

5. Cars arrive at a one-bay car wash facility where the interarrival time is exponential, with a mean of 10 minutes. Arriving cars line up in a single lane that can accommodate at most five waiting cars. If the lane is full, newly arriving cars will go elsewhere. It takes between 10 and 15 minutes, uniformly distributed, to wash a car. Simulate the system for 960 minutes, and estimate the time a car spends in the facility.

BIBLIOGRAPHY

- ✓ Banks, J., J. Carson, B. Nelson, and D. Nicol, *Discrete-Event System Simulation*, 4th ed., Prentice Hall, Upper Saddle River, NJ, 2005.
- Box, G., and M. Muller, "A Note on the Generation of Random Normal Deviates," *Annals of Mathematical Statistics*, Vol. 29, pp. 610-611, 1958.
- Elizandro, D., and H. Taha, *Simulation of Industrial Systems: Discrete Event Simulation Using Excel/VBA*, Taylor and Francis, New York, 2008.
- ✓ Law, A., *Simulation Modeling & Analysis*, 4th ed., McGraw-Hill, New York, 2007.
- Rubenstein, R., B. Melamed, and A. Shapiro, *Modern Simulation and Modeling*, Wiley, New York, 1998.
- Taha, H., *Simulation Modeling and SIMNET*, Prentice Hall, Upper Saddle River, NJ, 1988.

CHAPTER 20

Classical Optimization Theory

資料來源：

Operations Research
An Introduction 9th ed
Hamdy A. Taha.

20.1 UNCONSTRAINED PROBLEMS

An extreme point of a function $f(\mathbf{X})$ defines either a maximum or a minimum of the function. Mathematically, a point $\mathbf{X}_0 = (x_1^0, \dots, x_j^0, \dots, x_n^0)$ is a maximum if

$$f(\mathbf{X}_0 + \mathbf{h}) \leq f(\mathbf{X}_0)$$

for all $\mathbf{h} = (h_1, \dots, h_j, \dots, h_n)$, where $|h_j|$ is sufficiently small for all j . In a similar manner, \mathbf{X}_0 is a minimum if

$$f(\mathbf{X}_0 + \mathbf{h}) \geq f(\mathbf{X}_0)$$

Figure 20.1 illustrates the maxima and minima of a single-variable function $f(x)$ defined in the range $a \leq x \leq b$. The points x_1, x_2, x_3, x_4 , and x_6 are all extrema of $f(x)$, with x_1, x_3 , and x_6 as maxima and x_2 and x_4 as minima. The value of $f(x_6) = \max\{f(x_1), f(x_3), f(x_6)\}$ is a **global** or **absolute** maximum, and $f(x_1)$ and $f(x_3)$ are **local** or **relative** maxima. Similarly, $f(x_4)$ is a local minimum and $f(x_2)$ is a global minimum.

Although x_1 (in Figure 20.1) is a (local) maximum point, it differs from remaining local maxima in that the value of f corresponding to at least one point in the neighborhood of x_1 equals $f(x_1)$. In this respect, x_1 is a **weak maximum**, whereas x_3 and x_6 are **strong maxima**. In general, for \mathbf{h} as defined earlier, \mathbf{X}_0 is a weak maximum if $f(\mathbf{X}_0 + \mathbf{h}) \leq f(\mathbf{X}_0)$ and a strong maximum if $f(\mathbf{X}_0 + \mathbf{h}) < f(\mathbf{X}_0)$. This first derivative (slope) of f equals zero at all extrema. This

In Figure 20.1, the first derivative (slope) of f equals zero at all extrema. This property is also satisfied at **inflection** and **saddle** points, such as x_5 . If a point with zero slope (gradient) is not an extremum (maximum or minimum), then it must be an inflection or a saddle point.

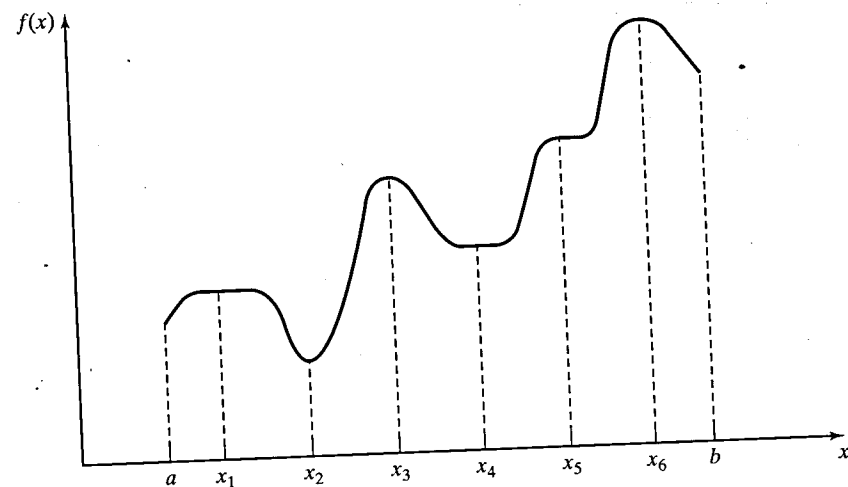


FIGURE 20.1
Examples of extreme points for a single-variable function

20.1.1 Necessary and Sufficient Conditions

This section develops the necessary and sufficient conditions for an n -variable function $f(\mathbf{X})$ to have extrema. It is assumed that the first and second partial derivatives of $f(\mathbf{X})$ are continuous for all \mathbf{X} .

Theorem 20.1-1 A necessary condition for \mathbf{X}_0 to be an extreme point of $f(\mathbf{X})$ is that

$$\nabla f(\mathbf{X}_0) = 0$$

Because the necessary condition is also satisfied at inflection and saddle points, it is more appropriate to refer to the points obtained from the solution of $\nabla f(\mathbf{X}_0) = 0$ as **stationary points**. The next theorem establishes the sufficiency conditions for \mathbf{X}_0 to be an extreme point.

Theorem 20.1-2 A sufficient condition for a stationary point \mathbf{X}_0 to be an extremum is that the Hessian matrix \mathbf{H} evaluated at \mathbf{X}_0 satisfy the following conditions:

- (i) \mathbf{H} is positive definite if \mathbf{X}_0 is a minimum point.
- (ii) \mathbf{H} is negative definite if \mathbf{X}_0 is a maximum point.

Example 20.1-1

Consider the function

$$f(x_1, x_2, x_3) = x_1 + 2x_3 + x_2x_3 - x_1^2 - x_2^2 - x_3^2$$

The necessary condition $\nabla f(\mathbf{X}_0) = 0$ gives

$$\frac{\partial f}{\partial x_1} = 1 - 2x_1 = 0$$

$$\frac{\partial f}{\partial x_2} = x_3 - 2x_2 = 0$$

$$\frac{\partial f}{\partial x_3} = 2 + x_2 - 2x_3 = 0$$

The solution of these simultaneous equations is

$$\mathbf{X}_0 = \left(\frac{1}{2}, \frac{2}{3}, \frac{4}{3} \right)$$

To determine the type of the stationary point, consider

$$\mathbf{H}|_{\mathbf{X}_0} = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_1 \partial x_3} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \frac{\partial^2 f}{\partial x_2 \partial x_3} \\ \frac{\partial^2 f}{\partial x_3 \partial x_1} & \frac{\partial^2 f}{\partial x_3 \partial x_2} & \frac{\partial^2 f}{\partial x_3^2} \end{pmatrix}_{\mathbf{X}_0} = \begin{pmatrix} -2 & 0 & 0 \\ 0 & -2 & 1 \\ 0 & 1 & -2 \end{pmatrix}$$

The principal minor determinants of $\mathbf{H}|_{\mathbf{X}_0}$ have the values -2 , 4 , and -6 , respectively. Thus, as shown in Section D.3, $\mathbf{H}|_{\mathbf{X}_0}$ is negative-definite, and $\mathbf{X}_0 = \left(\frac{1}{2}, \frac{2}{3}, \frac{4}{3} \right)$ represents a maximum point.

In general, if $\mathbf{H}|_{\mathbf{X}_0}$ is indefinite, \mathbf{X}_0 must be a saddle point. For nonconclusive cases, \mathbf{X}_0 may or may not be an extremum, and the sufficiency condition becomes rather involved, because higher-order terms in Taylor's expansion must be considered.

The sufficiency condition established by Theorem 20.1-2 applies to single-variable functions as follows. Given that y_0 is a stationary point, then

- (i) y_0 is a maximum if $f''(y_0) < 0$.
- (ii) y_0 is a minimum if $f''(y_0) > 0$.

If $f''(y_0) = 0$, higher-order derivatives must be investigated as the following theorem requires.

Theorem 20.1-3 Given y_0 , a stationary point of $f(y)$, if the first $(n - 1)$ derivatives are zero and $f^{(n)}(y_0) \neq 0$, then

- (i) If n is odd, y_0 is an inflection point.
- (ii) If n is even, then y_0 is a minimum if $f^{(n)}(y_0) > 0$ and a maximum if $f^{(n)}(y_0) < 0$.

FIGURE 20.2

Extreme points of $f(y) = y^4$ and $g(y) = y^3$

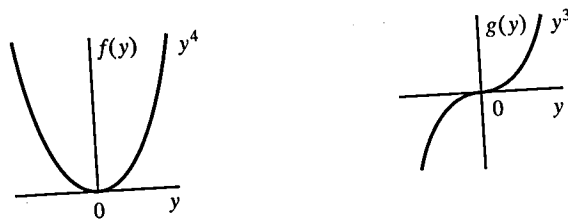
**Example 20.1-2**

Figure 20.2 graphs the following two functions:

$$f(y) = y^4$$

$$g(y) = y^3$$

For $f(y) = y^4$, $f'(y) = 4y^3 = 0$, which yields the stationary point $y_0 = 0$. Now

$$f'(0) = f''(0) = f^{(3)}(0) = 0, f^{(4)}(0) = 24 > 0$$

Hence, $y_0 = 0$ is a minimum point (see Figure 20.2).

For $g(y) = y^3$, $g'(y) = 3y^2 = 0$, which yields $y_0 = 0$ as a stationary point. Also

$$g'(0) = g''(0) = g^{(3)}(0) = 6 \neq 0$$

Thus, $y_0 = 0$ is an inflection point.

PROBLEM SET 20.1A

1. Determine the extreme points of the following functions.

* (a) $f(x) = x^3 + x$

* (b) $f(x) = x^4 + x^2$

(c) $f(x) = 4x^4 - x^2 + 5$

(d) $f(x) = (3x - 2)^2(2x - 3)^2$

(e) $f(x) = 6x^5 - 4x^3 + 10$

2. Determine the extreme points of the following functions.

(a) $f(\mathbf{X}) = x_1^3 + x_2^3 - 3x_1x_2$

(b) $f(\mathbf{X}) = 2x_1^2 + x_2^2 + x_3^2 + 6(x_1 + x_2 + x_3) + 2x_1x_2x_3$

3. Verify that the function

$$f(x_1, x_2, x_3) = 2x_1x_2x_3 - 4x_1x_3 - 2x_2x_3 + x_1^2 + x_2^2 + x_3^2 - 2x_1 - 4x_2 + 4x_3$$

has the stationary points $(0, 3, 1)$, $(0, 1, -1)$, $(1, 2, 0)$, $(2, 1, 1)$, and $(2, 3, -1)$. Use the sufficiency condition to identify the extreme points.

- *4. Solve the following simultaneous equations by converting the system to a nonlinear objective function with no constraints.

$$x_2 - x_1^2 = 0$$

$$x_2 - x_1 = 2$$

[Hint: $\min f^2(x_1, x_2)$ occurs at $f'(x_1, x_2) = 0$.]

20.1.2 The Newton-Raphson Method

In general, the necessary condition $\nabla f(x) \neq 0$ may be highly nonlinear and, hence, difficult to solve. The Newton-Raphson method is an iterative algorithm for solving simultaneous nonlinear equations.

Consider the simultaneous equations

$$f_i(\mathbf{X}) = 0, i = 1, 2, \dots, m$$

Let \mathbf{X}^k be a given point. Then by Taylor's expansion

$$f_i(\mathbf{X}) \approx f_i(\mathbf{X}_k) + \nabla f_i(\mathbf{X}_k)(\mathbf{X} - \mathbf{X}_k), i = 1, 2, \dots, m$$

Thus, the original equations, $f_i(\mathbf{X}) = 0, i = 1, 2, \dots, m$, may be approximated as

$$f_i(\mathbf{X}_k) + \nabla f_i(\mathbf{X}_k)(\mathbf{X} - \mathbf{X}_k) = 0, i = 1, 2, \dots, m$$

These equations may be written in matrix notation as

$$\mathbf{A}_k + \mathbf{B}_k(\mathbf{X} - \mathbf{X}_k) = 0$$

If \mathbf{B}_k is nonsingular, then

$$\mathbf{X} = \mathbf{X}_k - \mathbf{B}_k^{-1}\mathbf{A}_k$$

The idea of the method is to start from an initial point \mathbf{X}_0 , and then use the equation above to determine a new point. The process may or may not converge depending on the selection of the starting point. Convergence occurs when two successive points, \mathbf{X}_k and \mathbf{X}_{k+1} , are approximately equal (within specified acceptable tolerance).

A geometric interpretation of the method is illustrated by a single-variable function in Figure 20.3. The relationship between x_k and x_{k+1} for a single-variable function $f(x)$ reduces to

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$$

The terms may be arranged as $f'(x_k) = \frac{f(x_k)}{x_k - x_{k+1}}$ —meaning that x_{k+1} is determined from the slope of $f(x)$ at x_k , where $\tan \theta = f'(x_k)$, as the figure shows.

Figure 20.3 demonstrates that convergence is not always possible. If the initial point is a , the method will diverge. In general, it may be necessary to attempt a number of initial points before convergence is achieved.

Example 20.1-3

To demonstrate the use of the Newton-Raphson method, consider the function

$$g(x) = (3x - 2)^2(2x - 3)^2$$

To determine the stationary points of $g(x)$, we need to solve

$$f(x) = g'(x) = 72x^3 - 234x^2 + 241x - 78 = 0$$

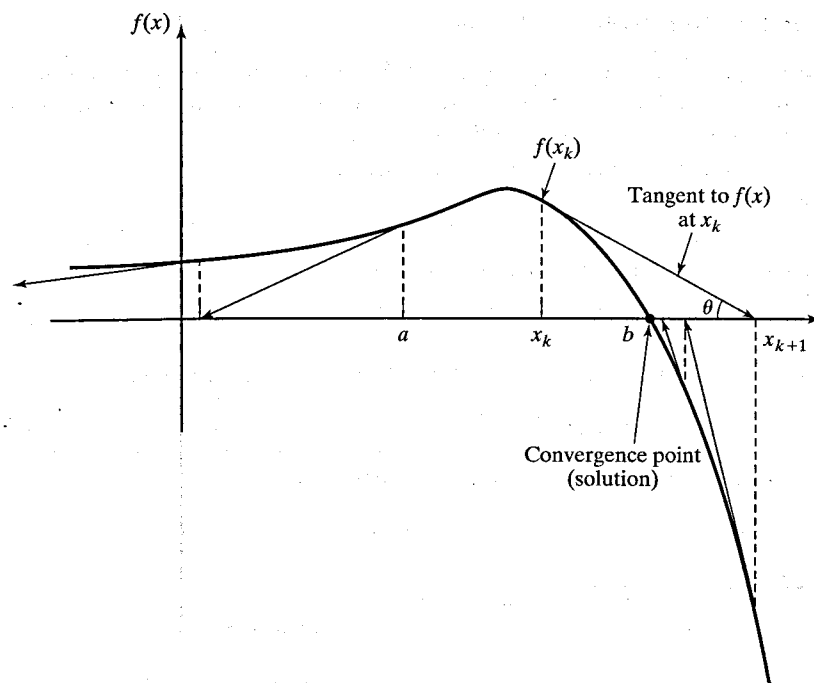


FIGURE 20.3
Illustration of the iterative process in the Newton-Raphson method

Thus, for the Newton-Raphson method, we have

$$f'(x) = 216x^2 - 468x + 241$$

$$x_{k+1} = x_k - \frac{72x^3 - 234x^2 + 241x - 78}{216x^2 - 468x + 241}$$

Starting with $x_0 = 10$, the following table provides the successive iterations:

k	x_k	$\frac{f(x_k)}{f'(x_k)}$	x_{k+1}
0	10.000000	2.978923	7.032108
1	7.032108	1.976429	5.055679
2	5.055679	1.314367	3.741312
3	3.741312	0.871358	2.869995
4	2.869995	0.573547	2.296405
5	2.296405	0.371252	1.925154
6	1.925154	0.230702	1.694452
7	1.694452	0.128999	1.565453
8	1.565453	0.054156	1.511296
9	1.511296	.0108641	1.500432
10	1.500432	.00043131	1.500001

The method converges to $x = 1.5$. Actually, $f(x)$ has three stationary points at $x = \frac{2}{3}$, $x = \frac{13}{12}$, and $x = \frac{3}{2}$. The remaining two points can be found by attempting different values for initial x_0 . In fact, $x_0 = .5$ and $x_0 = 1$ should yield the missing stationary points (try it!).

Excel Moment

Template *excelNewtonRaphson.xls* can be used to solve any single-variable equation. It requires entering $\frac{f(x)}{f'(x)}$ in cell C3. For Example 20.1-3, we enter

$$=(72*A3^3-234*A3^2+241*A3-78)/(216*A3^2-468*A3+241)$$

The variable x is replaced with A3. The template allows setting a tolerance limit Δ , which specifies the allowable difference between x_k and x_{k+1} that signals the termination of the iterations. You are encouraged to use different initial points, x_0 , to get a feel of how the method works.

PROBLEM SET 20.1B

1. Use *NewtonRaphson.xls* to solve Problem 1(c), Set 20.1a.
2. Solve Problem 2(b), Set 20.1a, by the Newton-Raphson method.

20.2 CONSTRAINED PROBLEMS

This section deals with the optimization of constrained continuous functions. Section 20.2.1 introduces the case of equality constraints, and Section 20.2.2 deals with inequality constraints. The presentation in Section 20.2.1 is covered for the most part in Beightler and Associates (1979, pp. 45–55).

20.2.1 Equality Constraints

This section presents two methods: the **Jacobian** and the **Lagrangian**. The **Lagrangian** method can be developed logically from the Jacobian. This relationship provides an interesting economic interpretation of the Lagrangian method.

Constrained derivatives (Jacobian) method. Consider the problem

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

subject to

$$g(\mathbf{X}) = 0$$

where

$$\mathbf{X} = (x_1, x_2, \dots, x_n)$$

$$\mathbf{g} = (g_1, g_2, \dots, g_m)^T$$

The functions $f(\mathbf{X})$ and $g(\mathbf{X})$, $i = 1, 2, \dots, m$, are twice continuously differentiable.

The idea of using constrained derivatives is to develop a closed-form expression for the first partial derivatives of $f(\mathbf{X})$ at all points satisfying $\mathbf{g}(\mathbf{X}) = \mathbf{0}$. The corresponding stationary points are identified as the points at which these partial derivatives vanish. The sufficiency conditions introduced in Section 20.1 can then be used to check the identity of stationary points.

To clarify the proposed concept, consider $f(x_1, x_2)$ illustrated in Figure 20.4. This function is to be minimized subject to the constraint

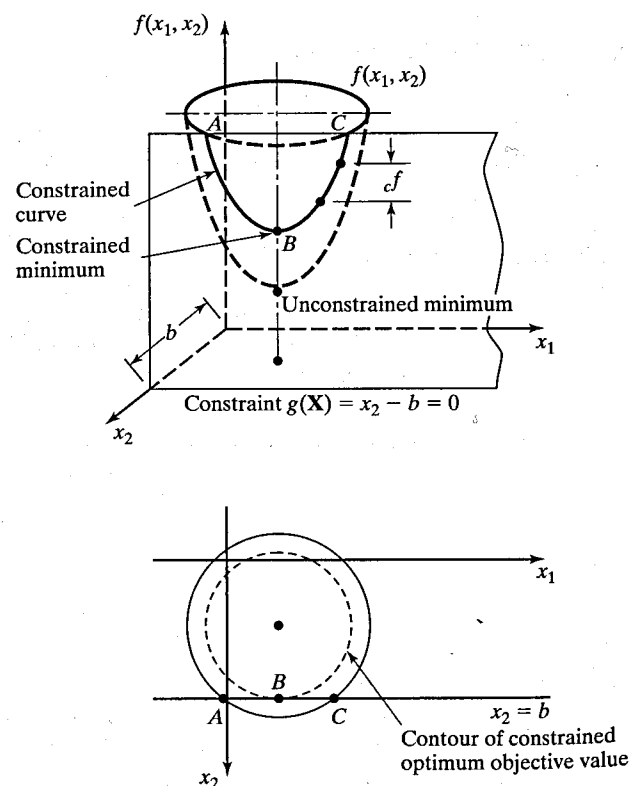
$$g_1(x_1, x_2) = x_2 - b = 0$$

where b is a constant. From Figure 20.4, the curve designated by the three points A , B , and C represents the values of $f(x_1, x_2)$ satisfying the given constraint. The constrained derivatives method defines the gradient of $f(x_1, x_2)$ at any point on the curve ABC . Point B at which the constrained derivative vanishes is a stationary point for the constrained problem.

The method is now developed mathematically. By Taylor's theorem, for $\mathbf{X} + \Delta\mathbf{X}$ in the feasible neighborhood of \mathbf{X} , we have

$$f(\mathbf{X} + \Delta\mathbf{X}) - f(\mathbf{X}) = \nabla f(\mathbf{X})\Delta\mathbf{X} + O(\Delta x_j^2)$$

FIGURE 20.4
Demonstration of the idea of the
Jacobian method



and

$$g(\mathbf{X} + \Delta\mathbf{X}) - g(\mathbf{X}) = \nabla g(\mathbf{X})\Delta\mathbf{X} + O(\Delta x_j^2)$$

As $\Delta x_j \rightarrow 0$, the equations reduce to

$$\partial f(\mathbf{X}) = \nabla f(\mathbf{X}) \partial\mathbf{X}$$

and

$$\partial g(\mathbf{X}) = \nabla g(\mathbf{X}) \partial\mathbf{X}$$

For feasibility, we must have $\mathbf{g}(\mathbf{X}) = \mathbf{0}$, $\partial g(\mathbf{X}) = \mathbf{0}$. Hence

$$\partial f(\mathbf{X}) - \nabla f(\mathbf{X}) \partial\mathbf{X} = 0$$

$$\nabla g(\mathbf{X}) \partial\mathbf{X} = 0$$

This gives $(m + 1)$ equations in $(n + 1)$ unknowns, $\partial f(\mathbf{X})$ and $\partial\mathbf{X}$. Note that $\partial f(\mathbf{X})$ is a dependent variable whose value is determined once $\partial\mathbf{X}$ is known. This means that, in effect, we have m equations in n unknowns.

If $m > n$, at least $(m - n)$ equations are redundant. Eliminating redundancy, the system reduces to $m \leq n$. If $m = n$, the solution is $\partial\mathbf{X} = \mathbf{0}$, and \mathbf{X} has no feasible neighborhood, which means that the solution space consists of one point only. The remaining case ($m < n$) requires further elaboration.

Define

$$\mathbf{X} = (\mathbf{Y}, \mathbf{Z})$$

such that

$$\mathbf{Y} = (y_1, y_2, \dots, y_m), \mathbf{Z} = (z_1, z_2, \dots, z_{n-m})$$

The vectors \mathbf{Y} and \mathbf{Z} represent the *dependent* and *independent* variables, respectively. Rewriting the gradient vectors of f and g in terms of \mathbf{Y} and \mathbf{Z} , we get

$$\nabla f(\mathbf{Y}, \mathbf{Z}) = (\nabla_{\mathbf{Y}} f, \nabla_{\mathbf{Z}} f)$$

$$\nabla g(\mathbf{Y}, \mathbf{Z}) = (\nabla_{\mathbf{Y}} g, \nabla_{\mathbf{Z}} g)$$

Define

$$\mathbf{J} = \nabla_{\mathbf{Y}} \mathbf{g} = \begin{pmatrix} \nabla_{\mathbf{Y}} g_1 \\ \vdots \\ \nabla_{\mathbf{Y}} g_m \end{pmatrix}$$

$$\mathbf{C} = \nabla_{\mathbf{Z}} \mathbf{g} = \begin{pmatrix} \nabla_{\mathbf{Z}} g_1 \\ \vdots \\ \nabla_{\mathbf{Z}} g_m \end{pmatrix}$$

$\mathbf{J}_{m \times m}$ is called the **Jacobian matrix** and $\mathbf{C}_{m \times n-m}$ the **control matrix**. The Jacobian \mathbf{J} is assumed nonsingular. This is always possible because the given m equations are independent by definition. The components of the vector \mathbf{Y} must thus be selected such that \mathbf{J} is nonsingular.

The original set of equations in $\partial f(\mathbf{X})$ and $\partial \mathbf{X}$ may be written as

$$\partial f(\mathbf{Y}, \mathbf{Z}) = \nabla_{\mathbf{Y}} f \partial \mathbf{Y} + \nabla_{\mathbf{Z}} f \partial \mathbf{Z}$$

and

$$\mathbf{J} \partial \mathbf{Y} = -\mathbf{C} \partial \mathbf{Z}$$

Given \mathbf{J} is nonsingular, it follows that

$$\partial \mathbf{Y} = -\mathbf{J}^{-1} \mathbf{C} \partial \mathbf{Z}$$

Substituting for $\partial \mathbf{Y}$ in the equation for $\partial f(\mathbf{X})$ gives ∂f as a function of $\partial \mathbf{Z}$ —that is,

$$\partial f(\mathbf{Y}, \mathbf{Z}) = (\nabla_{\mathbf{Z}} f - \nabla_{\mathbf{Y}} f \mathbf{J}^{-1} \mathbf{C}) \partial \mathbf{Z}$$

From this equation, the constrained derivative with respect to the independent vector \mathbf{Z} is given by

$$\nabla_c f = \frac{\partial_c f(\mathbf{Y}, \mathbf{Z})}{\partial_c \mathbf{Z}} = \nabla_{\mathbf{Z}} f - \nabla_{\mathbf{Y}} f \mathbf{J}^{-1} \mathbf{C}$$

where $\nabla_c f$ is the **constrained gradient** vector of f with respect to \mathbf{Z} . Thus, $\nabla_c f(\mathbf{Y}, \mathbf{Z})$ must be null at the stationary points.

The sufficiency conditions are similar to those developed in Section 20.1. The (constrained) Hessian matrix corresponds to the independent vector \mathbf{Z} , and the elements of the Hessian matrix must be the *constrained* second derivatives.

Example 20.2-1

Consider the following problem:

$$f(\mathbf{X}) = x_1^2 + 3x_2^2 + 5x_1x_3^2$$

$$g_1(\mathbf{X}) = x_1x_3 + 2x_2 + x_2^2 - 11 = 0$$

$$g_2(\mathbf{X}) = x_1^2 + 2x_1x_2 + x_3^2 - 14 = 0$$

Given the feasible point $\mathbf{X}^0 = (1, 2, 3)$, we wish to study the variation in $f (= \partial_c f)$ in the feasible neighborhood of \mathbf{X}^0 .

Let

$$\mathbf{Y} = (x_1, x_3) \quad \text{and} \quad \mathbf{Z} = x_2$$

Thus,

$$\nabla_{\mathbf{Y}} f = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_3} \right) = (2x_1 + 5x_3^2, 10x_1x_3)$$

$$\nabla_{\mathbf{Z}} f = \frac{\partial f}{\partial x_2} = 6x_2$$

$$\mathbf{J} = \begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_3} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_3} \end{pmatrix} = \begin{pmatrix} x_3 & x_1 \\ 2x_1 + 2x_2 & 2x_3 \end{pmatrix}$$

$$\mathbf{C} = \begin{pmatrix} \frac{\partial g_1}{\partial x_2} \\ \frac{\partial g_2}{\partial x_2} \end{pmatrix} = \begin{pmatrix} 2x_2 + 2 \\ 2x_2 \end{pmatrix}$$

Suppose that we need to estimate $\partial_c f$ in the feasible neighborhood of the feasible point $\mathbf{X}_0 = (1, 2, 3)$, given a small change $\partial x_2 = .01$ in the independent variable x_2 . We have

$$\mathbf{J}^{-1} \mathbf{C} = \begin{pmatrix} 3 & 1 \\ 6 & 6 \end{pmatrix}^{-1} \begin{pmatrix} 6 \\ 2 \end{pmatrix} = \begin{pmatrix} \frac{6}{12} & -\frac{1}{12} \\ -\frac{6}{12} & \frac{3}{12} \end{pmatrix} \begin{pmatrix} 6 \\ 2 \end{pmatrix} \approx \begin{pmatrix} 2.83 \\ -2.50 \end{pmatrix}$$

Hence, the incremental value of constrained f is given as

$$\partial_c f = (\nabla_{\mathbf{Z}} f - \nabla_{\mathbf{Y}} f \mathbf{J}^{-1} \mathbf{C}) \partial \mathbf{Z} = \left(6(2) - (47, 30) \begin{pmatrix} 2.83 \\ -2.50 \end{pmatrix} \right) \partial x_2 = -46.01 \partial x_2$$

By specifying the value of ∂x_2 for the *independent* variable x_2 , feasible values of ∂x_1 and ∂x_3 are determined for the dependent variables x_1 and x_3 using the formula

$$\partial \mathbf{Y} = -\mathbf{J}^{-1} \mathbf{C} \partial \mathbf{Z}$$

Thus, for $\partial x_2 = .01$,

$$\begin{pmatrix} \partial x_1 \\ \partial x_3 \end{pmatrix} = -\mathbf{J}^{-1} \mathbf{C} \partial x_2 = \begin{pmatrix} -.0283 \\ .0250 \end{pmatrix}$$

We now compare the value of $\partial_c f$ as computed above with the difference $f(\mathbf{X}_0 + \partial \mathbf{X}) - f(\mathbf{X}_0)$, given $\partial x_2 = .01$.

$$\mathbf{X}_0 + \partial \mathbf{X} = (1 - .0283, 2 + .01, 3 + .025) = (.9717, 2.01, 3.025)$$

This yields

$$f(\mathbf{X}_0) = 58, f(\mathbf{X}_0 + \partial \mathbf{X}) = 57.523$$

or

$$f(\mathbf{X}_0 + \partial \mathbf{X}) - f(\mathbf{X}_0) = -.477$$

The amount $-.477$ compares favorably with $\partial_c f = -46.01 \partial x_2 = -.4601$. The difference between the two values is the result of the linear approximation in computing $\partial_c f$ at \mathbf{X}_0 .

PROBLEM SET 20.2A

1. Consider Example 20.2-1.
 - (a) Compute $\partial_c f$ by the two methods presented in the example, using $\partial x_2 = .001$ instead of $\partial x_2 = .01$. Does the effect of linear approximation become more negligible with the decrease in the value of ∂x_2 ?
 - *(b) Specify a relationship among the elements of $\partial \mathbf{X} = (\partial x_1, \partial x_2, \partial x_3)$ at the feasible point $\mathbf{X}_0 = (1, 2, 3)$ that will keep the point $\mathbf{X}_0 + \partial \mathbf{X}$ feasible.
 - (c) If $\mathbf{Y} = (x_2, x_3)$ and $\mathbf{Z} = x_1$, what is the value of ∂x_1 that will produce the same value of $\partial_c f$ given in the example?

Example 20.2-2

This example illustrates the use of constrained derivatives. Consider the problem

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

subject to

$$g_1(\mathbf{X}) = x_1 + x_2 + 3x_3 - 2 = 0$$

$$g_2(\mathbf{X}) = 5x_1 + 2x_2 + x_3 - 5 = 0$$

We determine the constrained extreme points as follows. Let

$$\mathbf{Y} = (x_1, x_2) \text{ and } \mathbf{Z} = x_3$$

Thus,

$$\nabla_{\mathbf{Y}} f = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2} \right) = (2x_1, 2x_2), \quad \nabla_{\mathbf{Z}} f = \frac{\partial f}{\partial x_3} = 2x_3$$

$$\mathbf{J} = \begin{pmatrix} 1 & 1 \\ 5 & 2 \end{pmatrix}, \quad \mathbf{J}^{-1} = \begin{pmatrix} -\frac{2}{3} & \frac{1}{3} \\ \frac{5}{3} & -\frac{1}{3} \end{pmatrix}, \quad \mathbf{C} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$$

Hence,

$$\begin{aligned} \nabla_c f &= \frac{\partial_c f}{\partial_c x_3} = 2x_3 - (2x_1, 2x_2) \begin{pmatrix} -\frac{2}{3} & \frac{1}{3} \\ \frac{5}{3} & -\frac{1}{3} \end{pmatrix} \begin{pmatrix} 3 \\ 1 \end{pmatrix} \\ &= \frac{10}{3}x_1 - \frac{28}{3}x_2 + 2x_3 \end{aligned}$$

The equations for determining the stationary points are thus given as

$$\nabla_c f = 0$$

$$g_1(\mathbf{X}) = 0$$

$$g_2(\mathbf{X}) = 0$$

or

$$\begin{pmatrix} 10 & -28 & 6 \\ 1 & 1 & 3 \\ 5 & 2 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \\ 5 \end{pmatrix}$$

The solution is

$$\mathbf{X}_0 \approx (.81, .35, .28)$$

The identity of this stationary point is checked using the sufficiency condition. Given that x_3 is the independent variable, it follows from $\nabla_c f$ that

$$\frac{\partial_c^2 f}{\partial_c x_3^2} = \frac{10}{3} \left(\frac{dx_1}{dx_3} \right) - \frac{28}{3} \left(\frac{dx_2}{dx_3} \right) + 2 = \left(\frac{10}{3}, -\frac{28}{3} \right) \begin{pmatrix} \frac{dx_1}{dx_3} \\ \frac{dx_2}{dx_3} \end{pmatrix} + 2$$

From the Jacobian method,

$$\begin{pmatrix} \frac{dx_1}{dx_3} \\ \frac{dx_2}{dx_3} \end{pmatrix} = -\mathbf{J}^{-1}\mathbf{C} = \begin{pmatrix} \frac{5}{3} \\ -\frac{14}{3} \end{pmatrix}$$

Substitution gives $\frac{\partial_c^2 f}{\partial_c x_3^2} = \frac{460}{9} > 0$. Hence, \mathbf{X}_0 is the minimum point.

Sensitivity analysis in the Jacobian method. The Jacobian method can be used to study the effect of small changes in the right-hand side of the constraints on the optimal value of f . Specifically, what is the effect of changing $g_i(\mathbf{X}) = 0$ to $g_i(\mathbf{X}) = \partial g_i$ on the optimal value of f ? This type of investigation is called **sensitivity analysis** and is similar to that carried out in linear programming (see Chapters 3 and 4). However, sensitivity analysis in nonlinear programming is valid only in the small neighborhood of the extreme point. The development will be helpful in studying the Lagrangean method.

We have shown previously that

$$\partial f(\mathbf{Y}, \mathbf{Z}) = \nabla_{\mathbf{Y}} f \partial \mathbf{Y} + \nabla_{\mathbf{Z}} f \partial \mathbf{Z}$$

$$\partial \mathbf{g} = \mathbf{J} \partial \mathbf{Y} + \mathbf{C} \partial \mathbf{Z}$$

Given $\partial \mathbf{g} \neq 0$, then

$$\partial \mathbf{Y} = \mathbf{J}^{-1} \partial \mathbf{g} - \mathbf{J}^{-1} \mathbf{C} \partial \mathbf{Z}$$

Substituting in the equation for $\partial f(\mathbf{Y}, \mathbf{Z})$ gives

$$\partial f(\mathbf{Y}, \mathbf{Z}) = \nabla_{\mathbf{Y}} f \mathbf{J}^{-1} \partial \mathbf{g} + \nabla_{\mathbf{Z}} f \partial \mathbf{Z}$$

where

$$\nabla_c f = \nabla_{\mathbf{Z}} f - \nabla_{\mathbf{Y}} f \mathbf{J}^{-1} \mathbf{C}$$

as defined previously. The expression for $\partial f(\mathbf{Y}, \mathbf{Z})$ can be used to study variation in f in the feasible neighborhood of a feasible point \mathbf{X}_0 resulting from small changes $\partial \mathbf{g}$ and $\partial \mathbf{Z}$.

At the extreme (indeed, any stationary) point $\mathbf{X}_0 = (\mathbf{Y}_0, \mathbf{Z}_0)$, the constrained gradient $\nabla_c f$ must vanish. Thus

$$\partial f(\mathbf{Y}_0, \mathbf{Z}_0) = \nabla_{\mathbf{Y}_0} f \mathbf{J}^{-1} \partial \mathbf{g}(\mathbf{Y}_0, \mathbf{Z}_0)$$

or

$$\frac{\partial f}{\partial \mathbf{g}} = \nabla_{\mathbf{Y}_0} f \mathbf{J}^{-1}$$

The effect of the small change ∂g on the *optimum* value of f can be studied by evaluating the rate of change of f with respect to g . These rates are usually referred to as **sensitivity coefficients**.

Example 20.2-3

Consider the same problem of Example 20.2-2. The optimum point is given by $\mathbf{X}_0 = (x_{01}, x_{02}, x_{03}) = (.81, .35, .28)$. Given $\mathbf{Y}_0 = (x_{01}, x_{02})$, then

$$\nabla_{\mathbf{Y}_0} f = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2} \right) = (2x_{01}, 2x_{02}) = (1.62, .70)$$

Consequently,

$$\left(\frac{\partial f}{\partial g_1}, \frac{\partial f}{\partial g_2} \right) = \nabla_{\mathbf{Y}_0} f \mathbf{J}^{-1} = (1.62, .7) \begin{pmatrix} -\frac{2}{3} & \frac{1}{3} \\ \frac{5}{3} & -\frac{1}{3} \end{pmatrix} = (.0876, .3067)$$

This means that for $\partial g_1 = 1$, f will increase *approximately* by .0876. Similarly, for $\partial g_2 = 1$, f will increase *approximately* by .3067.

PROBLEM SET 20.2B

- Suppose that Example 20.2-2 is solved in the following manner. First, use the constraints to express x_1 and x_2 in terms of x_3 ; then use the resulting equations to express the objective function in terms of x_3 only. By taking the derivative of the new objective function with respect to x_3 , we can determine the points of maxima and minima.
 - Would the derivative of the new objective function (expressed in terms of x_3) be different from that obtained by the Jacobian method?
 - How does the suggested procedure differ from the Jacobian method?
- Apply the Jacobian method to Example 20.2-1 by selecting $\mathbf{Y} = (x_2, x_3)$ and $\mathbf{Z} = (x_1)$.
- Solve by the Jacobian method:

$$\text{Minimize } f(\mathbf{X}) = \sum_{i=1}^n x_i^2$$

subject to

$$\prod_{i=1}^n x_i = C$$

where C is a positive constant. Suppose that the right-hand side of the constraint is changed to $C + \delta$, where δ is a small positive quantity. Find the corresponding change in the optimal value of f .

- Solve by the Jacobian method:

$$\text{Minimize } f(\mathbf{X}) = 5x_1^2 + x_2^2 + 2x_1x_2$$

subject to

$$g(\mathbf{X}) = x_1x_2 - 10 = 0$$

- Find the change in the optimal value of $f(\mathbf{X})$ if the constraint is replaced by $x_1x_2 - 9.99 = 0$.
 - Find the change in value of $f(\mathbf{X})$ in the neighborhood of the feasible point $(2, 5)$, given that $x_1x_2 = 9.99$ and $\partial x_1 = .01$.
5. Consider the problem:

$$\text{Maximize } f(\mathbf{X}) = x_1^2 + 2x_2^2 + 10x_3^2 + 5x_1x_2$$

subject to

$$g_1(\mathbf{X}) = x_1 + x_2^2 + 3x_2x_3 - 5 = 0$$

$$g_2(\mathbf{X}) = x_1^2 + 5x_1x_2 + x_3^2 - 7 = 0$$

Apply the Jacobian method to find $\partial f(\mathbf{X})$ in the neighborhood of the feasible point $(1, 1, 1)$. Assume that this neighborhood is specified by $\partial g_1 = -.01$, $\partial g_2 = .02$, and $\partial x_1 = .01$.

6. Consider the problem

$$\text{Minimize } f(\mathbf{X}) = x_1^2 + x_2^2 + x_3^2 + x_4^2$$

subject to

$$g_1(\mathbf{X}) = x_1 + 2x_2 + 3x_3 + 5x_4 - 10 = 0$$

$$g_2(\mathbf{X}) = x_1 + 2x_2 + 5x_3 + 6x_4 - 15 = 0$$

- Show that by selecting x_3 and x_4 as independent variables, the Jacobian method fails to provide a solution and, state the reason.
- *Solve the problem using x_1 and x_3 as independent variables, and apply the sufficiency condition to determine the type of the resulting stationary point.
- Determine the sensitivity coefficients, given the solution in (b).

Lagrangean method. In the Jacobian method, let the vector $\boldsymbol{\lambda}$ represent the sensitivity coefficients—that is

$$\boldsymbol{\lambda} = \nabla_{\mathbf{Y}_0} \mathbf{J}^{-1} = \frac{\partial f}{\partial \mathbf{g}}$$

Thus,

$$\partial f - \boldsymbol{\lambda} \partial \mathbf{g} = 0$$

This equation satisfies the necessary conditions for stationary points because $\frac{\partial f}{\partial \mathbf{g}}$ is computed such that $\nabla_{\mathbf{X}} f = \mathbf{0}$. A more convenient form for presenting these equations is to take their partial derivatives with respect to all x_j . This yields

$$\frac{\partial}{\partial x_j} (f - \boldsymbol{\lambda} \mathbf{g}) = 0, \quad j = 1, 2, \dots, n$$

The resulting equations together with the constraint equations $\mathbf{g}(\mathbf{X}) = \mathbf{0}$ yield the feasible values of \mathbf{X} and $\boldsymbol{\lambda}$ that satisfy the *necessary* conditions for stationary points.

The given procedure defines the *Lagrangian method* for identifying the stationary points of optimization problems with *equality* constraints. Let

$$L(\mathbf{X}, \lambda) = f(\mathbf{X}) - \lambda g(\mathbf{X})$$

The function L is called the **Lagrangian function** and the elements of the vector λ constitute the **Lagrange multipliers**. By definition, these multipliers have the same interpretation as the sensitivity coefficients of the Jacobian method

The equations

$$\frac{\partial L}{\partial \lambda} = 0, \frac{\partial L}{\partial \mathbf{X}} = 0$$

give the necessary conditions for determining stationary points of $f(\mathbf{X})$ subject to $g(\mathbf{X}) = 0$. Sufficiency conditions for the Lagrangean method exist, but they are generally computationally difficult.

Example 20.2-4

Consider the problem of Example 20.2-2. The Lagrangean function is

$$L(\mathbf{X}, \lambda) = x_1^2 + x_2^2 + x_3^2 - \lambda_1(x_1 + x_2 + 3x_3 - 2) - \lambda_2(5x_1 + 2x_2 + x_3 - 5)$$

This yields the following necessary conditions:

$$\frac{\partial L}{\partial x_1} = 2x_1 - \lambda_1 - 5\lambda_2 = 0$$

$$\frac{\partial L}{\partial x_2} = 2x_2 - \lambda_1 - 2\lambda_2 = 0$$

$$\frac{\partial L}{\partial x_3} = 2x_3 - 3\lambda_1 - \lambda_2 = 0$$

$$\frac{\partial L}{\partial \lambda_1} = -(x_1 + x_2 + 3x_3 - 2) = 0$$

$$\frac{\partial L}{\partial \lambda_2} = -(5x_1 + 2x_2 + x_3 - 5) = 0$$

The solution to these simultaneous equations yields

$$\mathbf{X}_0 = (x_1, x_2, x_3) = (.8043, .3478, .2826)$$

$$\lambda = (\lambda_1, \lambda_2) = (.0870, .3043)$$

This solution combines the results of Examples 20.2-2 and 20.2-3. The values of the Lagrange multipliers, as given by the vector λ , equal the sensitivity coefficients obtained in Example 20.2-3. The result shows that these coefficients are independent of the specific choice of the dependent vector \mathbf{Y} in the Jacobian method.

PROBLEM SET 20.2C

1. Solve the following linear programming problem by both the Jacobian and the Lagrangean methods:

$$\text{Maximize } f(\mathbf{X}) = 5x_1 + 3x_2$$

subject to

$$g_1(\mathbf{X}) = x_1 + 2x_2 + x_3 - 6 = 0$$

$$g_2(\mathbf{X}) = 3x_1 + x_2 + x_4 - 9 = 0$$

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

- *2. Find the optimal solution to the problem

$$\text{Minimize } f(\mathbf{X}) = x_1^2 + 2x_2^2 + 10x_3^2$$

subject to

$$g_1(\mathbf{X}) = x_1 + x_2^2 + x_3 - 5 = 0$$

$$g_2(\mathbf{X}) = x_1 + 5x_2 + x_3 - 7 = 0$$

Suppose that $g_1(\mathbf{X}) = .01$ and $g_2(\mathbf{X}) = .02$. Find the corresponding change in the optimal value of $f(\mathbf{X})$.

3. Solve Problem 6, Set 20.2b, by the Lagrangean method, and verify that the values of the Lagrange multipliers are the same as the sensitivity coefficients obtained in Problem 6, Set 20.2b.

20.2.2 Inequality Constraints—Karush-Kuhn-Tucker (KKT) Conditions¹

This section extends the Lagrangean method to problems with inequality constraints. The main contribution of the section is the development of the general Karush-Kuhn-Tucker (KKT) *necessary* conditions for determining the stationary points. These conditions are also sufficient under certain rules that will be stated later.

Consider the problem

$$\text{Maximize } z = f(\mathbf{X})$$

subject to

$$g(\mathbf{X}) \leq 0$$

The inequality constraints may be converted into equations by using *nonnegative* slack variables. Let $S_i^2 (\geq 0)$ be the slack quantity added to the i th constraint $g_i(\mathbf{X}) \leq 0$ and define

$$\mathbf{S} = (S_1, S_2, \dots, S_m)^T, \mathbf{S}^2 = (S_1^2, S_2^2, \dots, S_m^2)^T$$

¹W. Karush was the first to develop the KKT conditions in 1939 as part of an M.S. thesis at the University of Chicago. The same conditions were developed independently in 1951 by W. Kuhn and A. Tucker.

where m is the total number of inequality constraints. The Lagrangean function is thus given by

$$L(\mathbf{X}, \mathbf{S}, \boldsymbol{\lambda}) = f(\mathbf{X}) - \boldsymbol{\lambda} [\mathbf{g}(\mathbf{X}) + \mathbf{S}^2]$$

Given the constraints $\mathbf{g}(\mathbf{X}) \leq \mathbf{0}$, a necessary condition for optimality is that $\boldsymbol{\lambda}$ be non-negative (nonpositive) for maximization (minimization) problems. This result is justified by noting that the vector $\boldsymbol{\lambda}$ measures the rate of variation of f with respect to \mathbf{g} —that is,

$$\boldsymbol{\lambda} = \frac{\partial f}{\partial \mathbf{g}}$$

In the maximization case, as the right-hand side of the constraint $\mathbf{g}(\mathbf{X}) \leq \mathbf{0}$ increases from $\mathbf{0}$ to the vector $\partial \mathbf{g}$, the solution space becomes less constrained and hence f cannot decrease, meaning that $\boldsymbol{\lambda} \geq \mathbf{0}$. Similarly for minimization, as the right-hand side of the constraints increases, f cannot increase, which implies that $\boldsymbol{\lambda} \leq \mathbf{0}$. If the constraints are equalities, that is, $\mathbf{g}(\mathbf{X}) = \mathbf{0}$, then $\boldsymbol{\lambda}$ becomes unrestricted in sign (see Problem 2, Set 20.2d).

The restrictions on $\boldsymbol{\lambda}$ hold as part of the KKT necessary conditions. The remaining conditions will now be developed.

Taking the partial derivatives of L with respect to \mathbf{X} , \mathbf{S} , and $\boldsymbol{\lambda}$, we obtain

$$\frac{\partial L}{\partial \mathbf{X}} = \nabla f(\mathbf{X}) - \boldsymbol{\lambda} \nabla \mathbf{g}(\mathbf{X}) = \mathbf{0}$$

$$\frac{\partial L}{\partial S_i} = -2\lambda_i S_i = 0, \quad i = 1, 2, \dots, m$$

$$\frac{\partial L}{\partial \boldsymbol{\lambda}} = -(\mathbf{g}(\mathbf{X}) + \mathbf{S}^2) = \mathbf{0}$$

The second set of equations reveals the following results:

1. If $\lambda_i \neq 0$, then $S_i^2 = 0$. This result means that the corresponding resource is scarce (i.e., consumed completely).
2. If $S_i^2 > 0$, then $\lambda_i = 0$. This means resource i is not scarce and, hence, it has no effect on the value of f (i.e., $\lambda_i = \frac{\partial f}{\partial g_i} = 0$).

From the second and third sets of equations, we obtain

$$\lambda_i g_i(\mathbf{X}) = 0, \quad i = 1, 2, \dots, m$$

This new condition essentially repeats the foregoing argument, because if $\lambda_i > 0$, $g_i(\mathbf{X}) = 0$ or $S_i^2 = 0$; and if $g_i(\mathbf{X}) < 0$, $S_i^2 > 0$, and $\lambda_i = 0$.

The KKT necessary conditions for maximization problem are summarized as follows:

$$\boldsymbol{\lambda} \geq \mathbf{0}$$

$$\nabla f(\mathbf{X}) - \boldsymbol{\lambda} \nabla \mathbf{g}(\mathbf{X}) = \mathbf{0}$$

$$\lambda_i g_i(\mathbf{X}) = 0, \quad i = 1, 2, \dots, m$$

$$\mathbf{g}(\mathbf{X}) \leq \mathbf{0}$$

TABLE 20.1 Sufficiency of the KKT Conditions

Sense of optimization	Required conditions	
	Objective function	Solution space
Maximization	Concave	Convex set
Minimization	Convex	Convex set

These conditions apply to the minimization case as well, except that $\boldsymbol{\lambda}$ must be non-positive (verify!). In both maximization and minimization, the Lagrange multipliers corresponding to equality constraints are unrestricted in sign.

Sufficiency of the KKT conditions. The KKT necessary conditions are also sufficient if the objective function and the solution space satisfy the conditions in Table 20.1.

It is simpler to verify that a function is convex or concave than to prove that a solution space is a convex set. For this reason, we provide a *subset* of sufficiency conditions, which, though not as general as the ones in Table 20.1, are easier to apply in practice. To provide these conditions, we define the generalized nonlinear problems as

$$\text{Maximize or minimize } z = f(\mathbf{X})$$

subject to

$$g_i(\mathbf{X}) \leq 0, \quad i = 1, 2, \dots, r$$

$$g_i(\mathbf{X}) \geq 0, \quad i = r + 1, \dots, p$$

$$g_i(\mathbf{X}) = 0, \quad i = p + 1, \dots, m$$

$$L(\mathbf{X}, \mathbf{S}, \boldsymbol{\lambda}) = f(\mathbf{X}) - \sum_{i=1}^r \lambda_i [g_i(\mathbf{X}) + S_i^2] - \sum_{i=r+1}^p \lambda_i [g_i(\mathbf{X}) - S_i^2] - \sum_{i=p+1}^m \lambda_i g_i(\mathbf{X})$$

The parameter λ_i is the Lagrange multiplier associated with constraint i . The conditions for establishing the sufficiency of the KKT conditions are summarized in Table 20.2.

TABLE 20.2 Subset of KKT Sufficient Conditions

Sense of optimization	Required conditions			
	$f(\mathbf{X})$	$g_i(\mathbf{X})$	λ_i	
Maximization	Concave	Convex	≥ 0	$(1 \leq i \leq r)$
		Concave	≤ 0	$(r + 1 \leq i \leq p)$
		Linear	Unrestricted	$(p + 1 \leq i \leq m)$
Minimization	Convex	Convex	≤ 0	$(1 \leq i \leq r)$
		Concave	≥ 0	$(r + 1 \leq i \leq p)$
		Linear	Unrestricted	$(p + 1 \leq i \leq m)$

The conditions in Table 20.2 are a subset of the conditions in Table 20.1 because a solution space can be convex without satisfying the conditions in Table 20.2.

Table 20.2 is valid because the given conditions yield a concave Lagrangean function $L(\mathbf{X}, \mathbf{S}, \boldsymbol{\lambda})$ in case of maximization and a convex $L(\mathbf{X}, \mathbf{S}, \boldsymbol{\lambda})$ in case of minimization. This result is verified by noticing that if $g_i(x)$ is convex, then $\lambda_i g_i(x)$ is convex if $\lambda_i \geq 0$ and concave if $\lambda_i \leq 0$. Similar interpretations can be established for all the remaining conditions. Observe that a linear function is both convex and concave. Also, if a function f is concave, then $(-f)$ is convex, and vice versa.

Example 20.2-5

Consider the following minimization problem:

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

subject to

$$g_1(\mathbf{X}) = 2x_1 + x_2 - 5 \leq 0$$

$$g_2(\mathbf{X}) = x_1 + x_3 - 2 \leq 0$$

$$g_3(\mathbf{X}) = 1 - x_1 \leq 0$$

$$g_4(\mathbf{X}) = 2 - x_2 \leq 0$$

$$g_5(\mathbf{X}) = -x_3 \leq 0$$

This is a minimization problem, hence $\boldsymbol{\lambda} \leq 0$. The KKT conditions are thus given as

$$(\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5) \leq 0$$

$$(2x_1, 2x_2, 2x_3) - (\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5) \begin{pmatrix} 2 & 1 & 0 \\ 1 & 0 & 1 \\ -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} = 0$$

$$\lambda_1 g_1 = \lambda_2 g_2 = \dots = \lambda_5 g_5 = 0$$

$$g(\mathbf{X}) \leq 0$$

These conditions reduce to

$$\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5 \leq 0$$

$$2x_1 - 2\lambda_1 - \lambda_2 + \lambda_3 = 0$$

$$2x_2 - \lambda_1 + \lambda_4 = 0$$

$$2x_3 - \lambda_2 + \lambda_5 = 0$$

$$\lambda_1(2x_1 + x_2 - 5) = 0$$

$$\lambda_2(x_1 + x_3 - 2) = 0$$

$$\lambda_3(1 - x_1) = 0$$

$$\lambda_4(2 - x_2) = 0$$

$$\lambda_5 x_3 = 0$$

$$2x_1 + x_2 \leq 5$$

$$x_1 + x_3 \leq 2$$

$$x_1 \geq 1, x_2 \geq 2, x_3 \geq 0$$

The solution is $x_1 = 1, x_2 = 2, x_3 = 0, \lambda_1 = \lambda_2 = \lambda_5 = 0, \lambda_3 = -2, \lambda_4 = -4$. Because both $f(\mathbf{X})$ and the solution space $g(\mathbf{X}) \leq 0$ are convex, $L(\mathbf{X}, \mathbf{S}, \boldsymbol{\lambda})$ must be convex, and the resulting stationary point yields a global constrained minimum. The KKT conditions are central to the development of the nonlinear programming algorithms in Chapter 21.

PROBLEM SET 20.2D

1. Consider the problem:

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

subject to

$$g(\mathbf{X}) \geq 0$$

Show that the KKT conditions are the same as in Section 20.2.2, except that $\boldsymbol{\lambda} \leq 0$.

2. Consider the following problem:

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

subject to

$$g(\mathbf{X}) = 0$$

Show that the KKT conditions are

$$\nabla f(\mathbf{X}) - \boldsymbol{\lambda} \nabla g(\mathbf{X}) = 0$$

$$g(\mathbf{X}) = 0$$

$\boldsymbol{\lambda}$ unrestricted in sign

3. Write the KKT necessary conditions for the following problems.

(a) Maximize $f(\mathbf{X}) = x_1^3 - x_2^2 + x_1 x_3^2$

subject to

$$x_1 + x_2^2 + x_3 = 5$$

$$5x_1^2 - x_2^2 - x_3 \geq 2$$

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

(b) Minimize $f(\mathbf{X}) = x_1^4 + x_2^2 + 5x_1 x_2 x_3$

subject to

$$x_1^2 - x_2^2 + x_3^3 \leq 10$$

$$x_1^3 + x_2^2 + 4x_3^2 \geq 20$$

4. Consider the problem

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

subject to

$$g(\mathbf{X}) = 0$$

Given $f(\mathbf{X})$ is concave and $g_i(\mathbf{X}) (i = 1, 2, \dots, m)$ is a linear function, show that the KKT necessary conditions are also sufficient. Is this result true if $g_i(\mathbf{X})$ is a convex nonlinear function for all i ? Why?

5. Consider the problem

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

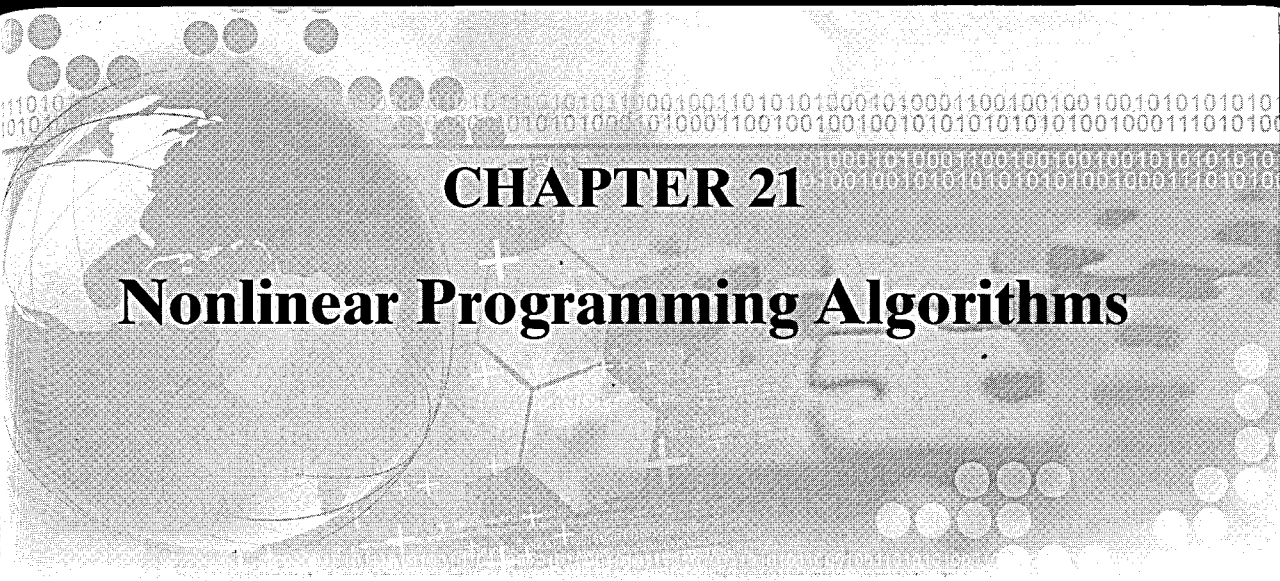
subject to

$$g_1(\mathbf{X}) \geq 0, g_2(\mathbf{X}) = 0, g_3(\mathbf{X}) \leq 0$$

Develop the KKT conditions, and give the stipulations under which the conditions are sufficient.

BIBLIOGRAPHY

Bazarra, M., H. Sherali, and C. Shetty, *Nonlinear Programming Theory and Algorithms*, 3rd ed., Wiley, New York, 2006.
Beightler, C., D. Phillips, and D. Wilde, *Foundations of Optimization*, 2nd ed., Prentice Hall, NJ, 1979.
Fletcher, R., *Practical Methods of Optimization*, 2nd ed., Wiley, New York, 2000.



CHAPTER 21

Nonlinear Programming Algorithms

21.1 UNCONSTRAINED ALGORITHMS

This section presents two types of algorithms for the unconstrained problem: *direct search* and *gradient*.

21.1.1 Direct Search Method

Direct search methods apply primarily to strictly unimodal single-variable functions. Although the case may appear trivial, Section 21.1.2 shows that optimization of single-variable functions is key in the development of the more general multivariable algorithm.

The idea of direct search methods is to identify the **interval of uncertainty** known to include the optimum solution point. The procedure locates the optimum by iteratively narrowing the interval of uncertainty to a desired level of accuracy.

Two closely related search algorithms are presented in this section: **dichotomous** and **golden section**. Both algorithms seek the maximization of a unimodal function $f(x)$ over the interval $a \leq x \leq b$ that includes the optimum point x^* . The two methods start with the initial interval of uncertainty $I_0 = (a, b)$.

General step i . Let $I_{i-1} = (x_L, x_R)$ be the current interval of uncertainty (at iteration 0, $x_L = a$ and $x_R = b$). The following table shows how x_1 and x_2 are determined:

Dichotomous method	Golden section method
$x_1 = \frac{1}{2}(x_R + x_L - \Delta)$	$x_1 = x_R - \left(\frac{\sqrt{5}-1}{2}\right)(x_R - x_L)$
$x_2 = \frac{1}{2}(x_R + x_L + \Delta)$	$x_2 = x_L + \left(\frac{\sqrt{5}-1}{2}\right)(x_R - x_L)$

The selection of x_1 and x_2 guarantees that $x_L < x_1 < x_2 < x_R$.