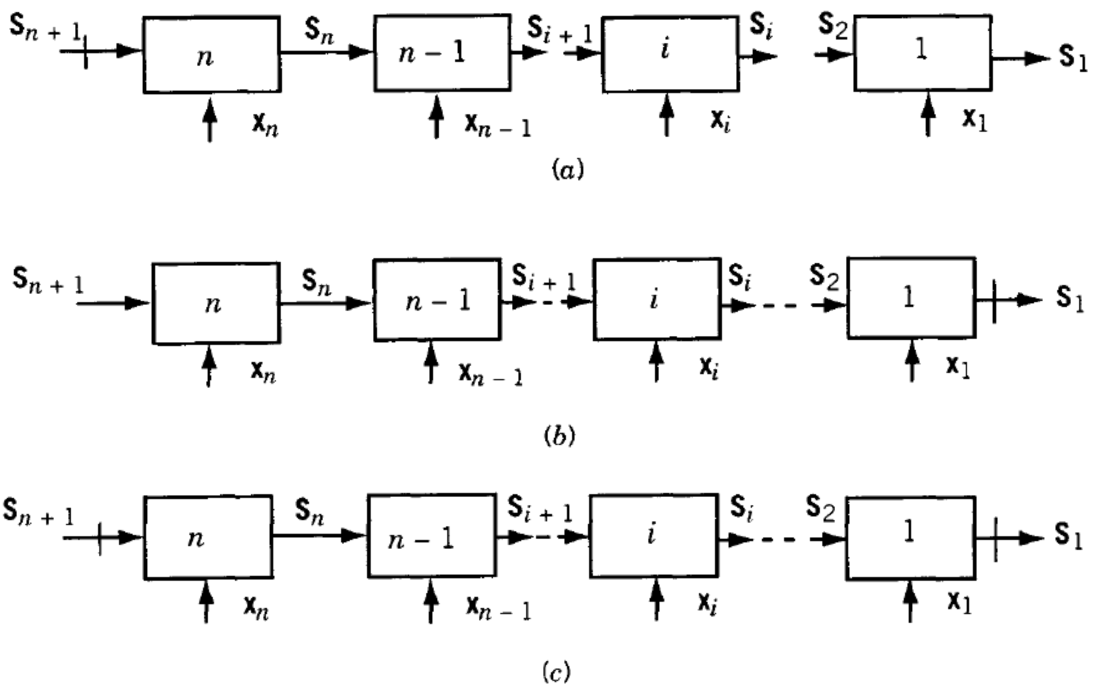


Figure 5 Types of multistage problems: *(a)* initial value problem; (b*)* final value problem; (c) boundary value problem.

1	***Multistage decision problems can be solved by direct application of the classical optimization techniques*.**	√
2	***A single-stage decision process can be represented as a rectangular block***	√

3	***The functions involved to be continuous and continuously differentiable, and the optimum points not to lie at the boundary points.***	√
4	***there are many practical problems that satisfy the separability condition. The objective function is said to be monotonic if for all values of a and b that make***	√

(4)In the Multistage decision problems, If the value of the initial state variable ……, is prescribed, the problem is called an initial value problem

Ⓐ	S_{n-1}	Ⓑ	S_n	Ⓒ	S_{n+1}	Ⓓ	S_{n+2}

(2) In the Multistage decision problems, If the value of the final state variable ………is prescribed, the problem is called a final value problem.

Ⓐ	S_n	Ⓑ	S_1	Ⓒ	S_2	Ⓓ	S_{n-1}

9.3 Concept Of Sub-optimization And Principle Of Optimality

[Q] Show how a dynamic programming problem can be reformulated as a sequence of sub- optimization problems, then define the Recurrence Relationship that joins them and obtain the ith subproblem in general state.

A dynamic programming problem can be stated as follows.

Find $x_1, x_2, \ldots, x_n$, which optimizes

$$f(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} R_i = \sum_{i=1}^{n} r_i(s_{i+1}, x_i)$$

and satisfies the design equations

$$s_i = t_i(s_{i+1}, x_i), \qquad i = 1, 2, \ldots, n$$

The dynamic programming makes use of the concept of suboptimization and the principle of optimality in solving this problem. The concept of suboptimization and the principle of optimality will be explained through the following example of an initial value problem.

Rem: In the subsequent discussion, the design variables *X1* and state variables s, are denoted as scalars for simplicity, although the theory is equally applicable even if they are vectors.

Example 9.1

Explain the concept of suboptimization in the context of the design of the water tank shown in Fig. 6a. The tank is required to have a capacity of 100,000 liters of water and is to be designed for minimum cost

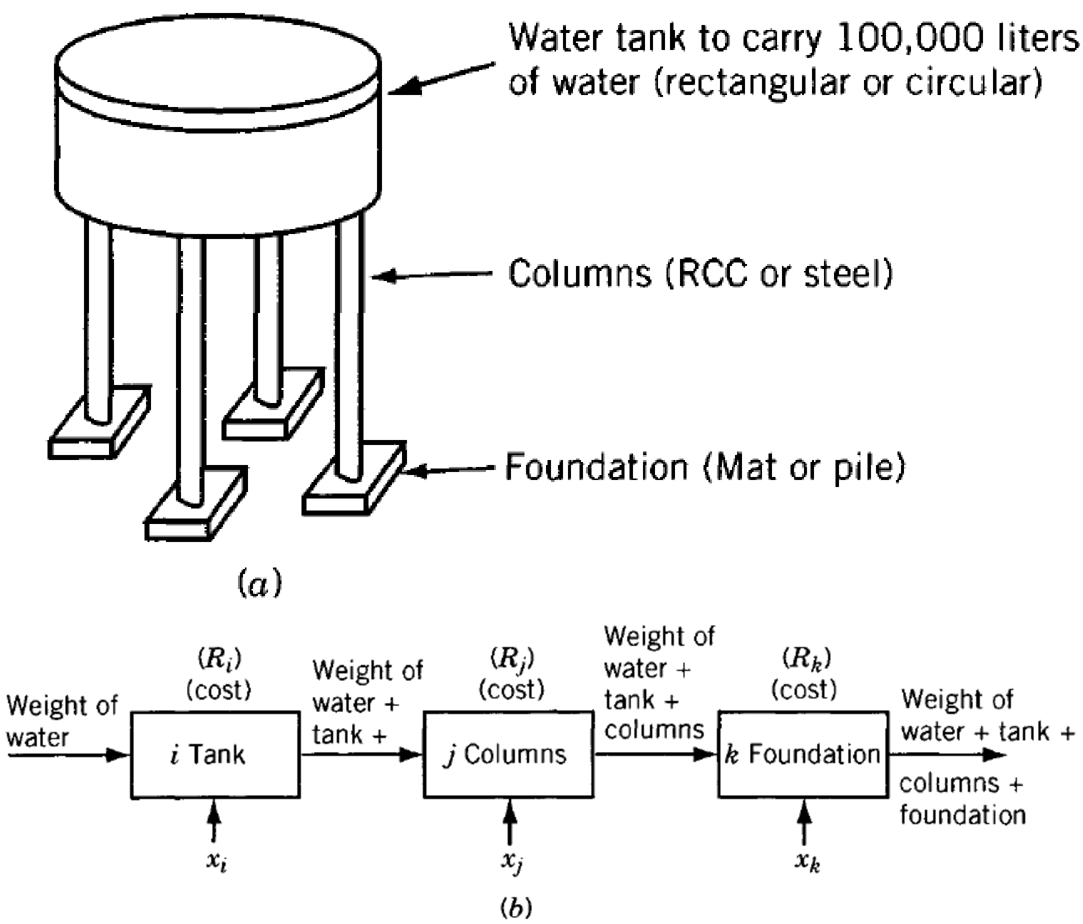


Figure 6 Water tank system.

SOLUTION

Instead of trying to optimize the complete system as a single unit, it would be desirable to *break* the system into components which could be optimized more or less individually.

For this breaking and component suboptimization, a logical procedure is to be used; otherwise, the procedure might result in a poor solution. This concept can be seen by breaking the system into three components: component i (tank), component *j* (columns), and component *k* (foundation).

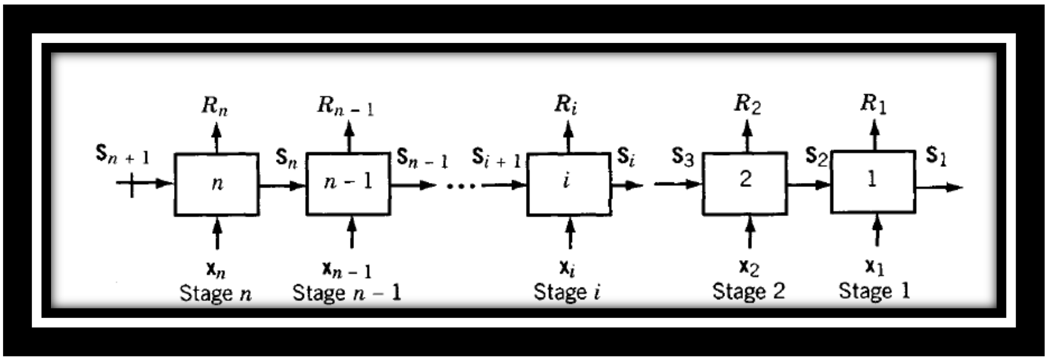
Consider the suboptimization of component *j* (columns) without a consideration of the other components. If the cost of steel is very high, the minimum cost design of component j may correspond to heavy concrete columns without reinforcement. Although this design may be acceptable for columns, the entire weight of the columns has to be carried by the foundation. This may result in a foundation that is prohibitively expensive.

This shows that the suboptimization of component *j* has adversely influenced the design of the following component *k*. This example shows that the design of any interior component affects the designs of all the subsequent (downstream) components. As such, it cannot be suboptimized without considering its effect on the downstream components.

The following mode of suboptimization can be adopted as a rational optimization strategy. Since the last component in a serial system influences no other component, it can be subopti-mized independently. Then the last two components can be considered together as a single (larger) component and can be suboptimized without adversely influencing any of the downstream components. This process can be continued to group any number of end components as a single (larger) end component and suboptimize them.

This process of suboptimization is shown in Fig. 7.

Since the suboptimizations are to be done in the reverse order, the components of the system are also numbered in the same manner for convenience (see Fig. 3).



The process of suboptimization was stated by Bellman as the principle of optimality:
An optimal policy (or a set of decisions) has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision.

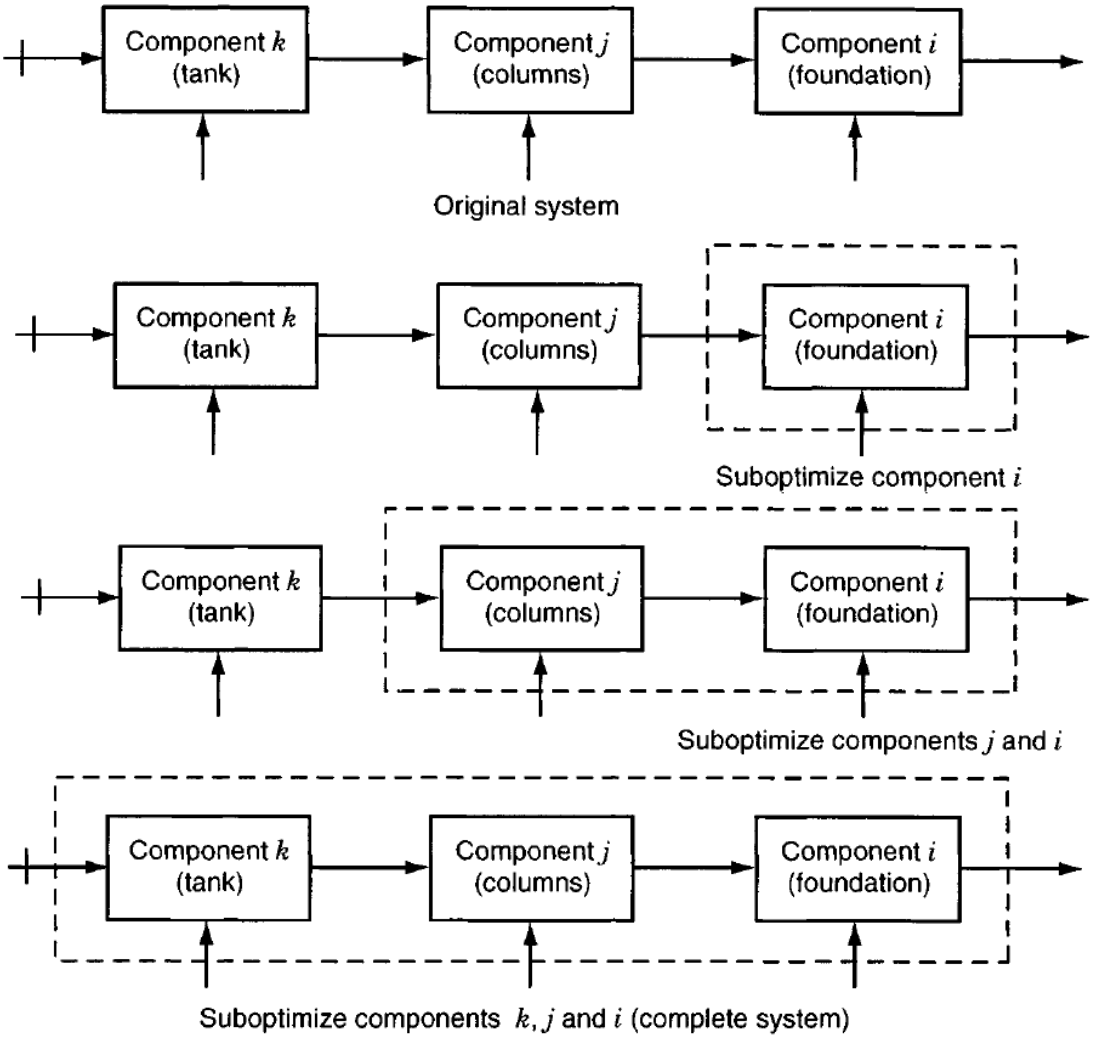


Figure 7 Suboptimization (principle of optimality).

Recurrence Relationship.

Suppose that the desired objective is to minimize the n-stage objective function f which is given by the sum of the individual stage returns:

$$\text{Minimize } f = R_n(x_n,s_{n+1}) + R_{n-1}(x_{n-1},s_n) + \cdots + R_1(x_1,s_2) \qquad (9.9)$$

where the state and decision variables are related as

$$s_i = t_i(s_{i+1},x_i), \qquad i = 1,2,\ldots,n \qquad (9.10)$$

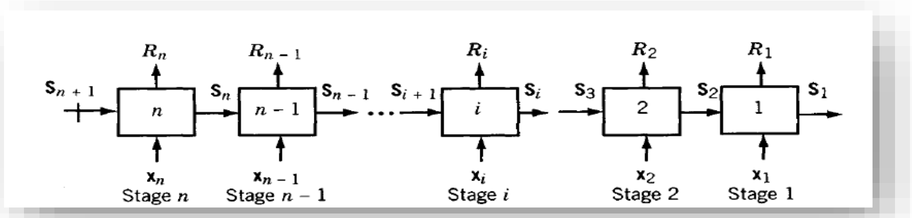
Consider the first subproblem by starting at the final stage, *i* = 1.
If the input to this stage S2 is specified, then according to the principle of optimality, *X1* must be selected to optimize *R1*. Irrespective of what happens to the other stages, *X1* must be selected such that *R1(x1,s2)* is an optimum for the input S2.

If the optimum is denoted as f_1^* , we have

$$f_1^*(s_2) = \underset{x_1}{\text{opt}}[R_1(x_1,s_2)] \qquad (9.11)$$

This is called a *one-stage policy* since once the input state *S2* is specified, the optimal values of *R1, X1,* and S1 are completely defined.
Thus Eq. (9.11) is a parametric equation giving the optimum f_1^* as a function of the input parameter ***S2.***

Next, consider the second subproblem by grouping the last two stages together.



If f_1^* denotes the optimum objective value of the second subproblem for a specified value of the input *S3*, we have

$$f_2^*(s_3) = \underset{x_1,x_2}{\text{opt}}[R_2(x_2,s_3) + R_1(x_1,s_2)] \qquad (9.12)$$

The principle of optimality requires that x_1 be selected so as to optimize R_1 for a given S_2. Since S_2 can be obtained once x_2 and S_3 are specified, Eq. (9.12) can be written as

$$f_2^*(s_3) = \underset{x_2}{\text{opt}}[R_2(x_2,s_3) + f_1^*(s_2)] \quad (9.13)$$

Thus f_2^* represents the optimal policy for the *two-stage subproblem.* It can be seen that the principle of optimality reduced the dimensionality of the problem from two [in Eq. (9.12)] to one [in Eq. (9.13)]. This can be seen more clearly by rewriting Eq. (9.13) using Eq. (9.10) as

$$s_i = t_i(s_{i+1},x_i), \quad i = 1,2,\ldots,n \quad (9.10$$

$$f_2^*(s_3) = \underset{x_2}{\text{opt}}[R_2(x_2,s_3) + f_1^*\{t_2(x_2,s_3)\}] \quad (9.14)$$

In this form it can be seen that for a specified input S3, the optimum is determined

solely by a suitable choice of the decision variable X2. Thus the optimization problem stated in Eq. (9.12), in which both *X2* and X1 are to be simultaneously varied to produce the optimum f*, is reduced to two subproblems defined by Eqs. (9.11) and (9.13). Since the optimization of each of these subproblems involves only a single decision variable, the optimization is, in general, much simpler.

This idea can be generalized and the i*th* subproblem defined by

$$f_i^*(s_{i+1}) = \underset{x_i,x_{i-1},\ldots,x_1}{\text{opt}} [R_i(x_i,s_{i+1}) + R_{i-1}(x_{i-1},s_i) + \cdots + R_1(x_1,s_2)] \quad (9.15)$$

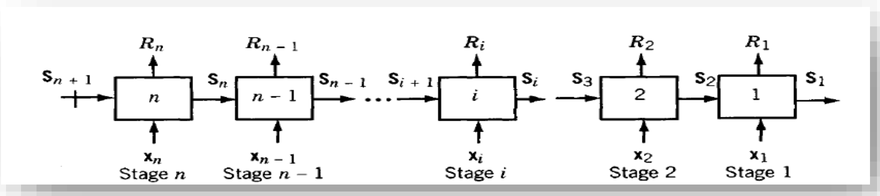
can be written as

$$f_i^*(s_{i+1}) = \underset{x_i}{\text{opt}}[R_i(x_i,s_{i+1}) + f_{i-1}^*(s_i)] \quad (9.16)$$

where f_{i-1}^* denotes the optimal value of the objective function corresponding to the last $i - 1$ stages, and s_i is the input to the stage $i - 1$. The original problem in Eq. (9.15) requires the simultaneous variation of i decision variables, $x_1, x_2, \ldots, x_i$, to determine the optimum value of $f_i = \sum_{k=1}^{i} R_k$ for any specified value of the input s_{i+1}. This problem, by using the principle of optimality, has been decomposed into i separate problems, each involving only one decision variable. Equation (9.16) is the desired recurrence relationship valid for $i = 2,3,\ldots,n$.

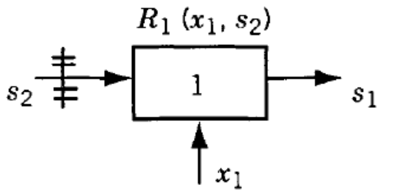
<u>4 computational procedure in dynamic Programming</u>

The use of the recurrence relationship derived in Section 3 in actual computations is discussed in this section. As stated, dynamic programming begins by suboptimizing the last component, numbered 1. This involves the determination of



$$f_1^*(s_2) = \underset{x_1}{\text{opt}}[R_1(x_1,s_2)] \qquad (9.17)$$

The best value of the decision variable x_1, denoted as x_1^*, is that which makes the return (or objective) function R_1 assume its optimum value, denoted by f_1^*. Both x_1^* and f_1^* depend on the condition of the input or feed that the component 1 receives from the upstream, that is, on s_2. Since the particular value s_2 will assume after the upstream components are optimized is not known at this time, this last-stage suboptimization problem is solved for a ''range'' of possible values of s_2 and the results are entered into a graph or a table. This graph or table contains a complete summary of the results of suboptimization of stage 1. In some cases, it may be possible to express f_1^* as a function of s_2.



(a)

s_2	x_1^*	f_1^*	s_1
—	—	—	—
—	—	—	—
—	—	—	—
—	—	—	—

(b) Summary of stage 1

Figure 8 Suboptimization of component 1 for various settings of the input state variable s_2.

If the calculations are to be performed on a computer, the results of suboptimization have to be stored in the form of a table in the computer. Figure 8 shows a typical table in which the results obtained from the suboptimization of stage 1 are entered.

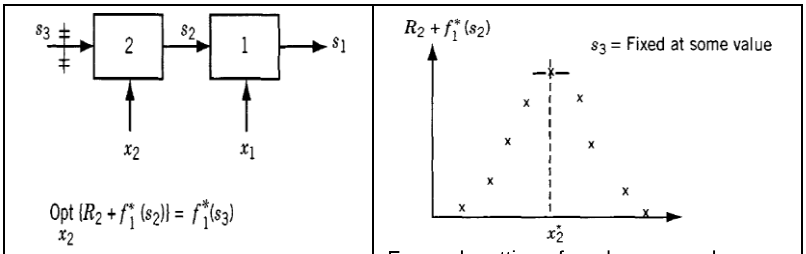
Next we move up the serial system to include the last two components. In this two-stage suboptimization, we have to determine

$$f_2^*(s_3) = \underset{x_2, x_1}{\text{opt}}[R_2(x_2,s_3) + R_1(x_1,s_2)] \tag{9.18}$$

Since all the information about component 1 has already been encoded in the table corresponding to f_1^*, this information can then be substituted for R_1 in Eq. (9.18) to get the following simplified statement:

$$f_2^*(s_3) = \underset{x_2}{\text{opt}}[R_2(x_2,s_3) + f_1^*(s_2)] \tag{9.19}$$

Thus the number of variables to be considered has been reduced from two (x_1 and x_2) to one (x_2). A range of possible values of s_3 must be considered and for each one, x_2^* must be found so as to optimize $[R_2 + f_1^*(s_2)]$. The results (x_2^* and f_2^* for different s_3) of this suboptimization are entered in a table as shown in Fig. 9.



For each setting of S_3, draw a graph as shown above to obtain the following:

s_3	x_2^*	f_2^*	s_2
—	—	—	—
—	—	—	—
—	—	—	—
—	—	—	—
—	—	—	—

(b) Summary of stages 2 and 1

Figure 9 Suboptimization of components 1 and 2 for various settings of the input state variable S_3.

Assuming that the suboptimization sequence has been carried on to include $i - 1$ of the end components, the next step will be to suboptimize the i end components. This requires the solution of

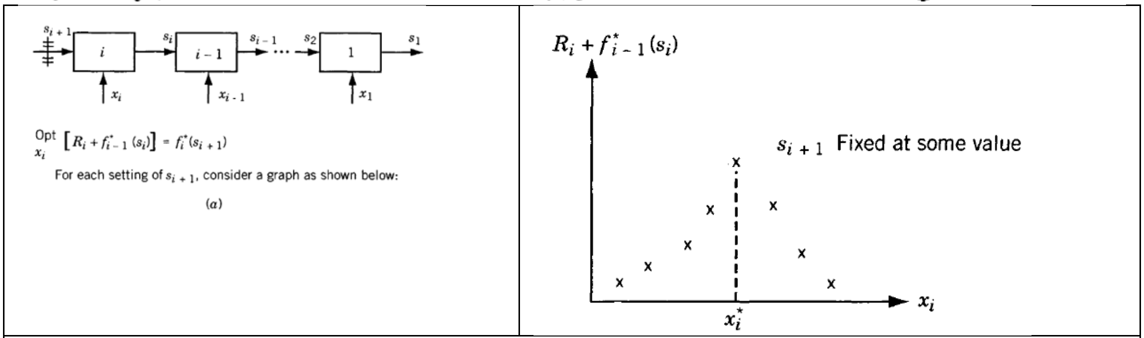
$$f_i^*(s_{i+1}) = \underset{x_i, x_{i-1}, \ldots, x_1}{\text{opt}} [R_i + R_{i-1} + \cdots + R_1] \tag{9.20}$$

However, again, all the information regarding the suboptimization of $i - 1$ end components is known and has been entered in the table corresponding to

f_{i-1}^*. Hence this information can be substituted in Eq. (9.20) to obtain

$$f_i^*(s_{i+1}) = \underset{x_i}{\text{opt}}[R_i(x_i,s_{i+1}) + f_{i-1}^*(s_i)] \tag{9.21}$$

Thus the dimensionality of the i-stage suboptimization has been reduced to 1, and the equation $s_i = t_i(s_{i+1},x_i)$ provides the functional relation between x_i and s_i. As before, a range of values of s_{i+1} are to be considered, and for each one, x_i^* is to be found so as to optimize $[R_i + f_{i-1}^*]$. A table showing the values of x_i^* and f_i^* for each of the values of s_{i+1} is made as shown in Fig. 9.10.



s_{i+1}	x_i^*	f_i^*	s_i
—	—	—	—
—	—	—	—
—	—	—	—

(b) Summary of stages *i, i-1 ...2,* and 1

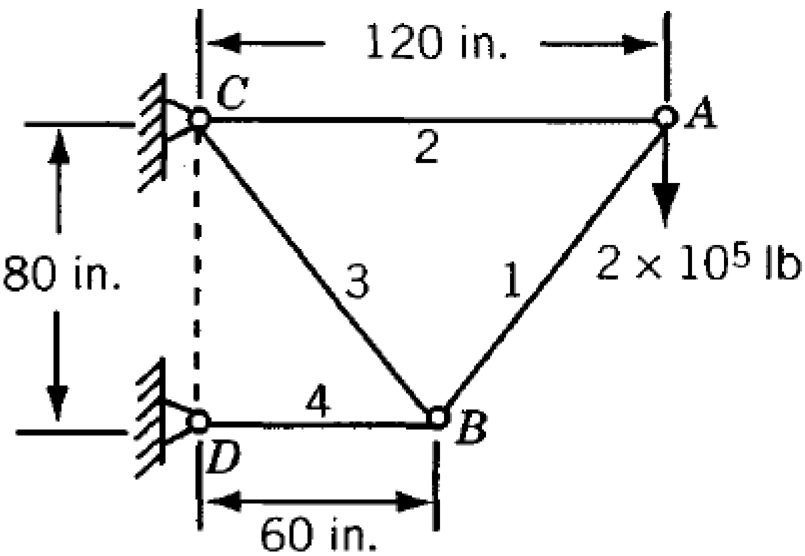
Figure 10 Suboptimization of components 1, 2, . . . , i for various settings of the input state variable *Si+1*

The suboptimization procedure above is continued until stage n is reached. At this stage only one value of s_{n+1} needs to be considered (for initial value problems), and the optimization of the n components completes the solution of the problem.

The final thing needed is to retrace the steps through the tables generated, to gather the complete set of x_i^* $(i = 1,2,\ldots,n)$ for the system. This can be done as follows. The nth suboptimization gives the values of x_n^* and f_n^* for the specified value of s_{n+1} (for initial value problem). The known design equation $s_n = t_n(s_{n+1}, x_n^*)$ can be used to find the input, s_n^*, to the $(n-1)$th stage. From the tabulated results for $f_{n-1}^*(s_n)$, the optimum values f_{n-1}^* and x_{n-1}^* corresponding to s_n^* can readily be obtained. Again the known design equation $s_{n-1} = t_{n-1}(s_n, x_{n-1}^*)$ can be used to find the input, s_{n-1}^*, to the $(n-2)$th stage. As before, from the tabulated results of $f_{n-2}^*(s_{n-1})$, the optimal values x_{n-2}^* and f_{n-2}^* corresponding to s_{n-1}^* can be found. This procedure is continued until the values x_1^* and f_1^* corresponding to s_2^* are obtained. Then the optimum solution vector of the original problem is given by $(x_1^*, x_2^*, \ldots, x_n^*)$ and the optimum value of the objective function by f_n^*.

5 Example Illustrating The Calculus Method Of Solution

Example 2 The four-bar truss shown in Fig. 11 is subjected to a vertical load of 2×10^5 lb at joint *A* as shown. Determine the cross-sectional areas of the members (bars) such that the total weight of the truss is minimum and the vertical deflection of joint *A* is equal to 0.5 in. Assume the unit weight as 0.01 lb/in^3 and the Young's modulus as 20×10^6 psi.



SOLUTION Let x_i denote the area of cross section of member i $(i = 1,2,3,4)$. The lengths of members are given by $l_1 = l_3 = 100$ in., $l_2 = 120$ in., and $l_4 = 60$ in. The weight of the truss is given by

$$f(x_1,x_2,x_3,x_4) = 0.01(100x_1 + 120x_2 + 100x_3 + 60x_4)$$

$$= x_1 + 1.2x_2 + x_3 + 0.6x_4 \qquad (E_1)$$

From structural analysis [9.5], the force developed in member i due to a unit load acting at joint $A(p_i)$, the deformation of member i (d_i), and the contribution of member i to the vertical deflection of A ($\delta_i = p_i d_i$) can be determined as follows:

Member i	p_i	$d_i = \frac{(\text{stress}_i) l_i}{E} = \frac{P p_i l_i}{x_i E}$ (in.)	$\delta_i = p_i d_i$ (in.)
1	−1.25	$-1.25/x_1$	$1.5625/x_1$
2	0.75	$0.9/x_2$	$0.6750/x_2$
3	1.25	$1.25/x_3$	$1.5625/x_3$
4	−1.50	$-0.9/x_4$	$1.3500/x_4$

The vertical deflection of joint A is given by

$$d_A = \sum_{i=1}^{4} \delta_i = \frac{1.5625}{x_1} + \frac{0.6750}{x_2} + \frac{1.5625}{x_3} + \frac{1.3500}{x_4} \qquad (E_2)$$

Thus the optimization problem can be stated as:

[Q] Consider the problem of four-bar truss, which is formulated mathematically as a non linear programming problem:

Minimize

$$f(\mathbf{X}) = x_1 + 1.2x_2 + x_3 + 0.6x_4$$

subject to

$$\frac{1.5625}{x_1} + \frac{0.6750}{x_2} + \frac{1.5625}{x_3} + \frac{1.3500}{x_4} = 0.5 \qquad (E_3)$$

$$x_1 \geq 0, \quad x_2 \geq 0, \quad x_3 \geq 0, \quad x_4 \geq 0$$

Show that the problem can be posed as a multistage decision problem. Then obtain the optimum solution of it

----

Answer:

Since the deflection of joint A is the sum of contributions of the various members, we can consider the 0.5 in. deflection as a resource to be allocated to the various activities x_i and the problem can be posed as a multistage decision problem as shown in Fig. 9.12. Let s_2 be the displacement (resource) available for allocation to the first member (stage 1), δ_1 the displacement contribution due to the first member, and $f_1^*(s_2)$ the minimum weight of the first member.

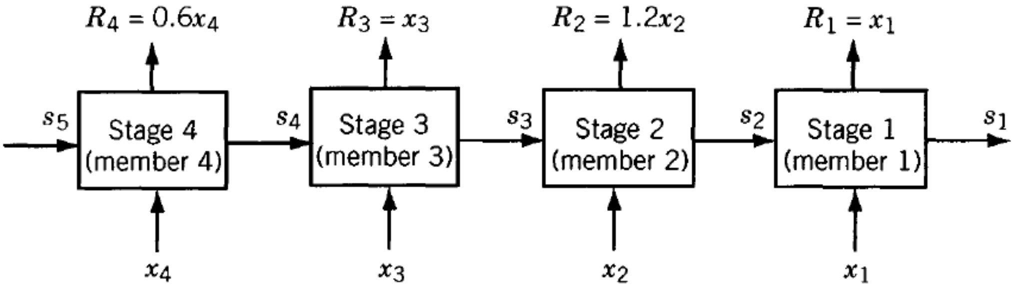


Figure 9.12 Example 9.2 as a four-stage decision problem.

Then

$$f_1^*(s_2) = \min[R_1 = x_1] = \frac{1.5625}{s_2} \tag{E$_4$}$$

such that

$$\delta_1 = \frac{1.5625}{x_1} \quad \text{and} \quad x_1 \geq 0$$

since $\delta_1 = s_2$, and

$$x_1^* = \frac{1.5625}{s_2} \tag{E$_5$}$$

Let s_3 be the displacement available for allocation to the first two members, δ_2 the displacement contribution due to the second member, and $f_2^*(s_3)$ the minimum weight of the first two members. Then we have, from the recurrence relationship of Eq. (9.16),

$$f_i^*(s_{i+1}) = \underset{x_i}{\text{opt}}[R_i(x_i,s_{i+1}) + f_{i-1}^*(s_i)] \tag{9.16}$$

$$f_2^*(s_3) = \min_{x_2 \geq 0}[R_2 + f_1^*(s_2)] \tag{E$_6$}$$

where s_2 represents the resource available after allocation to stage 2 and is given by

$$s_2 = s_3 - \delta_2 = s_3 - \frac{0.6750}{x_2}$$

Hence, from Eq. (E$_4$), we have

$$f_1^*(s_2) = \min[R_1 = x_1] = \frac{1.5625}{s_2} \qquad (E_4)$$

$$f_1^*(s_2) = f_1^*\left(s_3 - \frac{0.6750}{x_2}\right) = \left[1.5625 \Big/ \left(s_3 - \frac{0.6750}{x_2}\right)\right] \qquad (E_7)$$

Thus Eq. (E_6) becomes

$$f_2^*(s_3) = \min_{x_2 \geq 0}\left[1.2x_2 + \frac{1.5625}{s_3 - 0.6750/x_2}\right] \qquad (E_8)$$

Let

$$F(s_3,x_2) = 1.2x_2 + \frac{1.5625}{s_3 - 0.6750/x_2} = 1.2x_2 + \frac{1.5625x_2}{s_3x_2 - 0.6750}$$

For any specified value of s_3, the minimum of F is given by

$$\frac{\partial F}{\partial x_2} = 1.2 - \frac{(1.5625)(0.6750)}{(s_3x_2 - 0.6750)^2} = 0 \quad \text{or} \quad x_2^* = \frac{1.6124}{s_3} \qquad (E_9)$$

$$f_2^*(s_3) = 1.2x_2^* + \frac{1.5625}{s_3 - 0.6750/x_2^*} = \frac{1.9349}{s_3} + \frac{2.6820}{s_3} = \frac{4.6169}{s_3} \qquad (E_{10})$$

Let s_4 be the displacement available for allocation to the first three members. Let δ_3 be the displacement contribution due to the third member and $f_3^*(s_4)$ the minimum weight of the first three members. Then

$$f_3^*(s_4) = \min_{x_3 \geq 0}[x_3 + f_2^*(s_3)] \qquad (E_{11})$$

$$\frac{1.5625}{x_1} + \frac{0.6750}{x_2} + \frac{1.5625}{x_3} + \frac{1.3500}{x_4} = 0.5 \qquad (E_3)$$

where s_3 is the resource available after allocation to stage 3 and is given by

$$s_3 = s_4 - \delta_3 = s_4 - \frac{1.5625}{x_3}$$

From Eq. (E_{10}) we have

$$f_2^*(s_3) = \frac{4.6169}{s_4 - 1.5625/x_3} \qquad (E_{12})$$

and Eq. (E_{11}) can be written as

$$f_3^*(s_4) = \min_{x_3 \geq 0} \left[x_3 + \frac{4.6169x_3}{s_4x_3 - 1.5625} \right] \qquad (E_{13})$$

As before, by letting

$$F(s_4,x_3) = x_3 + \frac{4.6169x_3}{s_4x_3 - 1.5625} \qquad (E_{14})$$

the minimum of F, for any specified value of s_4, can be obtained as

$$\frac{\partial F}{\partial x_3} = 1.0 - \frac{(4.6169)(1.5625)}{(s_4x_3 - 1.5625)^2} = 0 \quad \text{or} \quad x_3^* = \frac{4.2445}{s_4} \qquad (E_{15})$$

$$f_3^*(s_4) = x_3^* + \frac{4.6169x_3^*}{s_4x_3^* - 1.5625} = \frac{4.2445}{s_4} + \frac{7.3151}{s_4} = \frac{11.5596}{s_4} \qquad (E_{16})$$

Finally, let s_5 denote the displacement available for allocation to the first four members. If δ_4 denotes the displacement contribution due to the fourth member, and $f_4^*(s_5)$ the minimum weight of the first four members, then

$$f_4^*(s_5) = \min_{x_4 \geq 0} [0.6x_4 + f_3^*(s_4)] \qquad (E_{17})$$

where the resource available after allocation to the fourth member (s_4) is given by

$$\frac{1.5625}{x_1} + \frac{0.6750}{x_2} + \frac{1.5625}{x_3} + \frac{1.3500}{x_4} = 0.5 \qquad (E_3)$$

$$s_4 = s_5 - \delta_4 = s_5 - \frac{1.3500}{x_4} \tag{E$_{18}$}$$

From Eqs. (E_{16}), (E_{17}), and (E_{18}), we obtain

$$f_4^*(s_5) = \min_{x_4 \geq 0} \left[0.6x_4 + \frac{11.5596}{s_5 - 1.3500/x_4} \right] \tag{E$_{19}$}$$

By setting

$$F(s_5,x_4) = 0.6x_4 + \frac{11.5596}{s_5 - 1.3500/x_4}$$

the minimum of $F(s_5,x_4)$, for any specified value of s_5, is given by

$$\frac{\partial F}{\partial x_4} = 0.6 - \frac{(11.5596)(1.3500)}{(s_5 x_4 - 1.3500)^2} = 0 \quad \text{or} \quad x_4^* = \frac{6.44}{s_5} \tag{E$_{20}$}$$

$$f_4^*(s_5) = 0.6x_4^* + \frac{11.5596}{s_5 - 1.3500/x_4^*} = \frac{3.864}{s_5} + \frac{16.492}{s_5} = \frac{20.356}{s_5} \tag{E$_{21}$}$$

Since the value of s_5 is specified as 0.5 in., the minimum weight of the structure can be calculated from Eq. (E_{21}) as

$$f_4^*(s_5 = 0.5) = \frac{20.356}{0.5} = 40.712 \text{ lb} \tag{E$_{22}$}$$

Once the optimum value of the objective function is found, the optimum values of the design variables can be found with the help of Eqs. (E_{20}), (E_{15}), (E_9), and (E_5) as

$$x_4^* = 12.88 \text{ in}^2$$

$$s_4 = s_5 - \frac{1.3500}{x_4^*} = 0.5 - 0.105 = 0.395 \text{ in.}$$

$$x_3^* = \frac{4.2445}{s_4} = 10.73 \text{ in}^2$$

$$s_3 = s_4 - \frac{1.5625}{x_3^*} = 0.3950 - 0.1456 = 0.2494 \text{ in.}$$

$$x_2^* = \frac{1.6124}{s_3} = 6.47 \text{ in}^2$$

$$s_2 = s_3 - \frac{0.6750}{x_2^*} = 0.2494 - 0.1042 = 0.1452 \text{ in.}$$

$$x_1^* = \frac{1.5625}{s_2} = 10.76 \text{ in}^2$$