4

BASIC CONCEPTS OF OPTIMIZATION

To UNDERSTAND THE strategy of optimization procedures, certain basic concepts must be described. In this chapter we examine the properties of objective functions and constraints to establish a basis for analyzing optimization problems. We identify those features that are desirable (and also undesirable) in the formulation of an optimization problem. Both qualitative and quantitative characteristics of functions are described. In addition, we present the necessary and sufficient conditions to guarantee that a supposed extremum is indeed a minimum or a maximum.

4.1   CONTINUITY OF FUNCTIONS

In carrying out analytical or numerical optimization you will find it preferable and more convenient to work with continuous functions of one or more variables than with functions containing discontinuities. Functions having continuous derivatives are also preferred. Case A in Figure 4.1 shows a discontinuous function. Is case B also discontinuous?

We define the property of continuity as follows. A function of a single variable x is continuous at a point x_0 if

$$f(x_0) \text{ exists}$$

$$\lim_{x \to x_0} f(x) \text{ exists}$$

$$\lim_{x \to x_0} f(x) = f(x_0)$$

If $f(x)$ is continuous at every point in region R, then $f(x)$ is said to be continuous throughout R. For case B in Figure 4.1, the function of x has a "kink" in it, but $f(x)$ does satisfy the property of continuity. However, $f'(x) \equiv df(x)/dx$ does not. Therefore, the function in case B is continuous but not continuously differentiable.
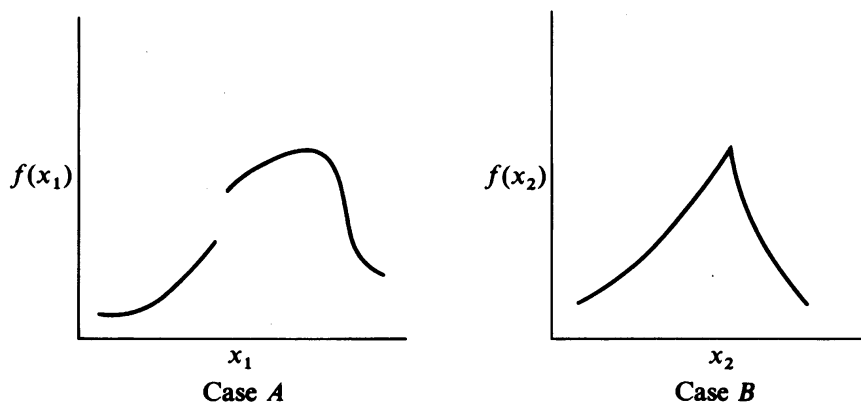
FIGURE 4.1
Functions with discontinuities in the function or derivatives.

EXAMPLE 4.1 ANALYSIS OF FUNCTIONS FOR CONTINUITY

Are the following functions continuous? (a) $f(x) = 1/x$; (b) $f(x) = \ln x$. In each case specify the range of x for which $f(x)$ and $f'(x)$ are continuous.

Solution

(a) $f(x) = 1/x$ is continuous except at $x = 0$; $f(0)$ is not defined. $f'(x) = -1/x^2$ is continuous except at $x = 0$.

(b) $f(x) = \ln x$ is continuous for $x > 0$. For $x \leq 0$, $\ln (x)$ is not defined. As to $f'(x) = 1/x$, see (a).

A discontinuity in a function may or may not cause difficulty in optimization. In case A in Figure 4.1, the maximum occurs reasonably far from the discontinuity which may or may not be encountered in the search for the optimum. In case B, if a method of optimization that does not use derivatives is employed, then the "kink" in $f(x)$ is probably unimportant, but methods employing derivatives might fail, because the derivative becomes undefined at the discontinuity and has different signs on each side of it. Hence a search technique approaches the optimum, but then oscillates about it rather than converges to it.

Objective functions that allow only discrete values of the independent variable(s) occur frequently in process design because the process variables assume only specific values rather than *continuous* ones. Examples are the cost per unit diameter of pipe, the cost per unit area for heat exchanger surface, or the insulation cost considered in Example 1.1. For a pipe, we might represent the installed cost as a function of the pipe diameter as shown in Figure 4.2 [see also Noltie (1978)]. For

Commercially available
pipe diameters
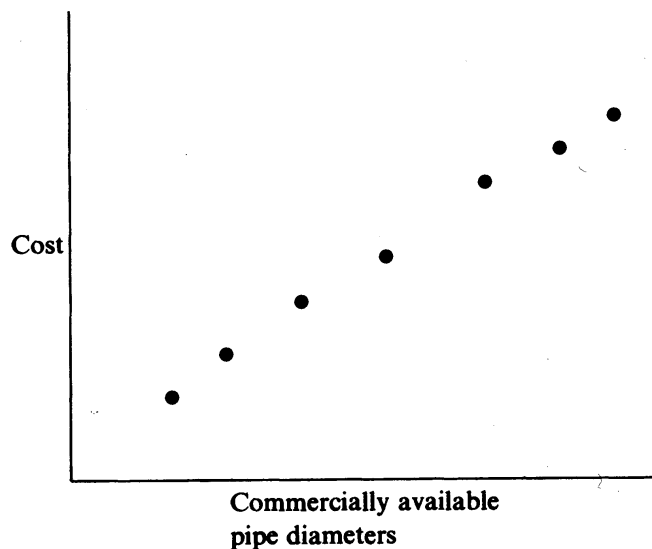
FIGURE 4.2
Installed pipe cost as a function of diameter.
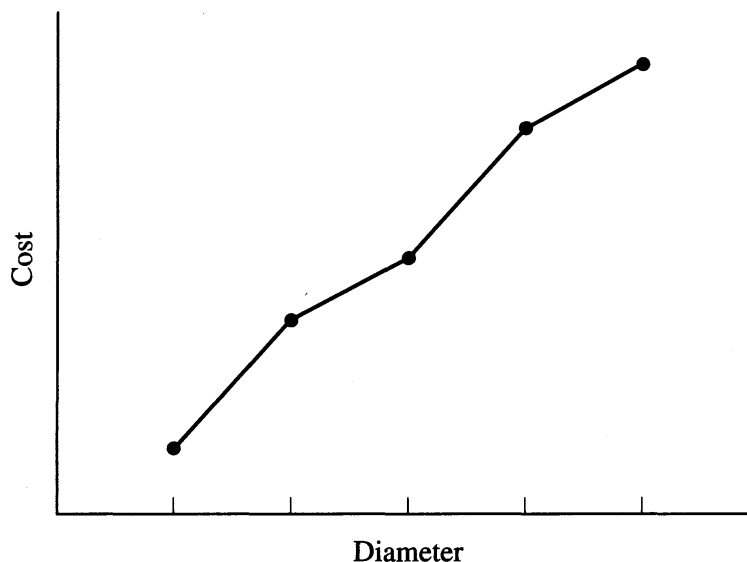
FIGURE 4.3
Piecewise linear approximation to cost function.

most purposes such a cost function can be approximated as a continuous function because of the relatively small differences in available pipe diameters. You can then disregard the discrete nature of the function and optimize the cost as if the diameter were a continuous variable. For example, extend the function of Figure 4.2 to a continuous range of diameters by interpolation. If linear interpolation is used, then the extended function usually has discontinuous derivatives at each of the original diameters, as shown in Figure 4.3. As mentioned earlier, this step can cause problems for derivative-based optimizers. A remedy is to interpolate with quadratic or cubic functions chosen so that their first derivatives are continuous at the break points. Such functions are called splines (Bartela et al., 1987). Once the optimum value of the diameter is obtained for the continuous function, the discretely valued diameter nearest to the optimum that is commercially available can be selected. A suboptimal value for installed cost results, but such a solution should be adequate for engineering purposes because of the narrow intervals between discrete values of the diameter.

EXAMPLE 4.2   OPTIMIZATION INVOLVING AN INTEGER-VALUED VARIABLE

Consider a catalytic regeneration cycle in which there is a simple trade-off between costs incurred during regeneration and the increased revenues due to the regenerated catalyst. Let x_1 be the number of days during which the catalyst is used in the reactor and x_2 be the number of days for regeneration. The reactor start-up crew is only available in the morning shift, so $x_1 + x_2$ must be an integer.

We assume that the reactor feed flow rate q (kg/day) is constant as is the cost of the feed C_1 ($/kg), the value of the product C_2 ($/kg), and the regeneration cost C_3

($/regeneration cycle). We further assume that the catalyst deteriorates gradually according to the linear relation

$$d = 1.0 - kx_1$$

where 1.0 represents the weight fraction conversion of feed at the start of the operating cycle, and k is the deterioration factor in units of weight fraction per day. Define an objective function and find the optimal value of x_1.

Solution. For one complete cycle of operation and regeneration, the objective function for the total profit per day comprises

$$\frac{Profit}{Day} = Product\ value - Feed\ cost$$

$$- (Regeneration\ cost\ per\ cycle) \cdot (Cycles\ per\ day)$$

or in the defined notation

$$f(\mathbf{x}) = \frac{qC_2 x_1 d_{avg} - qC_1 x_1 - C_3}{x_1 + x_2} \qquad (a)$$

where $d_{avg} = 1.0 - (kx_1/2)$.

The maximum daily profit for an entire cycle is obtained by maximizing Equation (a) with respect to x_1. As a first trial, we allow x_1 to be a continuous variable. When the first derivative of Equation (a) is set equal to zero and the resulting equation solved for x_1, the optimum is

$$x_1^{opt} = -x_2 + \left[x_2^2 + \left(\frac{2}{k}\right)\left(x_2 - \frac{C_1 x_2}{C_2} + \frac{C_3}{qC_2}\right)\right]^{1/2}$$

Suppose $x_2 = 2$, $k_1 = 0.02$, $q = 1000$, $C_2 = 1.0$, $C_1 = 0.4$, and $C_3 = 1000$. Then $x_1^{opt} = 12.97$ (rounded to 13 days if x_1 is an integer).

Clearly, treating x_1 as a continuous variable may be improper if x_1 is 1, 2, 3, and so on, but is probably satisfactory if x_1 is 15, 16, 17, and so on. You might specify x_1 in terms of shifts of 4–8 h instead of days to obtain finer subdivisions of time.

In real life, other problems involving discrete variables may not be so nicely posed. For example, if cost is a function of the number of discrete pieces of equipment, such as compressors, the optimization procedure cannot ignore the integer character of the cost function because usually only a small number of pieces of equipment are involved. You cannot install 1.54 compressors, and rounding off to 1 or 2 compressors may be quite unsatisfactory. This subject will be discussed in more detail in Chapter 9.

4.2 NLP PROBLEM STATEMENT

A general form for a nonlinear program (NLP) is

Minimize: $f(\mathbf{x})$

Subject to: $a_i \leq g_i(\mathbf{x}) \leq b_i$ $i = 1, \ldots, m$ (4.1)

and $l_j \leq x_j \leq u_j$ $j = 1, \ldots, n$

In this problem statement, $\mathbf{x}$ is a vector of n decision variables $(x_1, \ldots, x_n)$, f is the objective function, and the g_i are constraint functions. The a_i and b_i are specified lower and upper bounds on the constraint functions with $a_i \leq b_i$, and l_j, u_j are lower and upper bounds on the variables with $l_j \leq u_j$. If $a_i = b_i$, the ith constraint is an equality constraint. If the upper and lower limits on g_i correspond to $a_i = -\infty$ and $b_i = +\infty$, the constraint is unbounded. Similar comments apply to the variable bounds, with $l_j = u_j$ corresponding to a variable x_j whose value is fixed, and $l_j = -\infty$ and $u_j = +\infty$ specifying a free variable.

Problem 4.1 is nonlinear if one or more of the functions $f, g_1, \ldots, g_m$ are non-linear. It is *unconstrained* if there are no constraint functions g_i and no bounds on the x_i, and it is *bound-constrained* if only the x_i are bounded. In *linearly constrained* problems all constraint functions g_i are linear, and the objective f is nonlinear. There are special NLP algorithms and software for unconstrained and bound-constrained problems, and we describe these in Chapters 6 and 8. Methods and software for solving constrained NLPs use many ideas from the unconstrained case. Most modern software can handle nonlinear constraints, and is especially efficient on linearly constrained problems. A linearly constrained problem with a quadratic objective is called a quadratic program (QP). Special methods exist for solving QPs, and these are often faster than general purpose optimization procedures.

A vector $\mathbf{x}$ is *feasible* if it satisfies all the constraints. The set of all feasible points is called the *feasible region F*. If F is empty, the problem is *infeasible*, and if feasible points exist at which the objective f is arbitrarily large in a max problem or arbitrarily small in a min problem, the problem is *unbounded*. A point (vector) $\mathbf{x}^*$ is termed a local *extremum (minimum)* if

$$f(\mathbf{x}^*) \leq f(\mathbf{x})$$ (4.2)

for all $\mathbf{x}$ in a small neighborhood (region) N in F around $\mathbf{x}^*$ with $\mathbf{x}$ distinct from $\mathbf{x}^*$. Despite the fact that $\mathbf{x}^*$ is a local extremum, other extrema may exist outside the neighboorhood N meaning that the NLP problem may have more that one local minimum if the entire space of $\mathbf{x}$ is examined. Another important concept relates to the idea of a *global extremum*, the unique solution of the NLP problem. A *global minimum* occurs if Equation (4.2) holds for all $\mathbf{x} \in F$. Analogous concepts exist for *local maxima* and the *global maximum*. Most (but not all) algorithms for solving NLP problems locate a local extremum from a given starting point.

NLP geometry

A typical feasible region for a problem with two variables and the constraints

$$x_j \geq 0, \qquad g_i(\mathbf{x}) \leq 0, \qquad i = 1, 2, \qquad j = 1, 2$$
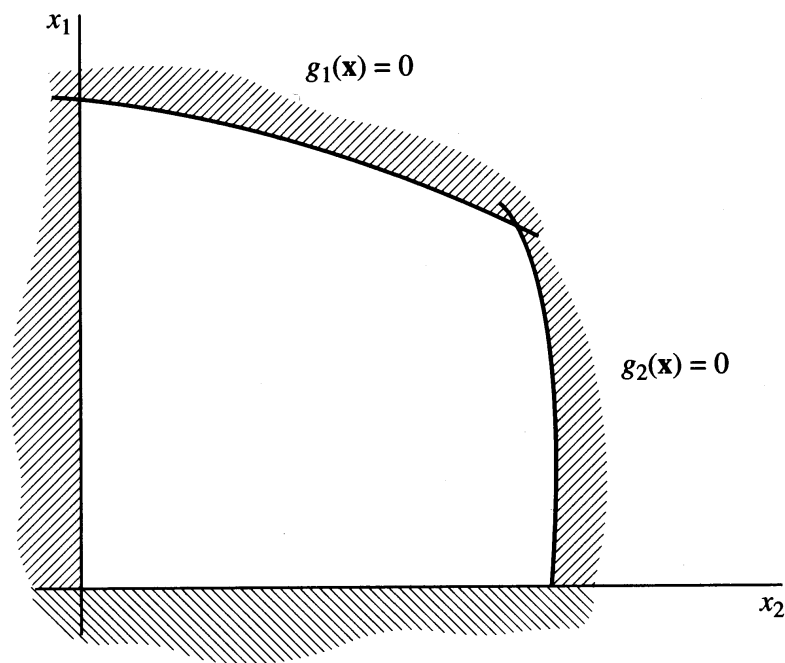
FIGURE 4.4
Feasible region (region not shaded and its boundaries).

is shown as the unshaded region in Figure 4.4. Its boundaries are the straight and curved lines $x_j = 0$ and $g_i(\mathbf{x}) = 0$ for $i = 1, 2, j = 1, 2$.

As another example, consider the problem

$$\text{Minimize} f = (x_1 - 3)^2 + (x_2 - 4)^2$$

subject to the linear constraints

$$x_1 \geq 0$$

$$x_2 \geq 0$$

$$5 - x_1 - x_2 \geq 0$$

$$-2.5 + x_1 - x_2 \leq 0$$

This problem is shown in Figure 4.5. The feasible region is defined by linear constraints with a finite number of corner points. The objective function, being non-linear, has contours (the concentric circles, *level sets*) of constant value that are not parallel lines, as would occur if it were linear. The minimum value of f corresponds to the contour of lowest value having at least one point in common with the feasible region, that is, at $x_1^* = 2, x_2^* = 3$. This is not an extreme point of the feasible set, although it is a boundary point. For linear programs the minimum is always at an extreme point, as shown in Chapter 7.

Furthermore, if the objective function of the previous problem is changed to

$$f_1 = (x_1 - 2)^2 + (x_2 - 2)^2$$
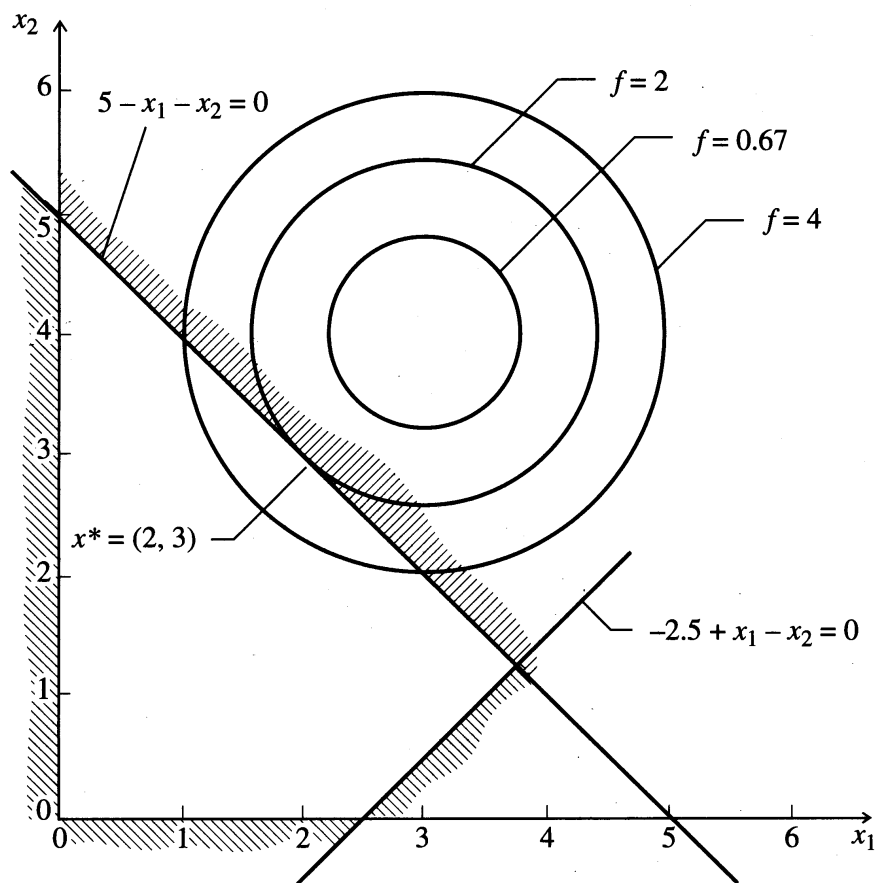
FIGURE 4.5
The minimum occurs on the boundary of the constraint set.

as depicted in Figure 4.6, the minimum is now at $x_1 = 2$, $x_2 = 2$, which is not a boundary point of the feasible region, but is the unconstrained minimum of the non-linear function and satisfies all the constraints.

Neither of the problems illustrated in Figures 4.5 and 4.6 had more than one optimum. It is easy, however, to construct nonlinear programs in which local optima occur. For example, if the objective function f_1 had two minima and at least one was interior to the feasible region, then the constrained problem would have two local minima. Contours of such a function are shown in Figure 4.7. Note that the minimum at the boundary point $x_1 = 3$, $x_2 = 2$ is the global minimum at $f = 3$; the feasible local minimum in the interior of the constraints is at $f = 4$.

Although the examples thus far have involved linear constraints, the chief non-linearity of an optimization problem often appears in the constraints. The feasible region then has curved boundaries. A problem with nonlinear constraints may have local optima, even if the objective function has only one unconstrained optimum. Consider a problem with a quadratic objective function and the feasible region shown in Figure 4.8. The problem has local optima at the two points a and b because no point of the feasible region in the immediate vicinity of either point yields a smaller value of f.
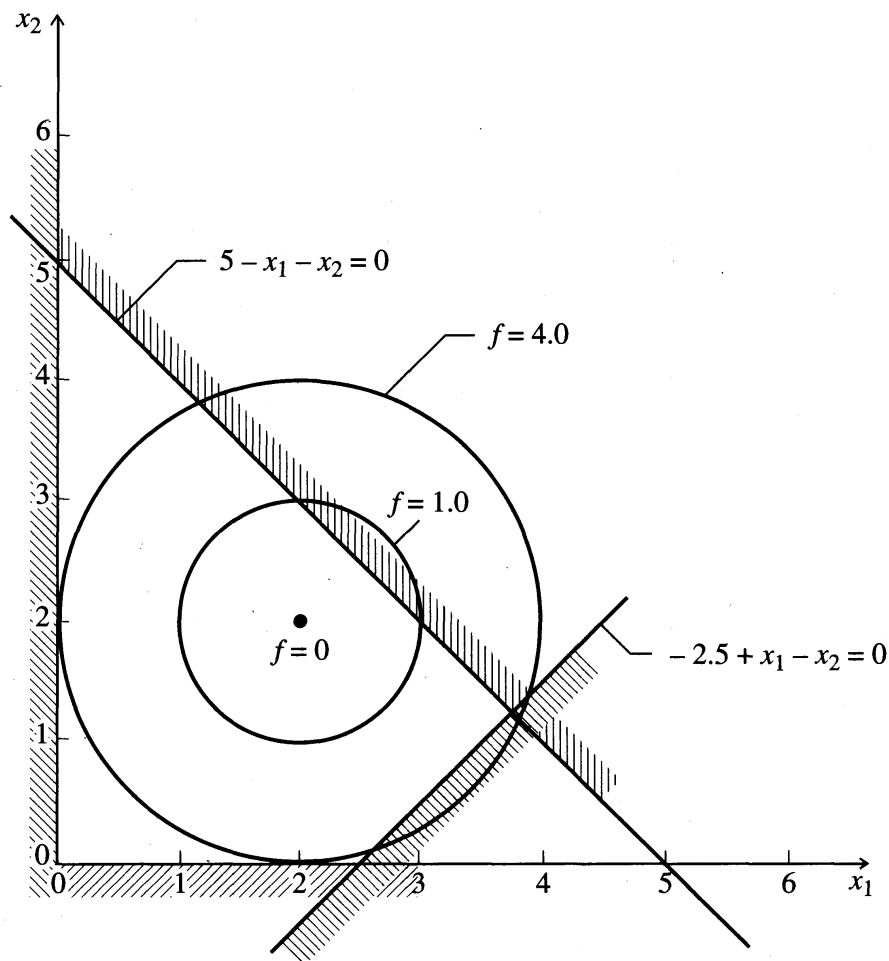
FIGURE 4.6
The minimum occurs in the interior of the constraint set.

In summary, the optimum of a nonlinear programming problem is, in general, not at an extreme point of the feasible region and may not even be on the boundary. Also, the problem may have local optima distinct from the global optimum. These properties are direct consequences of nonlinearity. A class of nonlinear problems can be defined, however, that are guaranteed to be free of distinct local optima. They are called convex programming problems and are considered in the following section.

4.3 CONVEXITY AND ITS APPLICATIONS

The concept of *convexity* is useful both in the theory and applications of optimization. We first define a *convex set*, then a *convex function*, and lastly look at the role played by convexity in optimization.

Convex set

A set of points (or a *region*) is defined as a *convex set* in n-dimensional space if, for all pairs of points x_1 and x_2 in the set, the straight-line segment joining them is also entirely in the set. Figure 4.9 illustrates the concept in two dimensions.
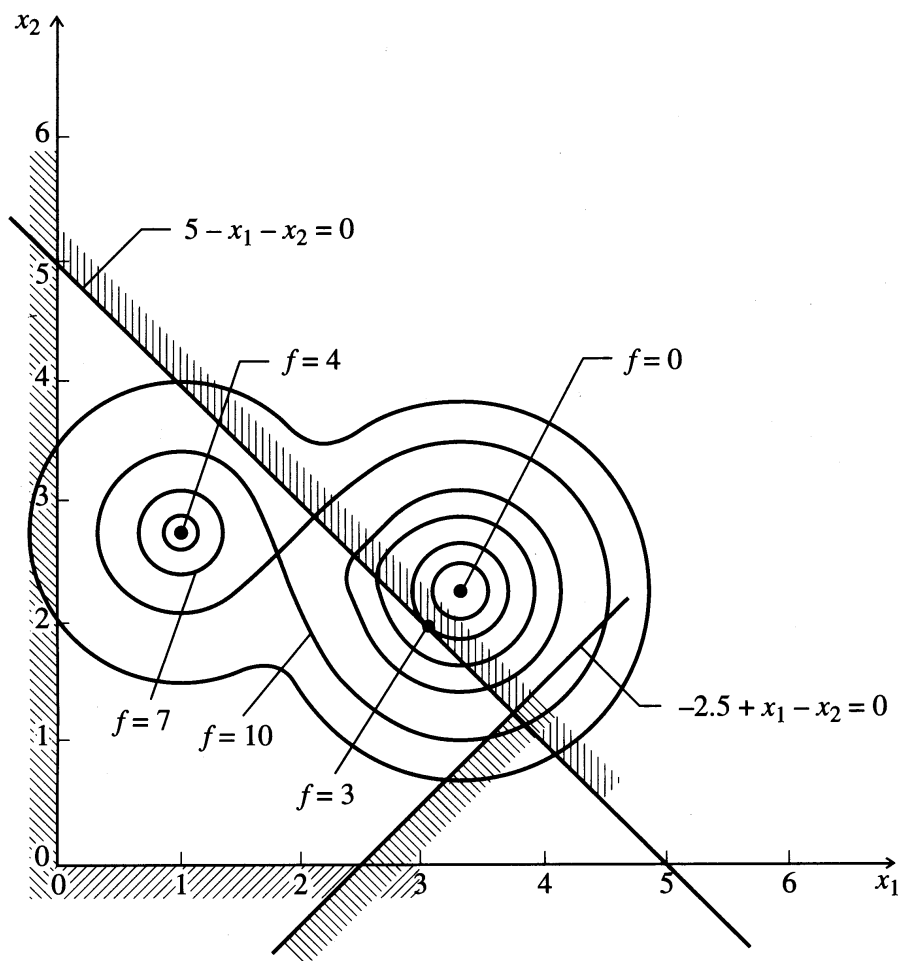
FIGURE 4.7
Local optima due to objective function.

A mathematical statement of a convex set is

For every pair of points $\mathbf{x}_1$ and $\mathbf{x}_2$ in a convex set, the point $\mathbf{x}$ given by a linear combination of the two points

$$\mathbf{x} = \gamma\mathbf{x}_1 + (1 - \gamma)\mathbf{x}_2 \quad 0 \le \gamma \le 1$$

is also in the set. The convex region may be *closed* (*bounded*) by a set of functions, such as the sets A and B in Figure 4.9 or may be *open* (*unbounded*) as in Figures 4.10 and 4.12. Also, the intersection of any number of convex set is a convex set.

Convex function

Next, let us examine the matter of a *convex function*. The concept of a convex function is illustrated in Figure 4.10 for a function of one variable. Also shown is a *concave function*, the negative of a convex function. (If $f(\mathbf{x})$ is convex, $-f(\mathbf{x})$ is concave.) A function $f(\mathbf{x})$ defined on a convex set F is said to be a *convex function* if the following relation holds

$$f[\gamma\mathbf{x}_1 + (1 - \gamma)\mathbf{x}_2] \le \gamma f(\mathbf{x}_1) + (1 - \gamma)f(\mathbf{x}_2)$$
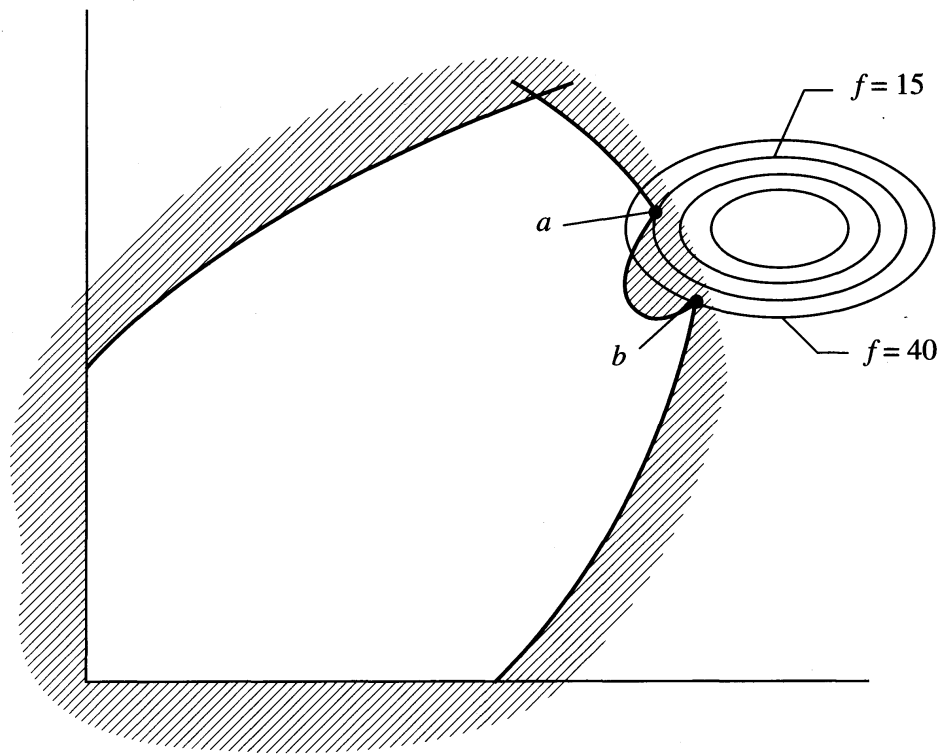
FIGURE 4.8
Local optima due to feasible region.

where γ is a scaler with the range $0 \le \gamma \le 1$. If only the inequality sign holds, the function is said to be not only convex but *strictly convex*. [If $f(\mathbf{x})$ is strictly convex, $-f(\mathbf{x})$ is *strictly concave*.] Figure 4.10 illustrates both a strictly convex and a strictly concave function. A convex function cannot have any value larger than the values of the function obtained by linear interpolation between $\mathbf{x}_1$ and $\mathbf{x}_2$ (the cord between $\mathbf{x}_1$ and $\mathbf{x}_2$ shown in the top figure in Figure 4.10). Linear functions are both convex and concave, but not strictly convex or concave, respectively. An important result of convexity is

If $f(\mathbf{x})$ is convex, then the set

$$R = \{\mathbf{x} | f(\mathbf{x}) \le k\}$$

is convex for all scalers k.

The result is illustrated in Figure 4.11 in which a convex quadratic function is cut by the plane $f(\mathbf{x}) = k$. The convex set R projected on to the $\mathbf{x}_1$–$\mathbf{x}_2$ plane comprises the boundary ellipse plus its interior.

The convex programming problem

An important result in mathematical programming evolves from the concept of convexity. For the nonlinear programming problem called the *convex programming problem*

$$\text{Minimize:} \quad f(\mathbf{x})$$
$$\text{Subject to:} \quad g_i(\mathbf{x}) \le 0 \quad i = 1, \ldots, m \tag{4.3}$$
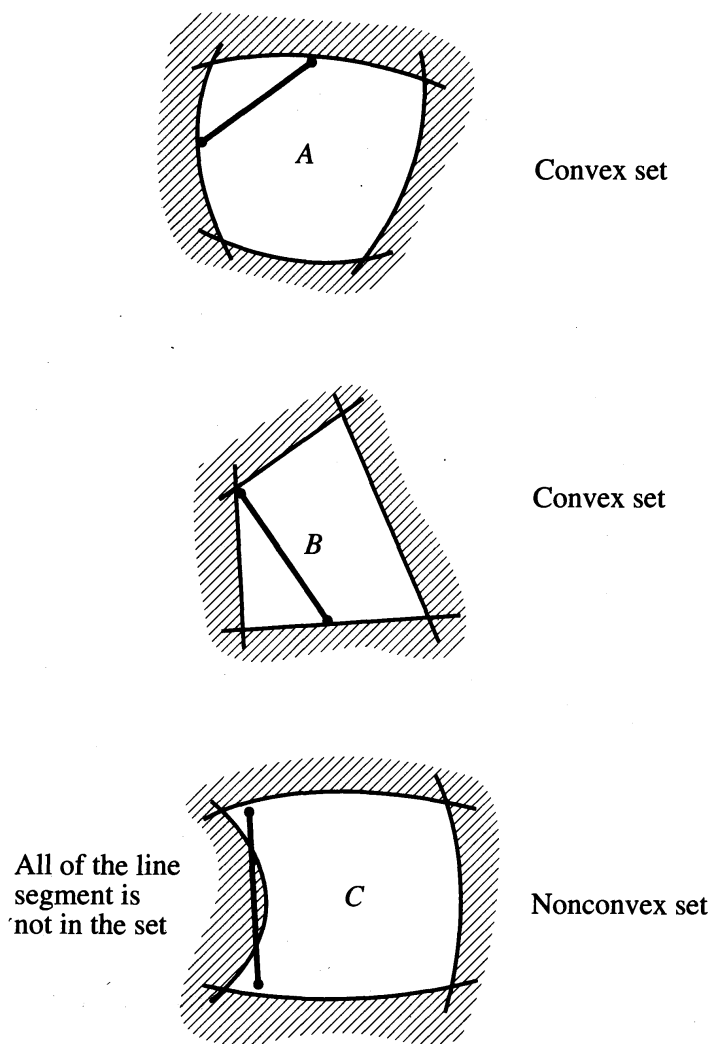
FIGURE 4.9
Convex and nonconvex sets.

in which (a) $f(\mathbf{x})$ is a convex function, and (b) each inequality constraint is a convex function (so that the constraints form a convex set), the following property can be shown to be true

The *local minimum* of $f(\mathbf{x})$ is also the *global minimum*.

Analogously, a local maximum is the global maximum of $f(\mathbf{x})$ if the objective function is concave and the constraints form a convex set.

Role of convexity

If the constraint set $\mathbf{g}(\mathbf{x})$ is nonlinear, the set

$$R = \{\mathbf{x}|\mathbf{g}(\mathbf{x}) = \mathbf{0}\}$$

is generally not convex. This is evident geometrically because most nonlinear functions have graphs that are curved surfaces. Hence the set R is usually a curved surface also, and the line segment joining any two points on this surface generally does not lie on the surface.

$f(x_1)$

$\gamma f(x_1) + (1 - \gamma) f(x_2)$

$f[\gamma x_1 + (1 - \gamma) x_2]$

$f(x_2)$

$\gamma x_1 + (1 - \gamma) x_2$

x_1 x_2

Convex function

Concave function
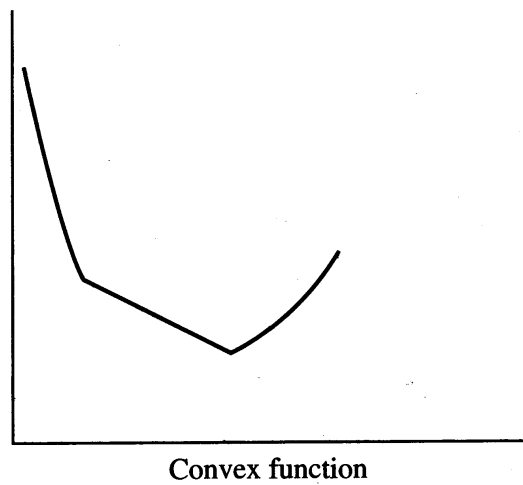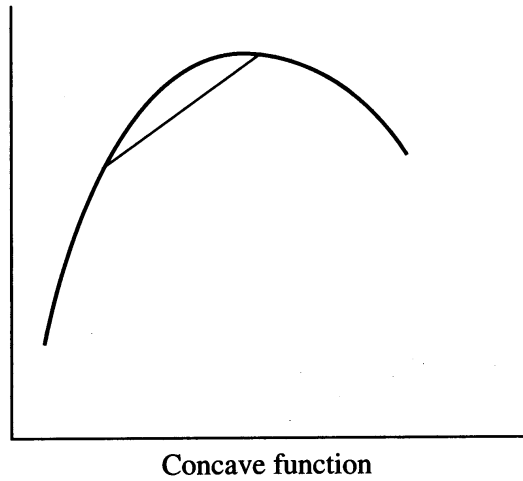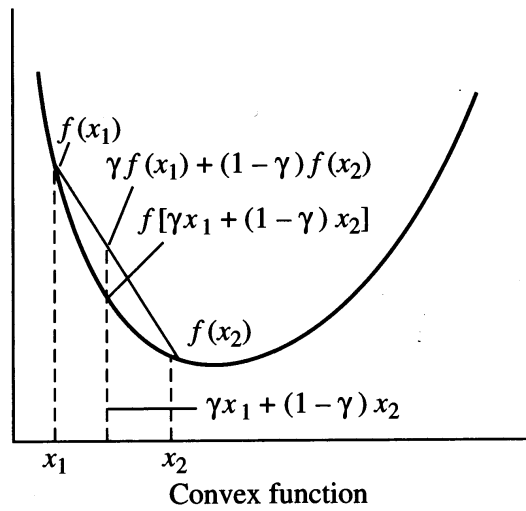
Convex function

FIGURE 4.10
Convex and concave functions of one variable.

FIGURE 4.11

Illustration of a convex set formed by a plane $f(\mathbf{x}) = k$ cutting a convex function.

As a consequence, the problem

Minimize: $f(\mathbf{x})$

Subject to: $g_i(\mathbf{x}) \le 0 \quad i = 1, \ldots, m$
$h_k(\mathbf{x}) = 0 \quad k = 1, \ldots, r < n$

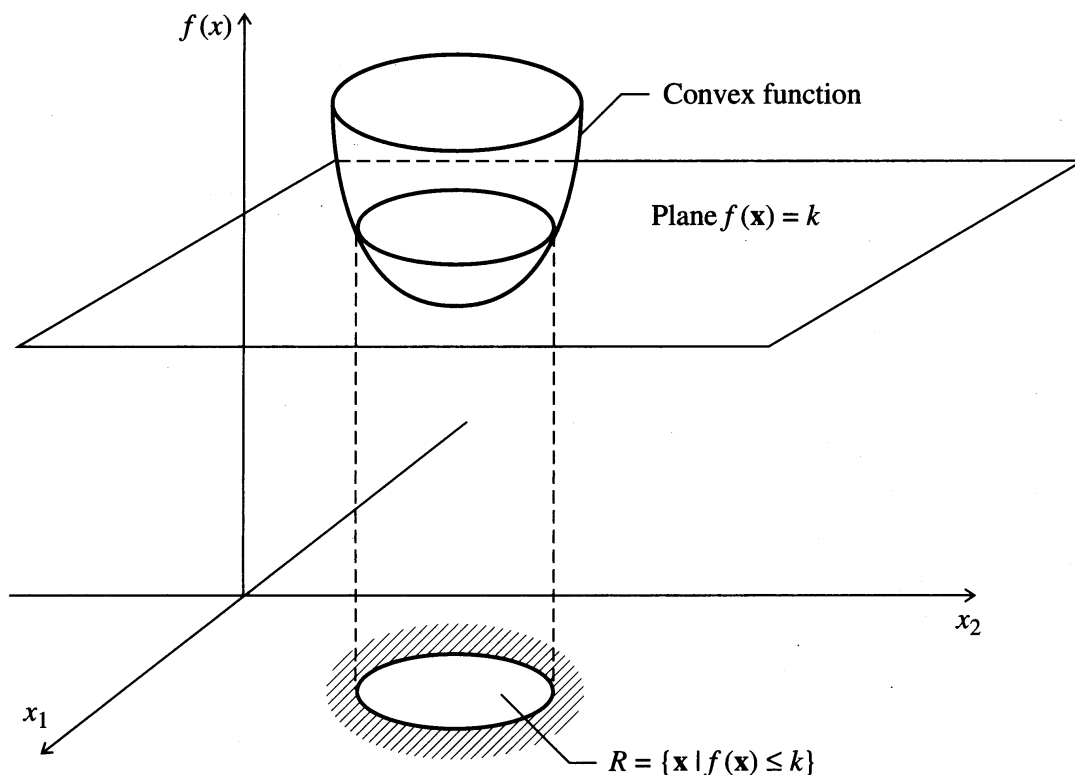may not be a convex programming problem in the variables $x_1, \ldots, x_n$ if any of the functions $h_k(\mathbf{x})$ are nonlinear. This, of course, does not preclude efficient solution of such problems, but it does make it more difficult to guarantee the absence of local optima and to generate sharp theoretical results.

In many cases the equality constraints may be used to eliminate some of the variables, leaving a problem with only inequality constraints and fewer variables. Even if the equalities are difficult to solve analytically, it may still be worthwhile solving them numerically. This is the approach taken by the generalized reduced gradient method, which is described in Section 8.7.

Although convexity is desirable, many real-world problems turn out to be non-convex. In addition, there is no simple way to demonstrate that a nonlinear problem is a convex problem for all feasible points. Why, then is convex programming studied? The main reasons are

1. When convexity is assumed, many significant mathematical results have been derived in the field of mathematical programming.
2. Often results obtained under assumptions of convexity can give insight into the properties of more general problems. Sometimes, such results may even be carried over to nonconvex problems, but in a weaker form.

For example, it is usually impossible to prove that a given algorithm will find the global minimum of a nonlinear programming problem unless the problem is convex. For nonconvex problems, however, many such algorithms find at least a local minimum. Convexity thus plays a role much like that of linearity in the study of dynamic systems. For example, many results derived from linear theory are used in the design of nonlinear control systems.

Determination of convexity and concavity

The definitions of convexity and a convex function are not directly useful in establishing whether a region or a function is convex because the relations must be applied to an unbounded set of points. The following is a helpful property arising from the concept of a convex set of points. A set of points $\mathbf{x}$ satisfying the relation

$$\mathbf{x}^T \mathbf{H}(\mathbf{x})\mathbf{x} \leq 1$$

is convex if the Hessian matrix $\mathbf{H}(\mathbf{x})$ is a real symmetric positive-semidefinite matrix. $\mathbf{H}(\mathbf{x})$ is another symbol for $\nabla^2 f(\mathbf{x})$, the matrix of second partial derivative of $f(\mathbf{x})$ with respect to each x_i

$$\mathbf{H}(\mathbf{x}) \equiv \mathbf{H} \equiv \nabla^2 f(\mathbf{x})$$

The status of $\mathbf{H}$ can be used to identify the character of extrema. A quadratic form $Q(\mathbf{x}) = \mathbf{x}^T \mathbf{H} \mathbf{x}$ is said to be *positive-definite* if $Q(\mathbf{x}) > 0$ for all $\mathbf{x} \neq \mathbf{0}$, and said to be *positive-semidefinite* if $Q(\mathbf{x}) \geq 0$ for all $\mathbf{x} \neq \mathbf{0}$. *Negative-definite* and *negative-semidefinite* are analogous except the inequality sign is reversed. If $Q(\mathbf{x})$ is positive-definite (semidefinite), $\mathbf{H}(\mathbf{x})$ is said to be a positive-definite (semidefinite) matrix. These concepts can be summarized as follows:

1. $\mathbf{H}$ is *positive-definite* if and only if $\mathbf{x}^T \mathbf{H} \mathbf{x}$ is > 0 for all $\mathbf{x} \neq \mathbf{0}$.
2. $\mathbf{H}$ is *negative-definite* if and only if $\mathbf{x}^T \mathbf{H} \mathbf{x}$ is < 0 for all $\mathbf{x} \neq \mathbf{0}$.
3. $\mathbf{H}$ is *positive-semidefinite* if and only if $\mathbf{x}^T \mathbf{H} \mathbf{x}$ is ≥ 0 for all $\mathbf{x} \neq \mathbf{0}$.
4. $\mathbf{H}$ is *negative-semidefinite* if and only if $\mathbf{x}^T \mathbf{H} \mathbf{x}$ is ≤ 0 for all $\mathbf{x} \neq \mathbf{0}$.
5. $\mathbf{H}$ is *indefinite* if $\mathbf{x}^T \mathbf{H} \mathbf{x} < 0$ for some $\mathbf{x}$ and > 0 for other $\mathbf{x}$.

It can be shown from a Taylor series expansion that if $f(\mathbf{x})$ has continuous second partial derivatives, $f(\mathbf{x})$ is concave if and only if its Hessian matrix is negative-semidefinite. For $f(\mathbf{x})$ to be strictly concave, $\mathbf{H}$ must be negative-definite. For $f(\mathbf{x})$ to be convex $\mathbf{H}(\mathbf{x})$ must be positive-semidefinite and for $f(\mathbf{x})$ to be strictly convex, $\mathbf{H}(\mathbf{x})$ must be positive-definite.

EXAMPLE 4.3   ANALYSIS FOR CONVEXITY AND CONCAVITY

For each of these functions
 (a) $f(x) = 3x^2$
 (b) $f(x) = 2x$
 (c) $f(x) = -5x^2$
 (d) $f(x) = 2x^2 - x^3$

determine if $f(x)$ is convex, concave, strictly convex, strictly concave, all, or none of these classes in the range $-\infty \leq x \leq \infty$.

Solution

(a) $f''(x) = 6$, always positive, hence $f(x)$ is both strictly convex and convex.

(b) $f''(x) = 0$ for all values of x, hence $f(x)$ is convex and concave. Note straight lines are both convex and concave simultaneously.

(c) $f''(x) = -10$, always negative, hence $f(x)$ is both strictly concave and concave.

(d) $f''(x) = 6 - 3x$; may be positive or negative depending on the value of x, hence $f(x)$ is not convex or concave over the entire range of x.

For a multivariate function, the nature of convexity can best be evaluated by examining the eigenvalues of $f(\mathbf{x})$ as shown in Table 4.1 We have omitted the indefinite case for $\mathbf{H}$, that is when $f(\mathbf{x})$ is neither convex or concave.

TABLE 4.1
Relationship between the character of $f(\mathbf{x})$ and the state of $\mathbf{H}(\mathbf{x})$

$f(\mathbf{x})$ is	$\mathbf{H}(\mathbf{x})$ is	All the eigenvalues of $\mathbf{H}(\mathbf{x})$ are
Strictly convex	Positive-definite	>0
Convex	Positive-semidefinite	≥ 0
Concave	Negative-semidefinite	≤ 0
Strictly concave	Negative-definite	<0

Now let us further illustrate the ideas presented in this section by some examples.

EXAMPLE 4.4   DETERMINATION OF POSITIVE-DEFINITENESS OF A FUNCTION

Classify the function $f(\mathbf{x}) = 2x_1^2 - 3x_1x_2 + 2x_2^2$ using the categories in Table 4.1, or state that it does not belong in any of the categories.

Solution

$$\frac{\partial f(x)}{\partial x_1} = 4x_1 - 3x_2 \qquad \frac{\partial^2 f(x)}{\partial x_1^2} = 4 \qquad \frac{\partial^2 f(x)}{\partial x_2^2} = 4$$

$$\frac{\partial f(x)}{\partial x_2} = -3x_1 + 4x_2 \qquad \frac{\partial^2 f(x)}{\partial x_1 \partial x_2} = \frac{\partial^2 f(x)}{\partial x_2 \partial x_1} = -3 \qquad \mathbf{H}(\mathbf{x}) = \begin{bmatrix} 4 & -3 \\ -3 & 4 \end{bmatrix}$$

The eigenvalues of $\mathbf{H}$ are 7 and 1, hence $\mathbf{H}(\mathbf{x})$ is positive-definite. Consequently, $f(\mathbf{x})$ is strictly convex (as well as convex).

EXAMPLE 4.5 DETERMINATION OF POSITIVE-DEFINITENESS OF A FUNCTION

Repeat the analysis of Example 4.4 for $f(\mathbf{x}) = x_1{}^2 + x_1x_2 + 2x_2 + 4$

Solution

$$\mathbf{H(x)} = \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix}$$

The eigenvalues are $1 + \sqrt{2}$ and $1 - \sqrt{2}$, or one positive or one negative value. Consequently, $f(\mathbf{x})$ does not fall into any of the categories in Table 4.1. We conclude that no unique extremum exists.

EXAMPLE 4.6 DETERMINATION OF CONVEXITY AND CONCAVITY

Determine if the following function

$$f(\mathbf{x}) = 2x_1 + 3x_2 + 6$$

is convex or concave.

Solution

$$\mathbf{H(x)} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

hence the function is both convex and concave.

EXAMPLE 4.7 DETERMINATION OF CONVEXITY OF A FUNCTION

Consider the following objective function: Is it convex?

$$f(\mathbf{x}) = 2x_1^2 + 2x_1x_2 + 1.5x_2^2 + 7x_1 + 8x_2 + 24$$

Solution

$$\frac{\partial^2 f(x)}{\partial x_1^2} = 4 \qquad \frac{\partial^2 f(x)}{\partial x_2^2} = 3 \qquad \frac{\partial^2 f(x)}{\partial x_1 \partial x_2} = \frac{\partial^2 f(x)}{\partial x_2 \partial x_1} = 2$$

Therefore the Hessian matrix is

$$\mathbf{H(x)} = \begin{bmatrix} 4 & 2 \\ 2 & 3 \end{bmatrix}$$

The eigenvalues of $\mathbf{H(x)}$ are 5.56 and 1.44. Because both eigenvalues are positive, the function is strictly convex (and convex, of course) for all values of x_1 and x_2.

EXAMPLE 4.8   DETECTION OF A CONVEX REGION

Does the following set of constraints that form a closed region form a convex region?

$$-x_1^2 + x_2 \geq 1$$

$$x_1 - x_2 \geq -2$$

Solution. A plot of the two functions indicates that the region circumscribed is closed. The arrows in Figure E4.8 designate the directions in which the inequalities hold. Write the inequality constraints as $g_i \geq 0$. Therefore

$$g_1(\mathbf{x}) = -x_1^2 + x_2 - 1 \geq 0$$

$$g_2(\mathbf{x}) = x_1 - x_2 + 2 \geq 0$$

That the enclosed region is convex can be demonstrated by showing that both $g_1(\mathbf{x})$ and $g_2(\mathbf{x})$ are concave functions:

$$\mathbf{H}[g_1(\mathbf{x})] = \begin{bmatrix} -2 & 0 \\ 0 & 0 \end{bmatrix} \qquad \text{negative definite}$$

$$\mathbf{H}[g_2(\mathbf{x})] = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \qquad \text{negative semidefinite}$$

Because all eigenvalues are zero or negative, according to Table 4.1 both g_1 and g_2 are concave and the region is convex.
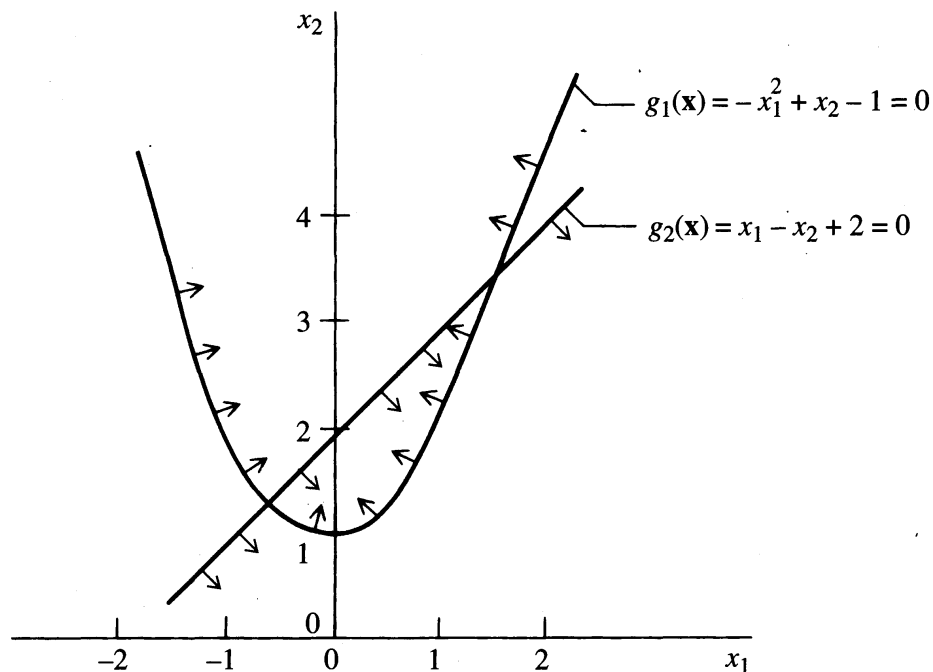
FIGURE E4.8
Convex region composed of two concave functions.

EXAMPLE 4.9   CONSTRUCTION OF A CONVEX REGION

Construct the region given by the following inequality constraints; is it convex?

$$x_1 \leq 6;\, x_2 \leq 6;\, x_1 \geq 0;\, x_1 + x_2 \leq 6;\, x_2 \geq 0$$

Solution. See Figure E4.9 for the region delineated by the inequality constraints. By visual inspection, the region is convex. This set of linear inequality constraints forms a convex region because all the constraints are concave. In this case the convex region is closed.
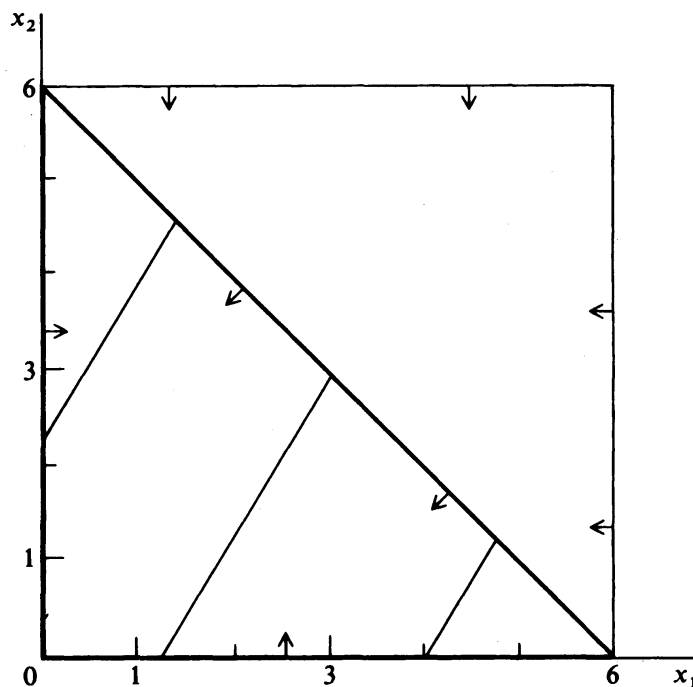
FIGURE E4.9
Diagram of region defined by linear inequality constraints.

4.4   INTERPRETATION OF THE OBJECTIVE FUNCTION IN TERMS OF ITS QUADRATIC APPROXIMATION

If a function of two variables is quadratic or approximated by a quadratic function $f(\mathbf{x}) = b_0 + b_1 x_1 + b_2 x_2 + b_{11} x_1^2 + b_{22} x_2^2 + b_{12} x_1 x_2$, then the eigenvalues of $\mathbf{H}(\mathbf{x})$ can be calculated and used to interpret the nature of $f(\mathbf{x})$ at $\mathbf{x}^*$. Table 4.2 lists some conclusions that can be reached by examining the eigenvalues of $\mathbf{H}(\mathbf{x})$ for a function of two variables, and Figures 4.12 through 4.15 illustrate the different types of surfaces corresponding to each case that arises for quadratic function. By

TABLE 4.2
Geometric interpretation of a quadratic function

Case	Eigenvalue relations	Signs e_1	e_2	Types of contours	Geometric interpretation	Character of center of contours	Figure
1	$e_1 = e_2$	−	−	Circles	Circular hill	Maximum	4.12
2	$e_1 = e_2$	+	+	Circles	Circular valley	Minimum	4.12
3	$e_1 > e_2$	−	−	Ellipses	Elliptical hill	Maximum	4.12
4	$e_1 > e_2$	+	+	Ellipses	Elliptical valley	Minimum	4.12
5	$|e_1| = |e_2|$	+	−	Hyperbolas	Symmetrical saddle	Saddle point	4.13
6	$|e_1| = |e_2|$	−	+	Hyperbolas	Symmetrical saddle	Saddle point	4.13
7	$e_1 > e_2$	+	−	Hyperbolas	Elongated saddle	Saddle point	4.13
8	$e_2 = 0$	−		Straight lines	Stationary ridge*	None	4.14
9	$e_2 = 0$	+		Straight lines	Stationary valley*	None	4.14
10	$e_2 = 0$	−		Parabolas	Rising ridge*‡	At ∞	4.15
11	$e_2 = 0$	+		Parabolas	Falling valley*‡	At ∞	4.15

*These are "degenerate" surfaces.
‡The condition of rising or falling must be evaluated from the linear terms in $f(\mathbf{x})$.
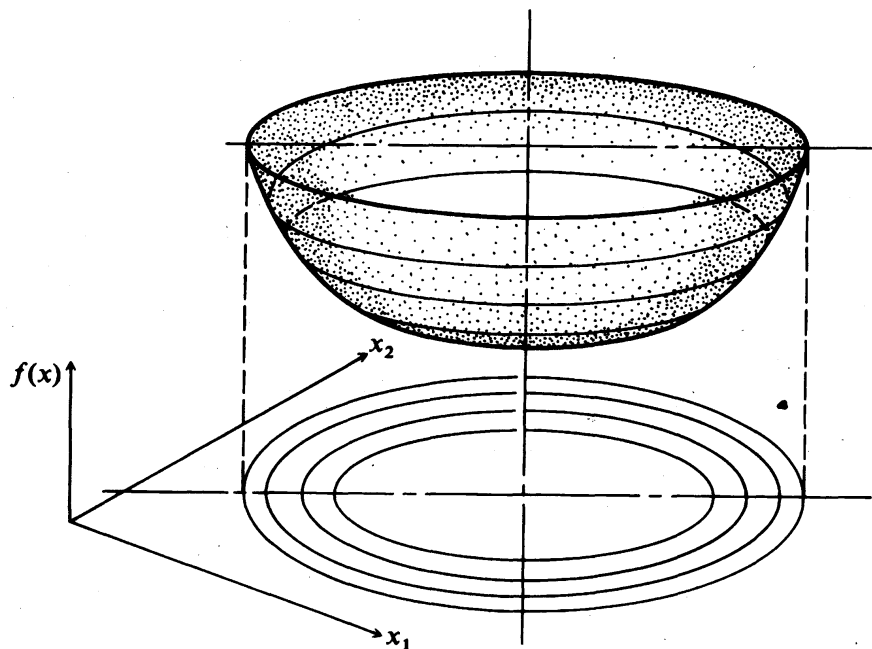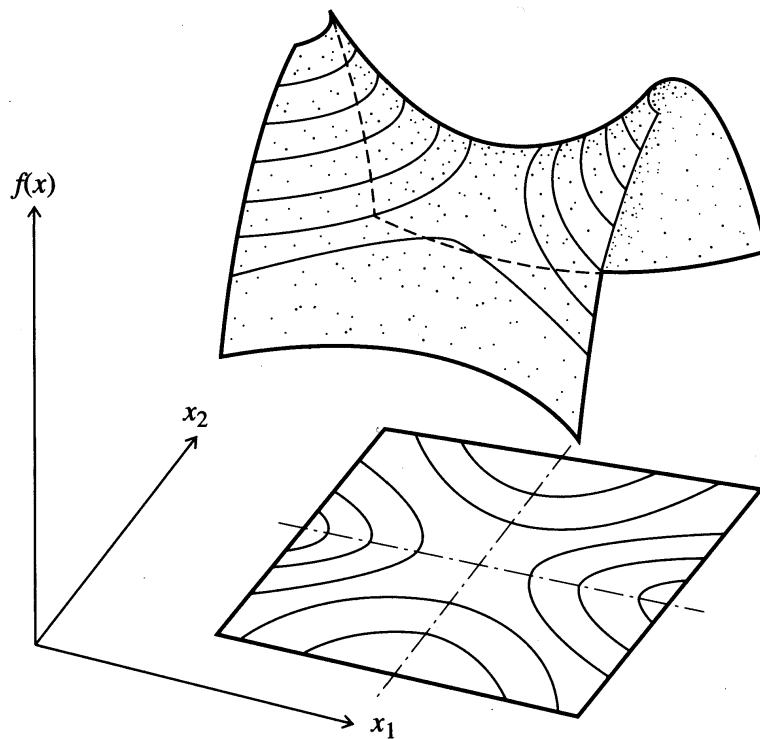

FIGURE 4.12
Geometry of a quadratic objective function of two independent
variables—elliptical contours. If the eigenvalues are equal, then the
contours are circles.

implication, analysis of a function of many variables via examination of the eigen-
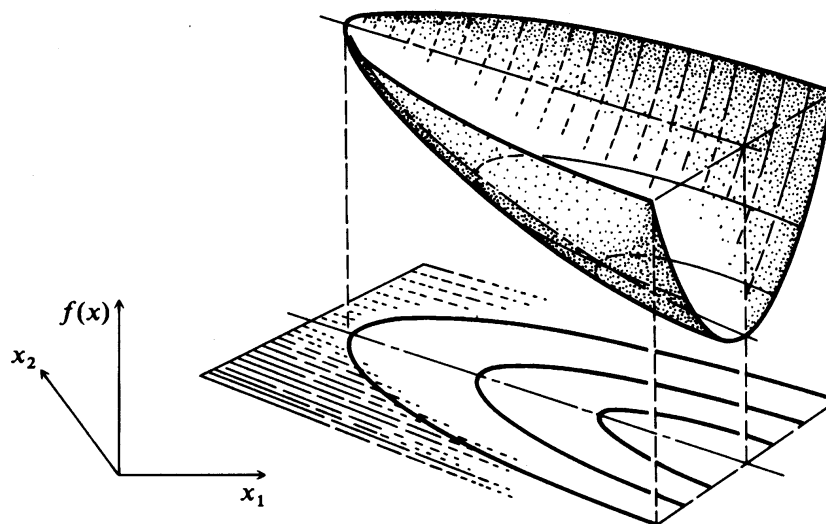values can be conducted, whereas contour plots are limited to functions of only two
or three variables.

FIGURE 4.13
Geometry of a quadratic objective function of two
independent variables—saddle point.
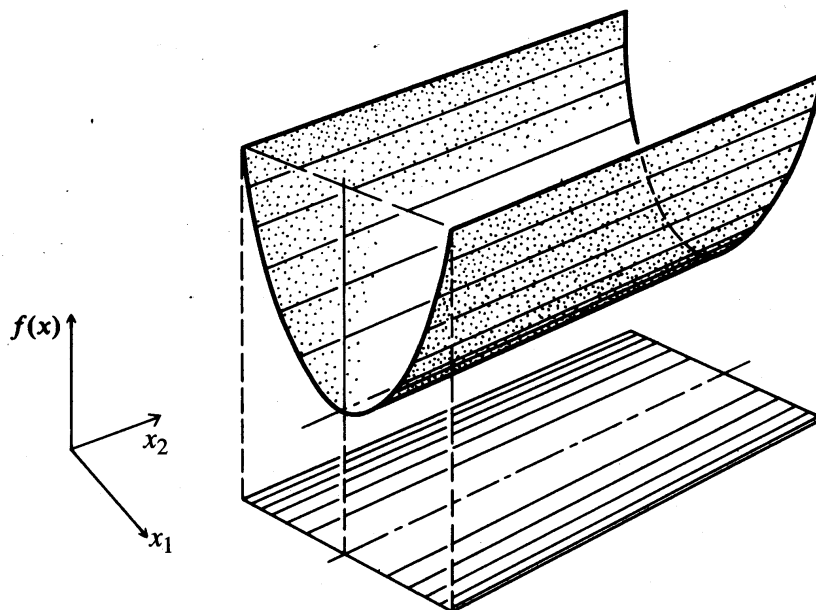
FIGURE 4.14
Geometry of a quadratic objective function of two independent
variables—stationary valley.

FIGURE 4.15

Geometry of second-order objective function of two independent variables—
falling valley.

Figure 4.12 corresponds to objective functions in well-posed optimization problems. In Table 4.2, cases 1 and 2 correspond to contours of $f(\mathbf{x})$ that are concentric circles, but such functions rarely occur in practice. Elliptical contours such as correspond to cases 3 and 4 are most likely for well-behaved functions. Cases 5 to 10 correspond to degenerate problems, those in which no finite maximum or minimum or perhaps nonunique optima appear.

For well-posed quadratic objective functions the contours always form a convex region; for more general nonlinear functions, they do not (see the next section for an example). It is helpful to construct contour plots to assist in analyzing the performance of multivariable optimization techniques when applied to problems of two or three dimensions. Most computer libraries have contour plotting routines to generate the desired figures.

As indicated in Table 4.2, the eigenvalues of the Hessian matrix of $f(\mathbf{x})$ indicate the shape of a function. For a positive-definite symmetric matrix, the eigenvectors (refer to Appendix A) form an orthonormal set. For example, in two dimensions, if the eigenvectors are $\mathbf{v}_1$ and $\mathbf{v}_2$, $\mathbf{v}_1^T \mathbf{v}_2 = 0$ (the eigenvectors are perpendicular to each other). The eigenvectors also correspond to the directions of the principal axes of the contours of $f(\mathbf{x})$.

One of the primary requirements of any successful optimization technique is the ability to move rapidly in a local region along a narrow valley (in minimization) toward the minimum of the objective function. In other words, an efficient algorithm selects a search direction that generally follows the axis of the valley rather than jumping back and forth across the valley. Valleys (ridges in maximization) occur quite frequently, at least locally, and these types of surfaces have the potential to slow down greatly the search for the optimum. A valley lies in the direction of the eigenvector associated with a small eigenvalue of the Hessian

matrix of the objective function. For example, if the Hessian matrix of a quadratic function is

$$\mathbf{H} = \begin{bmatrix} 1 & 0 \\ 0 & 10 \end{bmatrix}$$

then the eigenvalues are $e_1 = 1$ and $e_2 = 10$. The eigenvector associated with $e_1 = 1$, that is, the x_1 axis, is lined up with the valley in the ellipsoid. Variable transformation techniques can be used to allow the problem to be more efficiently solved by a search technique (see Chapter 6).

Valleys and ridges corresponding to cases 1 through 4 can lead to a minimum or maximum, respectively, but not for cases 8 through 11. Do you see why?

4.5   NECESSARY AND SUFFICIENT CONDITIONS FOR AN EXTREMUM OF AN UNCONSTRAINED FUNCTION

Figure 4.16 illustrates the character of $f(x)$ if the objective function is a function of a single variable. Usually we are concerned with finding the minimum or maximum of a multivariable function $f(\mathbf{x})$. The problem can be interpreted geometrically as finding the point in an n-dimension space at which the function has an extremum. Examine Figure 4.17 in which the contours of a function of two variables are displayed.

An optimal point $\mathbf{x}^*$ is completely specified by satisfying what are called the *necessary* and *sufficient conditions* for optimality. A condition N is necessary for a result R if R can be true only if the condition is true $(R \Rightarrow N)$. The reverse is not true, however, that is, if N is true, R is not necessarily true. A condition is sufficient for a result R if R is true if the condition is true $(S \Rightarrow R)$. A condition T is necessary and sufficient for result R if R is true if and only if T is true $(T \Leftrightarrow R)$.
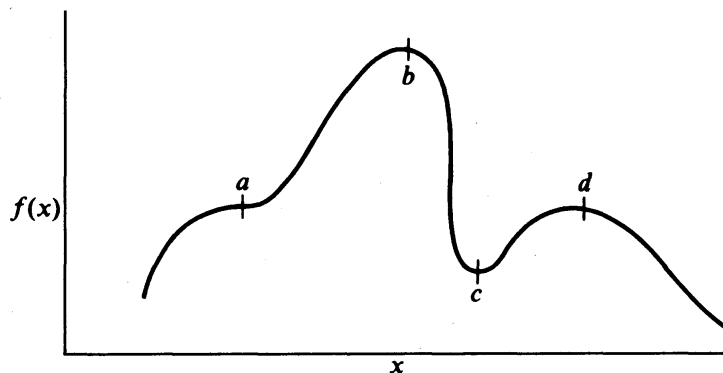
FIGURE 4.16

A function exhibiting different types of stationary points.

Key: a—inflection point (scalar equivalent to a saddle point);
b—global maximum (and local maximum); *c*—local minimum;
d—local maximum
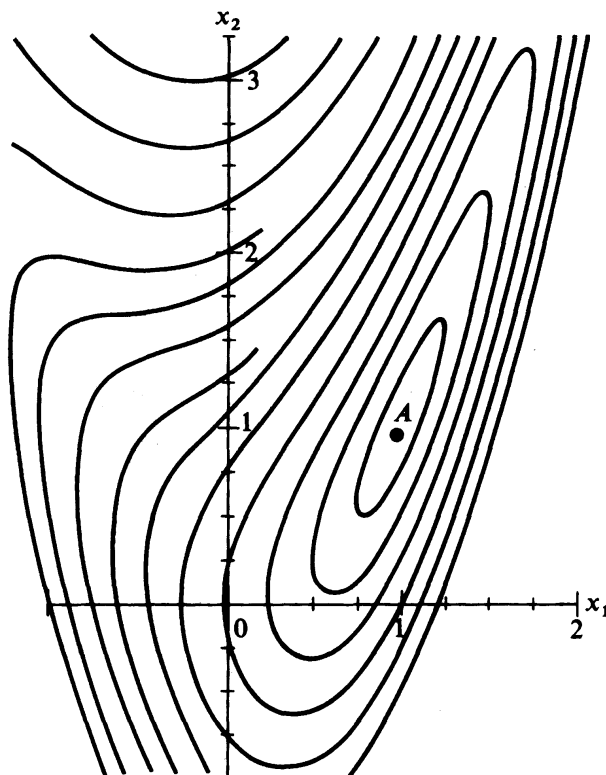
FIGURE 4.17a
A function of two variables with a single stationary point
(the extremum).

The easiest way to develop the necessary and sufficient conditions for a minimum or maximum of $f(\mathbf{x})$ is to start with a Taylor series expansion about the presumed extremum $\mathbf{x}^*$

$$f(\mathbf{x}) = f(\mathbf{x}^*) + \nabla^T f(\mathbf{x}^*)\,\Delta\mathbf{x} + \tfrac{1}{2}(\Delta\mathbf{x}^T)\,\nabla^2 f(\mathbf{x}^*)\Delta\mathbf{x} + O_3(\Delta\mathbf{x}) + \cdots \quad (4.4)$$

where $\Delta\mathbf{x} = \mathbf{x} - \mathbf{x}^*$, the perturbation of $\mathbf{x}$ from $\mathbf{x}^*$. We assume all terms in Equation (4.4) exist and are continuous, but will ignore the terms of order 3 or higher $[O_3(\Delta\mathbf{x})]$, and simply analyze what occurs for various cases involving just the terms through the second order.

We defined a local minimum as a point $\mathbf{x}^*$ such that no other point in the vicinity of $\mathbf{x}^*$ yields a value of $f(\mathbf{x})$ less than $f(\mathbf{x}^*)$, or

$$f(\mathbf{x}) - f(\mathbf{x}^*) \geq 0 \quad (4.5)$$

$\mathbf{x}^*$ is a global minimum if Equation (4.5) holds for any $\mathbf{x}$ in the n-dimensional space of $\mathbf{x}$. Similarly, $\mathbf{x}^*$ is a local maximum if

$$f(\mathbf{x}) - f(\mathbf{x}^*) \leq 0 \quad (4.6)$$
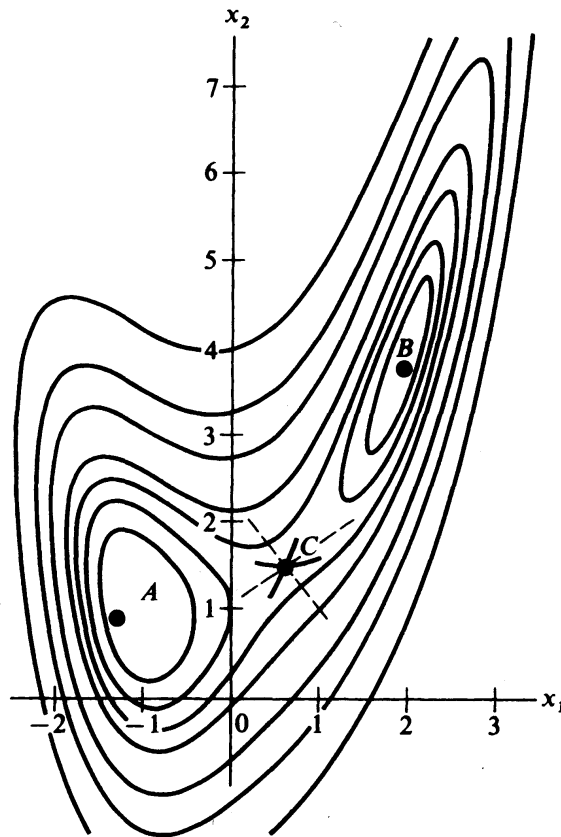
FIGURE 4.17b

A function of two variables with three stationary
points and two extrema, A and B.

Examine the second term on the right-hand side of Equation (4.4): $\nabla^T f(\mathbf{x}^*) \Delta\mathbf{x}$.
Because $\Delta\mathbf{x}$ is arbitrary and can have both plus and minus values for its elements,
we must insist that $\nabla f(\mathbf{x}^*) = \mathbf{0}$. Otherwise the resulting term added to $f(\mathbf{x}^*)$ would
violate Equation (4.5) for a minimum, or Equation (4.6) for a maximum. Hence, a
necessary condition for a minimum or maximum of $f(\mathbf{x})$ is that the gradient of $f(\mathbf{x})$
vanishes at $\mathbf{x}^*$

$$\nabla f(\mathbf{x}^*) = 0 \tag{4.7}$$

that is, $\mathbf{x}^*$ is a stationary point.

With the second term on the right-hand side of Equation (4.4) forced to be zero,
we next examine the third term: $\frac{1}{2}(\Delta\mathbf{x}^T)\nabla^2 f(\mathbf{x}^*)\Delta\mathbf{x}$. This term establishes the character of the stationary point (minimum, maximum, or saddle point). In Figure 4.17b,
A and B are minima and C is a saddle point. Note how movement along one of the
perpendicular search directions (dashed lines) from point C increases $f(\mathbf{x})$, whereas
movement in the other direction decreases $f(\mathbf{x})$. Thus, satisfaction of the necessary
conditions does not guarantee a minimum or maximum.

To establish the existence of a minimum or maximum at $\mathbf{x}^*$, we know from
Equation (4.4) with $\nabla f(\mathbf{x}^*) = \mathbf{0}$ and the conclusions reached in Section 4.3 concerning convexity that for $\Delta\mathbf{x} \neq 0$ we have the following outcomes

$\nabla^2 f(\mathbf{x}^*) \equiv \mathbf{H}(\mathbf{x}^*)$	$\Delta\mathbf{x}^T \nabla^2 f(\mathbf{x}^*) \Delta\mathbf{x}$	Near $\mathbf{x}^*, f(\mathbf{x}) - f(\mathbf{x}^*)$
Positive-define	>0	Increases
Positive-semidefinite	≥ 0	Possibly increases
Negative-definite	<0	Decreases
Negative-semidefinite	≤ 0	Possibly decreases
Indefinite	Both ≤ 0 and ≥ 0 depending on $\Delta\mathbf{x}$	Increases, decreases, neither

Consequently, $\mathbf{x}^*$ can be classified as

$\nabla^2 f(\mathbf{x}^*) \equiv \mathbf{H}(\mathbf{x}^*)$	$\mathbf{x}^*$
Positive-definite	Unique ("isolated") minimum
Negative-definite	Unique ("isolated") maximum

These two conditions are known as the *sufficiency conditions*.

In summary, the necessary conditions (items 1 and 2 in the following list) and the sufficient condition (3) to guarantee that $\mathbf{x}^*$ is an extremum are as follows:

1. $f(\mathbf{x})$ is twice differentiable at $\mathbf{x}^*$.
2. $\nabla f(\mathbf{x}^*) = \mathbf{0}$, that is, a stationary point exists at $\mathbf{x}^*$.
3. $\mathbf{H}(\mathbf{x}^*)$ is positive-definite for a minimum to exist at $\mathbf{x}^*$, and negative-definite for a maximum to exist at $\mathbf{x}^*$.

Of course, a minimum or maximum may exist at $\mathbf{x}^*$ even though it is not possible to demonstrate the fact using the three conditions. For example, if $f(x) = x^{4/3}$, $x^* = 0$ is a minimum but $\mathbf{H}(0)$ is not defined at $x^* = 0$, hence condition 3 is not satisfied.

EXAMPLE 4.10   CALCULATION OF A MINIMUM OF $f(x)$

Does $f(x) = x^4$ have an extremum? If so, what is the value of x^* and $f(x^*)$ at the extremum?

Solution

$$f'(x) = 4x^3 \qquad f''(x) = 12x^2$$

Set $f'(x) = 0$ and solve for x; hence $x = 0$ is a stationary point. Also, $f''(0) = 0$, meaning that condition 3 is *not* satisfied. Figure E4.10 is a plot of $f(\mathbf{x}) = x^4$. Thus, a minimum exists for $f(x)$ but the sufficiency condition is not satisfied.

If both first and second derivatives vanish at the stationary point, then further analysis is required to evaluate the nature of the function. For functions of a single variable, take successively higher derivatives and evaluate them at the stationary point. Continue this procedure until one of the higher derivatives is not zero (the nth one); hence, $f'(x^*), f''(x^*), \ldots, f^{(n-1)}(x^*)$ all vanish. Two cases must be analyzed:

1. If n is even, the function attains a maximum or a minimum; a positive sign of $f^{(n)}$ indicates a minimum, a negative sign a maximum.
2. If n is odd, the function exhibits a saddle point.

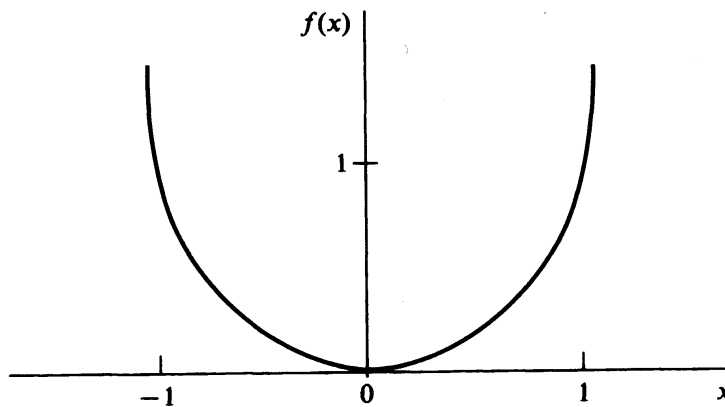For more details refer to Beveridge and Schechter (1970).

FIGURE E4.10

For application of these guidelines to $f(x) = x^4$, you will find $d^4 f(x)/dx^4 = 24$ for which n is even and the derivative is positive, so that a minimum exists.

EXAMPLE 4.11   CALCULATION OF EXTREMA

Identify the stationary points of the following function (Fox, 1971), and determine if any extrema exist.

$$f(\mathbf{x}) = 4 + 4.5x_1 - 4x_2 + x_1^2 + 2x_2^2 - 2x_1x_2 + x_1^4 - 2x_1^2x_2$$

Solution. For this function, three stationary points can be located by setting $\nabla f(\mathbf{x}) = 0$:

$$\frac{\partial f(x)}{\partial x_1} = 4.5 + 2x_1 - 2x_2 + 4x_1^3 - 4x_1x_2 = 0 \tag{a}$$

$$\frac{\partial f(x)}{\partial x_2} = -4 + 4x_2 - 2x_1 - 2x_1^2 = 0 \tag{b}$$

The set of nonlinear equations (a) and (b) has to be solved, say by Newton's method, to get the pairs (x_1, x_2) as follows:

Point	Stationary point (x_1, x_2)	$f(\mathbf{x})$	Hessian matrix eigenvalues		Classification
B	(1.941, 3.854)	0.9855	37.03	0.97	Local minimum
A	(−1.053, 1.028)	−0.5134	10.5.	3.5	Local minimum (also the global minimum)
C	(0.6117, 1.4929)	2.83	7.0	−2.56	Saddle point

Figure 4.17b shows contours for the objective function in this example. Note that the global minimum can only be identified by evaluating $f(\mathbf{x})$ for all the local minima.

For general nonlinear objective functions, it is usually difficult to ascertain the nature of the stationary points without detailed examination of each point.

EXAMPLE 4.12

In many types of processes such as batch constant-pressure filtration or fixed-bed ion exchange, the production rate decreases as a function of time. At some optimal time t^{opt}, production is terminated (at P^{opt}) and the equipment is cleaned. Figure E4.12a illustrates the cumulative throughput $P(t)$ as a function of time t for such a process. For one cycle of production and cleaning, the overall production rate is

$$R(t) = \frac{P(t)}{t + t_c} \qquad\qquad (a)$$

where $R(t)$ = the overall production rate per cycle (mass/time)

t_c = the cleaning time (assumed to be constant)

Determine the maximum production rate and show that P^{opt} is indeed the maximum throughout.

Solution. Differentiate $R(t)$ with respect to t, and equate the derivative to 0:

$$\frac{dR(t)}{dt} = \frac{-P(t) + [dP(t)/dt](t + t_c)}{(t + t_c)^2} = 0$$

$$P^{opt} = \frac{dP(t)}{dt}\bigg|_{opt}(t + t_c) \qquad\qquad (b)$$

The geometric interpretation of Equation (b) is the classical result (Walker et al., 1937) that the tangent to $P(t)$ at P^{opt} intersects the time axis at $-t_c$. Examine Figure E4.12b. The maximum overall production rate is
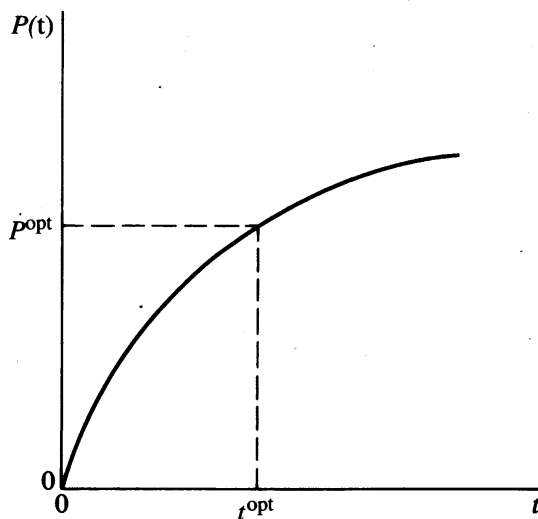
FIGURE E4.12a
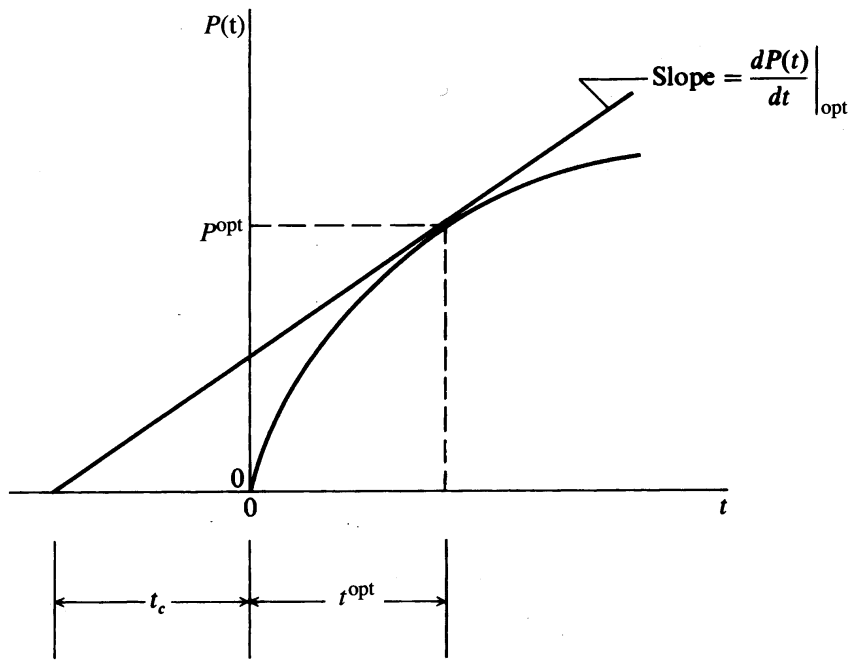
FIGURE E4.12b
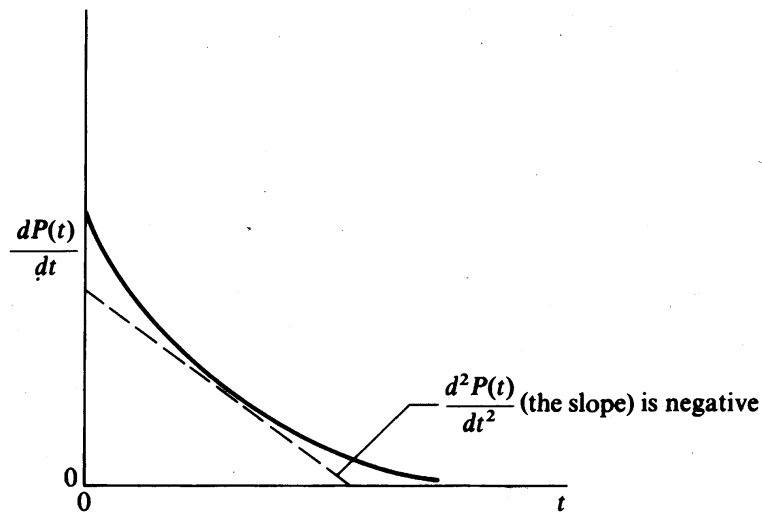
FIGURE E4.12c

$$R^{\text{opt}} = \frac{P^{\text{opt}}}{t^{\text{opt}} + t_c} \qquad\qquad (c)$$

Does P^{opt} meet the sufficiency condition to be a maximum? Is

$$\frac{d^2R(t)}{dt^2} = \frac{2P(t) - 2[dP(t)/dt](t + t_c) + [d^2P(t)/dt^2](t + t_c)^2}{(t + t_c)^3} < 0 \quad ? \qquad (d)$$

Rearrangement of (d) and introduction of (b) into (d), or the pair (P^{opt}, t^{opt}), gives

$$\frac{d^2P(t)}{dt^2}(t + t_c)^2 < 0$$

From Figure E4.12b we note in the range $0 < t < t^{opt}$ that $dP(t)/dt$ is always positive and decreasing so that $d^2P(t)/dt^2$ is always negative (see Figure E4.12c). Consequently, the sufficiency condition is met.

REFERENCES

Bartela, R.; J. Beatty; and B. Barsky. *An Introduction to Splines for Use in Computer Graphics and Geometric Modeling*. Morgan Kaufman Publishers, Los Altos, CA (1987).

Beveridge, G. S. G.; and R. S. Schechter. *Optimization: Theory and Practice*, McGraw-Hill, New York (1970), p. 126.

Fox, R. L. *Optimization Methods for Engineering Design*. Addison-Wesley, Reading, MA (1971), p. 42.

Happell, J.; and D. G. Jordan. *Chemical Process Economics*. Marcel Dekker, New York (1975), p. 178.

Kaplan, W. *Maxima and Minima with Applications: Practical Optimization and Duality*. Wiley, New York (1998).

Nash, S. G.; and A. Sofer. *Linear and Nonlinear Programming*. McGraw-Hill, New York (1996).

Noltie, C. B. *Optimum Pipe Size Selection*. Gulf Publ. Co., Houston, Texas (1978).

Walker, W. H.; W. K. Lewis; W. H. McAdams; and E. R. Gilliland. *Principles of Chemical Engineering*. 3d ed. McGraw-Hill, New York (1937), p. 357.

SUPPLEMENTARY REFERENCES

Amundson, N. R. *Mathematical Methods in Chemical Engineering: Matrices and Their Application*. Prentice-Hall, Englewood Cliffs, New Jersey (1966).

Avriel, M. Nonlinear Programming. Prentice-Hall, Englewood Cliffs, New Jersey (1976).

Bazarra, M. S.; H. D. Sherali; and C. M. Shetty. *Nonlinear Programming: Theory and Algorithms*. Wiley, New York (1993).

Borwein, J.; and A. S. Lewis. *Convex Analysis and Nonlinear Optimization*. Springer, New York (1999).

Jeter, M. W. *Mathematical Programming*. Marcel Dekker, New York (1986).

Luenberger, D. G. *Linear and Nonlinear Programming*. 2nd ed. Addison-Wesley, Menlo Park, CA (1984).

PROBLEMS

4.1 Classify the following functions as continuous (specify the range) or discrete:

(a) $f(x) = e^x$

(b) $f(x) = ax_{n-1} + b(x_0 - x_n)$ where x_n represents a stage in a distillation column

(c) $f(\mathbf{x}) = \dfrac{x_D - x_s}{1 + x_s}$ where x_D = concentration of vapor from a still and x_s is the concentration in the still

4.2 The future worth S of a series of n uniform payments each of amount P is

$$S = \frac{P}{i}[(1 + i)^n - 1]$$

where i is the interest rate per period. If i is considered to be the only variable, is it discrete or continuous? Explain. Repeat for n. Repeat for both n and i being variables.

4.3 In a plant the gross profit P in dollars is

$$P = nS - (nV + F)$$

where n = the number of units produced per year
S = the sales price in dollars per unit
V = the variable cost of production in dollars per unit
F = the fixed charge in dollars
Suppose that the average unit cost is calculated as

$$\text{Average unit cost} = \frac{nV + F}{n}$$

Discuss under what circumstances n can be treated as a continuous variable.

4.4 One rate of return is the ratio of net profit P to total investment

$$R = 100\,\frac{P(1 - t)}{I} = 100(1 - t)\frac{[S - (V + F/n)]}{I/n}$$

where t = the fraction tax rate
I = the total investment in dollars
Find the maximum R as a function of n for a given I if n is a continuous variable. Repeat if n is discrete. (See Problem 4.3 for other notation.)

4.5 Rewrite the following linear programming problems in matrix notation.

(a) Minimize: $\quad f(\mathbf{x}) = 3x_1 + 2x_2 + x_3$

Subject to: $\quad g_1(\mathbf{x}) = 2x_1 + 3x_2 + x_3 \geq 10$

$\qquad\qquad\quad g_2(\mathbf{x}) = x_1 + 2x_2 + x_3 \geq 15$

(b) Maximize: $\quad f(\mathbf{x}) = 5x_1 + 10x_2 + 12x_3$

Subject to: $\quad g_1(\mathbf{x}) = 15x_1 + 10x_2 + 10x_3 \leq 200$

$\qquad\qquad\quad g_2(\mathbf{x}) = x_1 \geq 0$

$\qquad\qquad\quad g_3(\mathbf{x}) = x_2 \geq 0$

$\qquad\qquad\quad g_4(\mathbf{x}) = x_3 \geq 0$

$\qquad\qquad\quad h_1(\mathbf{x}) = 10x_1 + 25x_2 + 20x_3 = 300$

4.6 Put the following nonlinear objective function into matrix notation by defining suitable matrices; $\mathbf{x} = [x_1\ x_2]^T$.

$$f(\mathbf{x}) = 3 + 2x_1 + 3x_2 + 2x_1^2 + 2x_1x_2 + 6x_2^2$$

4.7 Sketch the objective function and constraints of the following nonlinear programming problems.

(a) Minimize: $f(\mathbf{x}) = 2x_1^2 - 2x_1x_2 + 2x_2^2 - 6x_1 + 6$

Subject to: $g_1(\mathbf{x}) = x_1 + x_2 \leq 2$

(b) Minimize: $f(\mathbf{x}) = x_1^3 - 3x_1x_2 + 4$

Subject to: $g_1(\mathbf{x}) = 5x_1 + 2x_2 \geq 18$

$h_1(\mathbf{x}) = -2x_1 + x_2^2 = 5$

(c) Minimize: $f(\mathbf{x}) = -5x_1^2 + x_2^2$

Subject to: $g_1(\mathbf{x}) = \dfrac{x_1^2}{x_2^2} - \dfrac{1}{x_2} \leq -1$

$g_2(\mathbf{x}) = x_1 \geq 0$

$g_3(\mathbf{x}) = x_2 \geq 0$

4.8 Distinguish between the local and global extrema of the following objective function.

$$f(\mathbf{x}) = 2x_1^3 + x_2^2 + x_1^2x_2^2 + 4x_1x_2 + 3$$

4.9 Are the following vectors (a) feasible or nonfeasible vectors with regard to Problem 4.5b; (b) interior or exterior vectors?

(1) $\mathbf{x} = \begin{bmatrix} 5 & 2 & 10 \end{bmatrix}^T$

(2) $\mathbf{x} = \begin{bmatrix} 10 & 2 & 7.5 \end{bmatrix}^T$

(3) $\mathbf{x} = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}^T$

4.10 Shade the feasible region of the nonlinear programming problems of Problem 4.7. Is $\mathbf{x} = \begin{bmatrix} 1 & 1 \end{bmatrix}^T$ an interior, boundary, or exterior point in these problems?

4.11 What is the feasible region for $\mathbf{x}$ given the following constraints? Sketch the feasible region for the two-dimensional problems.

(a) $h_1(\mathbf{x}) = x_1 + x_2 - 3 = 0$

$h_2(\mathbf{x}) = 2x_1 - x_2 + 1 = 0$

(b) $h_1(\mathbf{x}) = x_1^2 + x_2^2 + x_3^2 = 0$

$h_2(\mathbf{x}) = x_1 + x_2 + x_3 = 0$

(c) $g_1(\mathbf{x}) = x_1 - x_2^2 - 2 \geq 0$

$g_2(\mathbf{x}) = x_1 - x_2 + 4 \geq 0$

(d) $h_1(\mathbf{x}) = x_1^2 + x_2^2 + 3$

$g_1(\mathbf{x}) = x_1 - x_2 + 2 \geq 0$

$g_2(\mathbf{x}) = x_1 \geq 0$

$g_3(\mathbf{x}) = x_2 \geq 0$

4.12 Two solutions to the nonlinear programming problem

Minimize: $f(\mathbf{x}) = 7x_1 - 6x_2 + 4x_3$

Subject to: $h_1(\mathbf{x}) = x_1^2 + 2x_2^2 + 3x_3^2 - 1 = 0$

$h_2(\mathbf{x}) = 5x_1 + 5x_2 - 3x_3 - 6 = 0$

have been reported, apparently a maximum and a minimum.

$$\mathbf{x} = \begin{bmatrix} 0.947 \\ 0.207 \\ -0.0772 \end{bmatrix} \qquad \mathbf{x} = \begin{bmatrix} 0.534 \\ 0.535 \\ -0.219 \end{bmatrix}$$

$$f(\mathbf{x}) = 5.08 \qquad f(\mathbf{x}) = -0.346$$

Verify that each of these $\mathbf{x}$ vectors is feasible.

4.13 The problem

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

Subject to: $x_1^2 + x_2^2 \leq 2$

is reported to have a local minimum at the point $\mathbf{x}^* = [1\ 1]^T$. Is this local optimum also a global optimum?

4.14 Under what circumstances is a local minimum guaranteed to be the global minimum? (Be brief.)

4.15 Are the following functions convex? Strictly convex? Why?

(a) $2x_1^2 + 2x_1x_2 + 3x_2^2 + 7x_1 + 8x_2 + 25$

What are the optimum values of x_1 and x_2?

(b) e^{5x}

4.16 Determine the convexity or concavity of the following objective functions:

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

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

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

4.17 Show that $f = e^{x_1} + e^{x_2}$ is convex. Is it also strictly convex?

4.18 Show that $f = |x|$ is convex.

4.19 Is the following region constructed by the four constraints convex? Closed?

$$x_2 \geq 1 - x_1$$

$$x_2 \leq 1 + 0.5x_1$$

$$x_1 \leq 2$$

$$x_2 \geq 0$$

4.20 Does the following set of constraints form a convex region?

$$g_1(x) = -(x_1^2 + x_2^2) + 9 \geq 0$$

$$g_2(x) = -x_1 - x_2 + 1 \geq 0$$

4.21 Consider the following problem:

Minimize: $f(\mathbf{x}) = x_1^2 + x_2$

Subject to: $g_1(\mathbf{x}) = x_1^2 + x_2^2 - 9 \leq 0$

$$g_2(\mathbf{x}) = (x_1 + x_2^2) - 1 \leq 0$$

$$g_3(\mathbf{x}) = (x_1 + x_2) - 1 \leq 0$$

Does the constraint set form a convex region? Is it closed? (*Hint*: A plot will help you decide.)

4.22 Is the following function convex, concave, neither, or both? Show your calculations.

$$f(\mathbf{x}) = \ln x_1 + \ln x_2$$

4.23 Sketch the region defined by the following inequality constraints. Is it a convex region? Is it closed?

$$x_1 + x_2 - 1 \geq 0$$

$$x_1 - x_2 + 1 \geq 0$$

$$2 - x_1 \qquad \geq 0$$

$$x_2 \qquad \geq 0$$

4.24 Does the following constraint set form a convex region (set)?

$$h(\mathbf{x}) = x_1^2 + x_2^2 - 9 = 0$$

$$g_1(\mathbf{x}) = -(x_1 + x_2^2) + 1 \geq 0$$

$$g_2(\mathbf{x}) = -(x_1 + x_2) + 1 \geq 0$$

4.25 Separable functions are those that can be expressed in the form

$$\psi(\mathbf{x}) = \sum_{i=1}^{n} \psi_i(\mathbf{x})$$

For example, $x_1^2 + x_2^2 + x_3^2$ is a separable function because

$$\psi(\mathbf{x}) = \sum x_i^2$$

Show that if the terms in a separable function are convex, the separable function is convex.

4.26 Is the following problem a convex programming problem?

Minimize:
$$f(\mathbf{x}) = 100x_1 + \frac{200}{x_1 x_2}$$

Subject to:
$$2x_2 + \frac{300}{x_1 x_2} \leq 1$$

$$x_1, x_2 \geq 0$$

4.27 Classify each of the following matrices as (a) positive-definite, (b) negative-definite, (c) neither.

(a) $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ (c) $\begin{bmatrix} 0.1 & 0 \\ 3 & 1 \end{bmatrix}$

(b) $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ (d) $\begin{bmatrix} 1 & -2 \\ -2 & 1 \end{bmatrix}$

4.28 Determine whether the following matrix is positive-definite, positive-semidefinite, negative-definite, negative-semidefinite, or none of the above. Show all calculations.

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

4.29 In designing a can to hold a specified amount of soda water, the cost function (to be minimized) for manufacturing one can is

$$f(D, h) = \pi D h + \frac{\pi}{2} D^2$$

and the constraints are

$$\frac{\pi}{4} D^2 h \geq 400$$

$$3.5 \leq D \leq 8 \qquad 8 \leq h \leq 18$$

Based on the preceding problem, answer the following; as far as possible for each answer use mathematics to support your statements:
(a) State whether $f(D, h)$ is unimodal (one extremum) or multimodal (more then one extremum).
(b) State whether $f(D, h)$ is continuous or not.
(c) State whether $f(D, h)$ is convex, concave, or neither.
(d) State whether or not $f(D, h)$ alone meets the necessary and sufficient conditions for a minimum to exist.
(e) State whether the constraints form a convex region.

4.30 A reactor converts an organic compound to product P by heating the material in the presence of an additive A (mole fraction = x_A). The additive can be injected into the reactor, while steam can be injected into a heating coil inside the reactor to provide heat. Some conversion can be obtained by heating without addition of A, and vice versa.

The product P can be sold for \$50/lb mol. For 1 lb mol of feed, the cost of the additive (in dollars/lb mol) as a function of x_A is given by the formula, $2.0 + 10x_A + 20x_A^2$. The cost of the steam (in dollars) as a function of S is $1.0 + 0.003S + 2.0 \times 10^{-6} S^2$. ($S =$ lb steam/lb mol feed). The yield equation is $y_p = 0.1 + 0.3x_A + 0.001S + 0.0001x_A S$; $y_p =$ lb mol product P/lb mol feed.

(a) Formulate the profit function (basis of 1.0 lb mol feed) in terms of x_A and S.

$$f = \text{Income} - \text{Costs}$$

The constraints are:

$$0 \le x_A \le 1 \qquad S \ge 0$$

(b) Is f a concave function? Demonstrate mathematically why it is or why it is not concave.
(c) Is the region of search convex? Why?

4.31 The objective function for the work requirement for a three-stage compressor can be expressed as (p is pressure)

$$f = \left(\frac{p_2}{p_1}\right)^{0.286} + \left(\frac{p_3}{p_2}\right)^{0.286} + \left(\frac{p_4}{p_3}\right)^{0.286}$$

$p_1 = 1$ atm and $p_4 = 10$ atm. The minimum occurs at a pressure ratio for each stage of $\sqrt[3]{10}$. Is f convex for $1 \le p_2 \le 10, 1 \le p_3 \le 10$?

4.32 In the following problem
(a) Is the objective function convex? (b) Is the constraint region convex?

Minimize: $f(\mathbf{x}) = 100x_1 + \dfrac{200}{x_1 x_2}$

Subject to: $g(\mathbf{x}) = 2x_2 + \dfrac{300}{x_1 + x_2} \ge 1$

$x_1 \ge 0$

$x_2 \ge 0$

4.33 Answer the questions below for the following problem; in each case *justify* your answer.

Minimize: $f(\mathbf{x}) = \frac{1}{4}x_1^4 - \frac{1}{2}x_1^2 - x_2$

Subject to: $x_1^2 + x_2^2 = 4$

$x_1 - x_2 \le 2$

(a) Is the problem a convex programming problem?
(b) Is the point $\mathbf{x} = [1\ 1]^T$ a feasible point?
(c) Is the point $\mathbf{x} = [2\ 2]^T$ an interior point?

4.34 Happel and Jordan (1975) reported an objective function (cost) for the design of a distillation column as follows:

$$f = 14720(100 - P) + 6560R - 30.2PR + 6560 - 30.2P$$

$$+ 19.5n\,(5000R - 23PR + 5000 - 23P)^{0.5}$$

$$+ 23.2\,[5000R - 23PR + 5000 - 23P]^{0.62}$$

where n = number of theoretical stages
R = reflux ratio
P = percent recovery in bottoms stream

They reported the optimum occurs at $R = 8$, $n = 55$, and $P = 99$. Is f convex at this point? Are there nearby regions where f is not convex?

4.35 Given a linear objective function,

$$f = x_1 + x_2$$

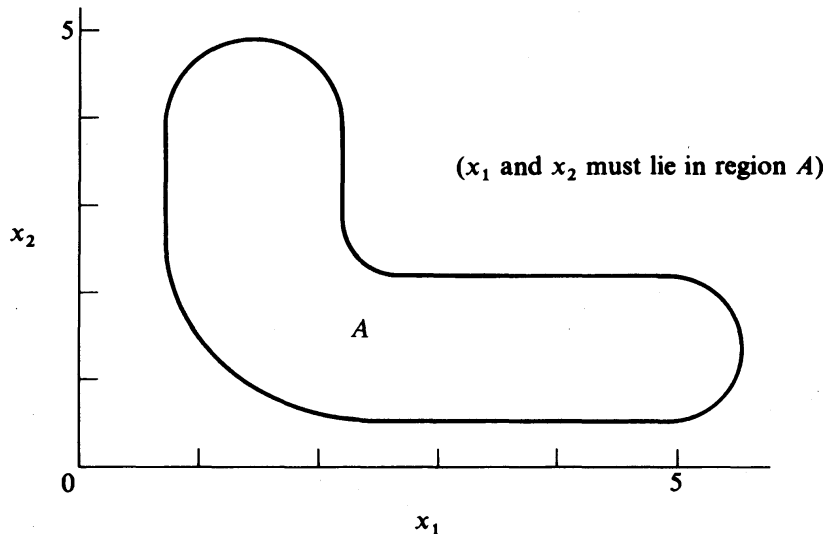(x_1 and x_2 must lie in region A)

FIGURE P4.35

explain why a nonconvex region such as region A in Figure P4.35 causes difficulties in the search for the maximum of f in the region. Why is region A not convex?

4.36 Consider the following objective function

$$f(\alpha) = \sum_{i=1}^{n} |x_i - \alpha|$$

Show that f is convex. *Hint:* Expand f for both n odd and n even. You can plot the function to assist in your analysis. Under what circumstances is

$$f(x) = \sum_{i=1}^{n} c_i |x - \alpha_i|$$

convex?

4.37 Classify the stationary points of

(a) $f = -x^4 + x^3 + 20$
(b) $f = x^3 + 3x^2 + x + 5$
(c) $f = x^4 - 2x^2 + 1$
(d) $f = x_1^2 - 8x_1x_2 + x_2^2$

according to Table 4.2

4.38 List stationary points and their classification (maximum, minimum, saddle point) of

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

(b) $f = x_1 + x_2 + x_1^2 - 4x_1x_2 + 2x_2^2$

4.39 State what type of surface is represented by

$$f(\mathbf{x}) = 2x_1^2x_2 - 2x_2^2 + x_1^3$$

at the stationary point $\mathbf{x} = [0 \ 0]^T$ (use Table 4.2).

4.40 Interpret the geometry of the following function at its stationary point in terms of Table 4.2

$$f(\mathbf{x}) = 3x_1^3x_2$$

4.41 Classify the following function in terms of the list in Table 4.2:

$$f(\mathbf{x}) = 10x_1 - x_1^2 + 10x_2 - x_2^2 - x_1x_2 + x_1^3 - 34$$

4.42 In crystal NaCl, each Na^+ or Cl^- ion is surrounded by 6 nearest neighbors of opposite charge and 12 nearest neighbors of the same charge. Two sets of forces oppose each other: the coulombic attraction and the hard-core repulsion. The potential energy $u(r)$ of the crystal is given by the Lennard–Jones potential expression,

$$u(r) = 4\in \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right]$$

where $\in, \sigma$ are constants, such that $\in > 0, \sigma > 0$.
(a) Does the Lennard–Jones potential $u(r)$ have a stationary point(s)? If it does, locate it (them).
(b) Identify the nature of the stationary point(s) min, max, etc.
(c) What is the magnitude of the potential energy at the stationary points?

4.43 Consider the function

$$y = (x - a)^2$$

Note that $x = a$ minimizes y. Let $z = x^2 - 4x + 16$. Does the solution to $x^2 - 4x + 16 = 0$,

$$x = \frac{4 \pm \sqrt{-48}}{2} = 2 \pm j2\sqrt{3}$$

minimize z? ($j = \sqrt{-1}$).

4.44 The following objective function can be seen by inspection to have a minimum at $x = 0$:

$$f(x) = |x^3|$$

Can the criteria of Section 4.5 be applied to test this outcome?

4.45 (a) Consider the objective function,

$$f = 6x_1^2 + x_2^3 + 6x_1x_2 + 3x_2^2$$

Find the stationary points and classify them using the Hessian matrix.
(b) Repeat for

$$f = 3x_1^2 + 6x_1 + x_2^2 + 6x_1x_2 + x_3 + 2x_3^2 + x_2x_3 + x_2$$

(c) Repeat for

$$f = a_0x_1 + a_1x_2 + a_2x_1^2 + a_3x_2^2 + a_4x_1x_2$$

4.46 An objective function is

$$f(\mathbf{x}) = (x_1 - 8)^2 + (x_2 - 5)^2 + 16$$

By inspection, you can find $\mathbf{x}^* = [8\ 5]^T$ yields the minimum of $f(\mathbf{x})$. Show that $\mathbf{x}^*$ meets the necessary and sufficient conditions for a minimum.

4.47 Analyze the function

$$f(\mathbf{x}) = \tfrac{1}{4}x^4 - \tfrac{1}{2}x^2$$

Find all of its stationary points and determine if they are maxima, minima, or inflection (saddle) points. Sketch the curve in the region of

$$-2 \le x \le 2$$

4.48 Determine if the following objective function

$$f(\mathbf{x}) = 2x_1^3 + x_2^2 + x_1^2x_2^2 + 4x_1x_2 + 3$$

has local minima or maxima. Classify each point clearly.

4.49 Is the following function unimodal (only one extremum) or multimodal (more than one extremum)?

$$f(\mathbf{x}) = \begin{cases} -x^2 & -\infty \le x \le 0 \\ -x^2 & 0 \le x \le 1 \\ e^{x-1} & 1 \le x \le \infty \end{cases}$$

4.50 Determine whether the solution $\mathbf{x} = [-0.87\ -0.8]^T$ for the objective function

$$f(x) = x_1^4 + 12x_2^3 - 15x_1^2 - 56x_2 + 60$$

is indeed a maximum.