Chapter 13

Further Topics in Optimization

This chapter deals with two major topics. The first is nonlinear programming, which extends the techniques of constrained optimization of Chap. 12 by allowing *inequality constraints* into the problem. In Chap. 12, the constraints must be satisfied as strict equalities; i.e., the constraints are always binding. Now we shall consider constraints that may not be binding in the solution; i.e., they may be satisfied as inequalities in the solution.

In the second part of this chapter, we revert back to the realm of classical-constrained optimization to discuss some topics left untouched in the previous chapters. These include the indirect objective function, the envelope theorem, and the concept of duality.

13.1 Nonlinear Programming and Kuhn-Tucker Conditions

In the history of methodological development, the first attempts at dealing with inequality constraints were concentrated on linear ones only. With linearity prevailing in the constraints as well as in the objective function, the resulting methodology is quite naturally christened *linear programming*. Despite the limitation of linearity, however, we could for the first time, explicitly specify the choice variables to be nonnegative, as is appropriate in most economic analysis. This represents a significant advance. Nonlinear programming, a later development, makes it possible even to handle nonlinear inequality constraints and nonlinear objective function. Thus it occupies a most important place in optimization methodology.

In the classical optimization problem, with no explicit restrictions on the signs of the choice variables, and with no inequalities in the constraints, the first-order condition for a relative or local extremum is simply that the first partial derivatives of the (smooth) Lagrangian function with respect to all the choice variables and the Lagrange multipliers be zero. In nonlinear programming, there exists a similar type of first-order condition, known as the *Kuhn-Tucker conditions.*[†] As we shall see, however, while the classical first-order condition is always necessary, the Kuhn-Tucker conditions cannot be accorded the

[†] H. W. Kuhn and A. W. Tucker, "Nonlinear Programming," in J. Neyman (ed.), *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press, Berkeley, California, 1951, pp. 481–492.

status of necessary conditions unless a certain proviso is satisfied. On the other hand, under certain specific circumstances, the Kuhn-Tucker conditions turn out to be *sufficient conditions,* or even *necessary-and-sufficient* conditions as well.

Since the Kuhn-Tucker conditions are the single most important analytical result in nonlinear programming, it is essential to have a proper understanding of those conditions as well as their implications. For the sake of expository convenience, we shall develop these conditions in two steps.

Step 1: Effect of Nonnegativity Restrictions

As the first step, consider a problem with nonnegativity restrictions on the choice variables, but with no other constraints. Taking the single-variable case, in particular, we have

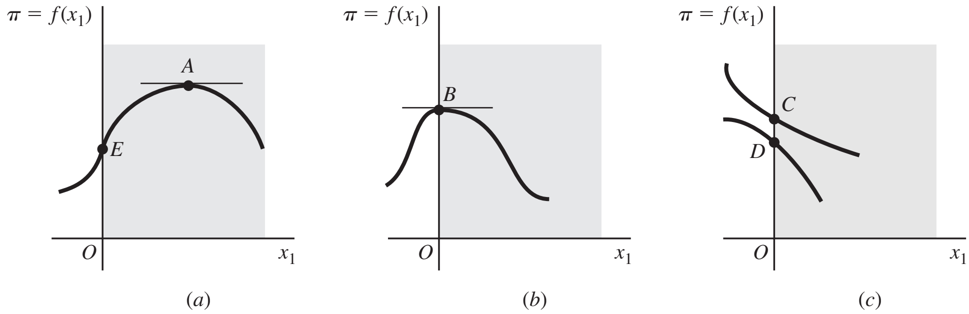
$$\begin{aligned} &\text{Maximize} && \pi = f(x_1) \\ &\text{subject to} && x_1 \geq 0 \end{aligned} \tag{13.1}$$

where the function f is assumed to be differentiable. In view of the restriction $x_1 \geq 0$, three possible situations may arise. First, if a local maximum of π occurs in the interior of the shaded feasible region in Fig. 13.1, such as point A in Fig. 13.1a, then we have an *interior solution.* The first-order condition in this case is $d\pi/dx_1 = f'(x_1) = 0$, same as in the classical problem. Second, as illustrated by point B in Fig. 13.1b, a local maximum can also occur on the vertical axis, where $x_1 = 0$. Even in this second case, where we have a *boundary solution,* the first-order condition $f'(x_1) = 0$ nevertheless remains valid. However, as a third possibility, a local maximum may in the present context take the position of point C or point D in Fig. 13.1c, because to qualify as a local maximum in problem (13.1), the candidate point merely has to be higher than the neighboring points *within* the feasible region. In view of this last possibility, the maximum point in a problem like (13.1) can be characterized, not only by the equation $f'(x_1) = 0$, but also by the inequality $f'(x_1) < 0$. Note on the other hand, that the opposite inequality $f'(x_1) > 0$ can safely be ruled out, for at a point where the curve is upward-sloping, we can never have a maximum, even if that point is located on the vertical axis, such as point E in Fig. 13.1a.

The upshot of the preceding discussion is that, in order for a value of x_1 to give a local maximum of π in problem (13.1), it must satisfy one of the following three conditions

$$f'(x_1) = 0 \quad \text{and} \quad x_1 > 0 \quad [\text{point } A] \tag{13.2}$$
$$f'(x_1) = 0 \quad \text{and} \quad x_1 = 0 \quad [\text{point } B] \tag{13.3}$$
$$f'(x_1) < 0 \quad \text{and} \quad x_1 = 0 \quad [\text{points } C \text{ and } D] \tag{13.4}$$

FIGURE 13.1



Actually, these three conditions can be consolidated into a single statement

$$f'(x_1) \le 0 \qquad x_1 \ge 0 \qquad \text{and} \qquad x_1 f'(x_1) = 0 \tag{13.5}$$

The first inequality in (13.5) is a summary of the information regarding $f'(x_1)$ enumerated in (13.2) through (13.4). The second inequality is a similar summary for x_1; in fact, it merely reiterates the nonnegativity restriction of the problem. And, as for the third part of (13.5), we have an equation which expresses an important feature common to (13.2) through (13.4), namely, that of the two quantities x_1 and $f'(x_1)$, *at least one* must take a zero value, so that the product of the two must be zero. This feature is referred to as the *complimentary slackness* between x_1 and $f'(x_1)$. Taken together, the three parts of (13.5) constitute the first-order necessary condition for a local maximum in a problem where the choice variable must be nonnegative. But going a step further, we can also take them to be necessary for a *global* maximum. This is because a global maximum must also be a local maximum and, as such, must also satisfy the necessary condition for a local maximum.

When the problem contains n choice variables:

$$\begin{aligned} &\text{Maximize} && \pi = f(x_1, x_2, \ldots, x_n) \\ &\text{subject to} && x_j \ge 0 \qquad (j = 1, 2, \ldots, n) \end{aligned} \tag{13.6}$$

The classical first-order condition $f_1 = f_2 = \cdots = f_n = 0$ must be similarly modified. To do this, we can apply the same type of reasoning underlying (13.5) to each choice variable x_j taken by itself. Graphically, this amounts to viewing the horizontal axis in Fig. 13.1 as representing each x_j in turn. The required modification of the first-order condition then readily suggests itself:

$$f_j \le 0 \qquad x_j \ge 0 \qquad \text{and} \qquad x_j f_j = 0 \qquad (j = 1, 2, \ldots, n) \tag{13.7}$$

where f_j is the partial derivative $\partial\pi/\partial x_j$.

Step 2: Effect of Inequality Constraints

With this background, we now proceed to the second step, and try to include inequality constraints as well. For simplicity, let us first deal with a problem with three choice variables ($n = 3$) and two constraints ($m = 2$):

$$\begin{aligned} &\text{Maximize} && \pi = f(x_1, x_2, x_3) \\ &\text{subject to} && g^1(x_1, x_2, x_3) \le r_1 \\ &&& g^2(x_1, x_2, x_3) \le r_2 \\ &\text{and} && x_1, x_2, x_3 \ge 0 \end{aligned} \tag{13.8}$$

which, with the help of two dummy variables s_1 and s_2, can be transformed into the equivalent form

$$\begin{aligned} &\text{Maximize} && \pi = f(x_1, x_2, x_3) \\ &\text{subject to} && g^1(x_1, x_2, x_3) + s_1 = r_1 \\ &&& g^2(x_1, x_2, x_3) + s_2 = r_2 \\ &\text{and} && x_1, x_2, x_3, s_1, s_2 \ge 0 \end{aligned} \tag{13.8'}$$

If the nonnegativity restrictions are absent, we may, in line with the classical approach, form the Lagrangian function:

$$Z' = f(x_1, x_2, x_3) + \lambda_1[r_1 - g^1(x_1, x_2, x_3) - s_1] \\ + \lambda_2[r_2 - g^2(x_1, x_2, x_3) - s_2] \qquad \textbf{(13.9)}$$

and write the first-order condition as

$$\frac{\partial Z'}{\partial x_1} = \frac{\partial Z'}{\partial x_2} = \frac{\partial Z'}{\partial x_3} = \frac{\partial Z'}{\partial s_1} = \frac{\partial Z'}{\partial s_2} = \frac{\partial Z'}{\partial \lambda_1} = \frac{\partial Z'}{\partial \lambda_2} = 0$$

But since the x_j and s_i variables do have to be nonnegative, the first-order condition on those variables should be modified in accordance with (13.7). Consequently, we obtain the following set of conditions instead:

$$\begin{array}{llll} \dfrac{\partial Z'}{\partial x_j} \le 0 & x_j \ge 0 & \text{and} & x_j \dfrac{\partial Z'}{\partial x_j} = 0 \\ \dfrac{\partial Z'}{\partial s_i} \le 0 & s_i \ge 0 & \text{and} & s_i \dfrac{\partial Z'}{\partial s_i} = 0 \\ \dfrac{\partial Z'}{\partial \lambda_i} = 0 & & & \begin{pmatrix} i = 1, 2 \\ j = 1, 2, 3 \end{pmatrix} \end{array} \qquad \textbf{(13.10)}$$

Note that the derivatives $\partial Z'/\partial \lambda_i$ are still to be set strictly equal to zero. (Why?)

Each line of (13.10) relates to a different type of variable. But we can consolidate the last two lines and, in the process, eliminate the dummy variable s_i from the first-order condition. Inasmuch as $\partial Z'/\partial s_i = -\lambda_i$, the second line of (13.10) tells us that we must have $-\lambda_i \le 0$, $s_i \ge 0$, and $-s_i\lambda_i = 0$, or equivalently,

$$s_i \ge 0 \qquad \lambda_i \ge 0 \qquad \text{and} \qquad s_i\lambda_i = 0 \qquad \textbf{(13.11)}$$

But the third line—a restatement of the constraints in (13.8′)—means that $s_i = r_i - g^i(x_1, x_2, x_3)$. By substituting the latter into (13.11), therefore, we can combine the second and third lines of (13.10) into

$$r_i - g^i(x_1, x_2, x_3) \ge 0 \qquad \lambda_i \ge 0 \qquad \text{and} \qquad \lambda_i[r_i - g^i(x_1, x_2, x_3)] = 0$$

This enables us to express the first-order condition (13.10) in an equivalent form *without* the dummy variables. Using the symbol g^i_j to denote $\partial g^i/\partial x_j$, we now write

$$\begin{array}{llll} \dfrac{\partial Z'}{\partial x_j} = f_j - \left(\lambda_1 g^1_j + \lambda_2 g^2_j\right) \le 0 & x_j \ge 0 & \text{and} & x_j \dfrac{\partial Z'}{\partial x_j} = 0 \\ r_i - g^i(x_1, x_2, x_3) \ge 0 & \lambda_i \ge 0 & \text{and} & \lambda_i[r_i - g^i(x_1, x_2, x_3)] = 0 \end{array} \qquad \textbf{(13.12)}$$

These, then, are the Kuhn-Tucker conditions for problem (13.8), or, more accurately, one version of the Kuhn-Tucker conditions, expressed in terms of the Lagrangian function Z' in (13.9).

Now that we know the results, though, it is possible to obtain the same set of conditions more directly by using a different Lagrangian function. Given the problem (13.9), let us ignore the nonnegativity restrictions as well as the inequality signs in the constraints and write the purely classical type of Lagrangian function Z:

$$Z = f(x_1, x_2, x_3) + \lambda_1[r_1 - g^1(x_1, x_2, x_3)] + \lambda_2[r_2 - g^2(x_1, x_2, x_3)] \qquad \textbf{(13.13)}$$

Then let us do the following: (1) set the partial derivatives $\partial Z/\partial x_j \le 0$, but $\partial Z/\partial \lambda_i \ge 0$, (2) impose nonnegativity restrictions on x_j and λ_i, and (3) require complementary slackness to prevail between each variable and the partial derivative of Z with respect to that variable, that is, require their product to vanish. Since the results of these steps, namely,

$$\begin{aligned} &\frac{\partial Z}{\partial x_j} = f_j - \left(\lambda_1 g_j^1 + \lambda_2 g_j^2\right) \le 0 \qquad x_j \ge 0 \qquad \text{and} \qquad x_j \frac{\partial Z}{\partial x_j} = 0 \\ &\frac{\partial Z}{\partial \lambda_i} = r_i - g^i(x_1, x_2, x_3) \ge 0 \qquad \lambda_i \ge 0 \qquad \text{and} \qquad \lambda_i \frac{\partial Z}{\partial \lambda_i} = 0 \end{aligned} \qquad \textbf{(13.14)}$$

are identical with (13.12), the Kuhn-Tucker conditions are expressible also in terms of the Lagrangian function Z (as against Z'). Note that, by switching from Z' to Z, we can not only arrive at the Kuhn-Tucker conditions more directly, but also identify the expression $r_i - g^i(x_1, x_2, x_3)$—which was left nameless in (13.12)—as the partial derivative $\partial Z/\partial \lambda_i$. In the subsequent discussion, therefore, we shall only use the (13.14) version of the Kuhn-Tucker conditions, based on the Lagrangian function Z.

Example 1 If we cast the familiar problem of utility maximization into the nonlinear programming mold, we may have a problem with an inequality constraint as follows:

$$\begin{aligned} &\text{Maximize} && U = U(x, y) \\ &\text{subject to} && P_x x + P_y y \le B \\ &\text{and} && x, y \ge 0 \end{aligned}$$

Note that, with the inequality constraint, the consumer is no longer required to spend the entire amount B.

To add a new twist to the problem, however, let us suppose that a ration has been imposed on commodity x equal to X_0. Then the consumer would face a second constraint, and the problem changes to

$$\begin{aligned} &\text{Maximize} && U = U(x, y) \\ &\text{subject to} && P_x x + P_y y \le B \\ & && x \le X_0 \\ &\text{and} && x, y \ge 0 \end{aligned}$$

The Lagrangian function is

$$Z = U(x, y) + \lambda_1(B - P_x x - P_y y) + \lambda_2(X_0 - x)$$

and the Kuhn-Tucker conditions are

$$\begin{aligned} &Z_x = U_x - P_x\lambda_1 - \lambda_2 \le 0 && x \ge 0 && \text{and} && xZ_x = 0 \\ &Z_y = U_y - P_y\lambda_1 \le 0 && y \ge 0 && \text{and} && yZ_y = 0 \\ &Z_{\lambda_1} = B - P_x y - P_y y \ge 0 && \lambda_1 \ge 0 && \text{and} && \lambda_1 Z_{\lambda_1} = 0 \\ &Z_{\lambda_2} = X_0 - x \ge 0 && \lambda_2 \ge 0 && \text{and} && \lambda_2 Z_{\lambda_2} = 0 \end{aligned}$$

It is useful to examine the implications of the third column of the Kuhn-Tucker conditions. The condition $\lambda_1 Z_{\lambda_1} = 0$, in particular, requires that

$$\lambda_1(B - P_x x - P_y y) = 0$$

Therefore, we must have either

$$\lambda_1 = 0 \qquad \text{or} \qquad B - P_x x - P_y y = 0$$

If we interpret λ_1 as the marginal utility of budget money (income), and if the budget constraint is nonbinding (satisfied as an inequality in the solution, with money left over), the marginal utility of B should be zero ($\lambda_1 = 0$).

Similarly, the condition $\lambda_2 \, Z_{\lambda_2} = 0$ requires that either

$$\lambda_2 = 0 \qquad \text{or} \qquad X_0 - x = 0$$

Since λ_2 can be interpreted as the marginal utility of relaxing the constraint, we see that if the ration constraint is nonbinding, the marginal utility of relaxing the constraint should be zero ($\lambda_2 = 0$).

This feature, referred to as complementary slackness, plays an essential role in the search for a solution. We shall now illustrate this with a numerical example:

$$\begin{aligned} &\text{Maximize} && U = xy \\ &\text{subject to} && x + y \le 100 \\ & && x \le 40 \\ &\text{and} && x, y \ge 0 \end{aligned}$$

The Lagrangian is

$$Z = xy + \lambda_1(100 - x - y) + \lambda_2(40 - x)$$

and the Kuhn-Tucker conditions become

$$\begin{aligned} &Z_x = y - \lambda_1 - \lambda_2 \le 0 && x \ge 0 && \text{and} && xZ_x = 0 \\ &Z_y = x - \lambda_1 \le 0 && y \ge 0 && \text{and} && yZ_y = 0 \\ &Z_{\lambda_1} = 100 - x - y \ge 0 && \lambda_1 \ge 0 && \text{and} && \lambda_1 Z_{\lambda_1} = 0 \\ &Z_{\lambda_2} = 40 - x \ge 0 && \lambda_2 \ge 0 && \text{and} && \lambda_2 Z_{\lambda_2} = 0 \end{aligned}$$

To solve a nonlinear programming problem, the typical approach is one of trial and error. We can, for example, start by trying a zero value for a choice variable. Setting a variable equal to zero always simplifies the marginal conditions by causing certain terms to drop out. If appropriate nonnegative values of Lagrange multipliers can then be found that satisfy all the marginal inequalities, the zero solution will be optimal. If, on the other hand, the zero solution violates some of the inequalities, then we must let one or more choice variables be positive. For every positive choice variable, we may, by complementary slackness, convert a weak inequality marginal condition into a strict equality. Properly solved, such an equality will lead us either to a solution, or to a contradiction that would then compel us to try something else. If a solution exists, such trials will eventually enable us to uncover it. We can also start by assuming one of the constraints to be nonbinding. Then the related Lagrange multiplier will be zero by complementary slackness and we have thus eliminated a variable. If this assumption leads to a contradiction, then we must treat the said constraint as a strict equality and proceed on that basis.

For the present example, it makes no sense to try $x = 0$ or $y = 0$, for then we would have $U = xy = 0$. We therefore assume both x and y to be nonzero, and deduce $Z_x = Z_y = 0$ from complementary slackness. This means

$$y - \lambda_1 - \lambda_2 = x - \lambda_1 (= 0)$$

so that

$$y - \lambda_2 = x.$$

Now, assume the ration constraint to be nonbinding in the solution, which implies that $\lambda_2 = 0$. Then we have $x = y$, and the given budget $B = 100$ yields the trial solution $x = y = 50$. But this solution violates the ration constraint $x \le 40$. Hence we must adopt the alternative assumption that the ration constraint is binding with $x^* = 40$. The budget constraint then allows the consumer to have $y^* = 60$. Moreover, since complementary slackness dictates that $Z_x = Z_y = 0$, we can readily calculate that $\lambda_1^* = 40$, and $\lambda_2^* = 20$.

Interpretation of the Kuhn-Tucker Conditions

Parts of the Kuhn-Tucker conditions (13.14) are merely a restatement of certain aspects of the given problem. Thus the conditions $x_j \ge 0$ merely repeat the nonnegativity restrictions, and the conditions $\partial Z/\partial \lambda_i \ge 0$ merely reiterate the constraints. To include these in (13.14), however, has the important advantage of revealing more clearly the remarkable symmetry between the two types of variables, x_j (choice variable) and λ_i (Lagrange multipliers). To each variable in each category, there corresponds a marginal condition—$\partial Z/\partial x_j \le 0$ or $\partial Z/\partial \lambda_i \ge 0$—to be satisfied by the optimal solution. Each of the variables must be nonnegative as well. And, finally, each variable is characterized by complementary slackness in relation to a particular partial derivative of the Lagrangian function Z. This means that, for each x_j, we must find in the optimal solution that *either* the marginal condition holds as an equality, as in the classical context, *or* the choice variable in question must take a zero value, *or* both. Analogously, for each λ_i, we must find in the optimal solution that *either* the marginal condition holds as an equality—meaning that the ith constraint is exactly satisfied—*or* the Lagrange multiplier vanishes, *or* both.

An even more explicit interpretation is possible when we look at the expanded expressions for $\partial Z/\partial x_j$ and $\partial Z/\partial \lambda_i$ in (13.14). Assume the problem to be the familiar production problem. Then we have

$f_j \equiv$ marginal gross profit of jth product

$\lambda_i \equiv$ shadow price of ith resource (the opportunity cost of using a unit of the ith resource)

$g_j^i \equiv$ amount of ith resource used up in producing the marginal unit of jth product

$\lambda_i g_j^i \equiv$ marginal imputed cost of ith resource incurred in producing a unit of jth product

$\sum_i \lambda_i g_j^i \equiv$ aggregate marginal imputed cost of jth product

Thus the marginal condition

$$\frac{\partial Z}{\partial x_j} = f_j - \sum_i \lambda_i g_j^i \le 0$$

requires that the marginal gross profit of the jth product be no greater than its aggregate marginal imputed cost; i.e., no *under*imputation is permitted. The complementary-slackness condition then means that, if the optimal solution calls for the active production of the jth product ($x_j^* > 0$), the marginal gross profit must be exactly equal to the aggregate marginal imputed cost ($\partial Z/\partial x_j^* = 0$), as would be the situation in the classical optimization problem. If, on the other hand, the marginal gross profit optimally falls short of the aggregate imputed cost ($\partial Z/\partial x_j^* < 0$), entailing *excess* imputation, then that product must not be produced ($x_j^* = 0$).[†] This latter situation is something that can never occur in the classical context, for if the marginal gross profit is less than the marginal imputed cost, then the output should in that framework be reduced all the way to the level where the marginal condition is satisfied as an equality. What causes the situation of $\partial Z/\partial x_j^* < 0$ to qualify as an optimal one here, is the explicit specification of nonnegativity in the present framework. For then the most we can do in the way of output reduction is to lower production to the level $x_j^* = 0$, and if we still find $\partial Z/\partial x_j^* < 0$ at the zero output, we stop there anyway.

As for the remaining conditions, which relate to the variables λ_i, their meanings are even easier to perceive. First of all, the marginal condition $\partial Z/\partial \lambda_i \geq 0$ merely requires the firm to stay within the capacity limitation of every resource in the plant. The complementary-slackness condition then stipulates that, if the *i*th resource is not fully used in the optimal solution ($\partial Z/\partial \lambda_i^* > 0$), the shadow price of that resource—which is never allowed to be negative—must be set equal to zero ($\lambda_i^* = 0$). On the other hand, if a resource has a positive shadow price in the optimal solution ($\lambda_i^* > 0$), then it is perforce a fully utilized resource ($\partial Z/\partial \lambda_i^* = 0$).

It is also possible, of course, to take the Lagrange-multiplier value λ_i^* to be a measure of how the optimal value of the objective function reacts to a slight relaxation of the *i*th constraint. In that light, complementary slackness would mean that, if the *i*th constraint is optimally not binding ($\partial Z/\partial \lambda_i^* > 0$), then relaxing that particular constraint will not affect the optimal value of the gross profit ($\lambda_i^* = 0$)—just as loosening a belt which is not constricting one's waist to begin with will not produce any greater comfort. If, on the other hand, a slight relaxation of the *i*th constraint (increasing the endowment of the *i*th resource) does increase the gross profit ($\lambda_i^* > 0$), then that resource constraint must in fact be binding in the optimal solution ($\partial Z/\partial \lambda_i^* = 0$).

The *n*-Variable, *m*-Constraint Case

The preceding discussion can be generalized in a straightforward manner to when there are *n* choice variables and *m* constraints. The Lagrangian function *Z* will appear in the more general form

$$Z = f(x_1, x_2, \ldots, x_n) + \sum_{i=1}^{m} \lambda_i [r_i - g^i(x_1, x_2, \ldots, x_n)] \qquad \textbf{(13.15)}$$

And the Kuhn-Tucker conditions will simply be

$$\begin{aligned} &\frac{\partial Z}{\partial x_j} \leq 0 \qquad x_j \geq 0 \qquad \text{and} \qquad x_j \frac{\partial Z}{\partial x_j} = 0 \qquad \text{[maximization]} \\ &\frac{\partial Z}{\partial \lambda_i} \geq 0 \qquad \lambda_i \geq 0 \qquad \text{and} \qquad \lambda_i \frac{\partial Z}{\partial \lambda_i} = 0 \qquad \begin{pmatrix} i = 1, 2, \ldots, m \\ j = 1, 2, \ldots, n \end{pmatrix} \end{aligned} \qquad \textbf{(13.16)}$$

Here, in order to avoid a cluttered appearance, we have not written out the expanded expressions for the partial derivatives $\partial Z/\partial x_j$ and $\partial Z/\partial \lambda_i$. But you are urged to write them out for a more detailed view of the Kuhn-Tucker conditions, similar to what was given in (13.14). Note that, aside from the change in the dimension of the problem, the Kuhn-Tucker conditions remain entirely the same. The interpretation of these conditions should naturally also remain the same.

[†] Remember that, given the equation $ab = 0$, where *a* and *b* are real numbers, we can legitimately infer that $a \neq 0$ implies $b = 0$, but it is not true that $a = 0$ implies $b \neq 0$, since $b = 0$ is also consistent with $a = 0$.

What if the problem is one of *minimization?* One way of handling it is to convert the problem into a maximization problem and then apply (13.6). To minimize C is equivalent to *maximizing* $-C$, so such a conversion is always feasible. But we must, of course, also reverse the constraint inequalities by multiplying every constraint through by -1. Instead of going through the conversion process, however, we may—again using the Lagrangian function Z as defined in (13.15)—directly apply the minimization version of the Kuhn-Tucker conditions as follows:

$$\begin{array}{llll} \dfrac{\partial Z}{\partial x_j} \geq 0 & x_j \geq 0 & \text{and} \quad x_j \dfrac{\partial Z}{\partial x_j} = 0 & \text{[minimization]} \\ \dfrac{\partial Z}{\partial \lambda_i} \leq 0 & \lambda_i \geq 0 & \text{and} \quad \lambda_i \dfrac{\partial Z}{\partial \lambda_i} = 0 & \begin{pmatrix} i = 1, 2, \ldots, m \\ j = 1, 2, \ldots, n \end{pmatrix} \end{array} \tag{13.17}$$

This you should compare with (13.16).

Reading (13.16) and (13.17) horizontally (*row*wise), we see that the Kuhn-Tucker conditions for both maximization and minimization problems consist of a set of conditions relating to the choice variables x_j (first row) and another set relating to the Lagrange multipliers λ_i (second row). Reading them vertically (*column*wise) on the other hand, we note that, for each x_j and λ_i, there is a marginal condition (first column), a nonnegativity restriction (second column), and a complementary-slackness condition (third column). In any given problem, the marginal conditions pertaining to the choice variables always differ, as a group, from the marginal conditions for the Lagrange multipliers in the sense of inequality they take.

Subject to the proviso to be explained in Sec. 13.2, the Kuhn-Tucker maximum conditions (13.16) and minimum conditions (13.17) are necessary conditions for a local maximum and local minimum, respectively. But since a global maximum (minimum) must also be a local maximum (minimum), the Kuhn-Tucker conditions can also be taken as necessary conditions for a global maximum (minimum), subject to the same proviso.

Example 2

Let us apply the Kuhn-Tucker conditions to solve a minimization problem:

$$\begin{array}{ll} \text{Minimize} & C = (x_1 - 4)^2 + (x_2 - 4)^2 \\ \text{subject to} & 2x_1 + 3x_2 \geq 6 \\ & -3x_1 - 2x_2 \geq -12 \\ \text{and} & x_1, x_2 \geq 0 \end{array}$$

The Lagrangian function for this problem is

$$Z = (x_1 - 4)^2 + (x_2 - 4)^2 + \lambda_1(6 - 2x_1 - 3x_2) + \lambda_2(-12 + 3x_1 + 2x_2)$$

Since the problem is one of minimization, the appropriate conditions are (13.17), which include the four marginal conditions

$$\begin{aligned} \frac{\partial Z}{\partial x_1} &= 2(x_1 - 4) - 2\lambda_1 + 3\lambda_2 \geq 0 \\ \frac{\partial Z}{\partial x_2} &= 2(x_2 - 4) - 3\lambda_1 + 2\lambda_2 \geq 0 \\ \frac{\partial Z}{\partial \lambda_1} &= 6 - 2x_1 - 3x_2 \leq 0 \\ \frac{\partial Z}{\partial \lambda_2} &= -12 + 3x_1 + 2x_2 \leq 0 \end{aligned} \tag{13.18}$$

plus the nonnegativity and complementary-slackness conditions.

To find a solution, we again use the trial-and-error approach, realizing that the first few trials may lead us into a blind alley. Suppose we first try $\lambda_1 > 0$ and $\lambda_2 > 0$ and check whether we can find corresponding x_1 and x_2 values that satisfy both constraints. With positive Lagrange multipliers, we must have $\partial Z/\partial\lambda_1 = \partial Z/\partial\lambda_2 = 0$. From the last two lines of (13.18), we can thus write

$$2x_1 + 3x_2 = 6 \qquad \text{and} \qquad 3x_1 + 2x_2 = 12$$

These two equations yield the trial solution $x_1 = 4\frac{4}{5}$ and $x_2 = -1\frac{1}{5}$, which violates the nonnegativity restriction on x_2.

Let us next try $x_1 > 0$ and $x_2 > 0$, which would imply $\partial Z/\partial x_1 = \partial Z/\partial x_2 = 0$ by complementary slackness. Then, from the first two lines of (13.18), we can write

$$2(x_1 - 4) - 2\lambda_1 + 3\lambda_2 = 0 \qquad \text{and} \qquad 2(x_2 - 4) - 3\lambda_1 + 2\lambda_2 = 0 \qquad \textbf{(13.19)}$$

Multiplying the first equation by 2, and the second equation by 3, then subtracting the latter from the former, we can eliminate λ_2 and obtain the result

$$4x_1 - 6x_2 + 5\lambda_1 + 8 = 0$$

By further assuming $\lambda_1 = 0$, we can derive the following relationship between x_1 and x_2:

$$x_1 - \frac{3}{2}x_2 = -2 \qquad \textbf{(13.20)}$$

In order to solve for the two variables, however, we need another relationship between x_1 and x_2. For this purpose, let us assume that $\lambda_2 \neq 0$, so that $\partial Z/\partial\lambda_2 = 0$. Then, from the last two lines of (13.18), we can write (after rearrangement)

$$3x_1 + 2x_2 = 12 \qquad \textbf{(13.21)}$$

Together, (13.20) and (13.21) yield another trial solution

$$x_1 = \frac{28}{13}\left(= 2\frac{2}{13}\right) > 0 \qquad x_2 = \frac{36}{13}\left(= 2\frac{10}{13}\right) > 0$$

Substituting these values into (13.19), and solving for the Lagrange multipliers, we get

$$\lambda_1 = 0 \qquad \lambda_2 = \frac{16}{13}\left(= 1\frac{3}{13}\right) > 0$$

Since the solution values for the four variables are all nonnegative and satisfy both constraints, they are acceptable as the final solution.

EXERCISE 13.1

1. Draw a set of diagrams similar to those in Fig. 13.1 for the minimization case, and deduce a set of necessary conditions for a local minimum corresponding to (13.2) through (13.4). Then condense these conditions into a single statement similar to (13.5).
2. (*a*) Show that, in (13.16), instead of writing

   $$\lambda_i \frac{\partial Z}{\partial \lambda_i} = 0 \qquad (i = 1, \ldots, m)$$

   as a set of m separate conditions, it is sufficient to write a single equation in the form of

   $$\sum_{i=1}^{m} \lambda_i \frac{\partial Z}{\partial \lambda_i} = 0$$

(*b*) Can we do the same for the following set of conditions?

$$x_j \frac{\partial Z}{\partial x_j} = 0 \qquad (j = 1, \ldots, n)$$

3. Based on the reasoning used in Prob. 2, which set (or sets) of conditions in (13.17) can be condensed into a single equation?
4. Suppose the problem is

   Minimize $\quad C = f(x_1, x_2, \ldots, x_n)$

   subject to $\quad g^i(x_1, x_2, \ldots, x_n) \geq r_i$

   and $\quad x_j \geq 0 \quad \begin{pmatrix} i = 1, 2, \ldots, m \\ j = 1, 2, \ldots, n \end{pmatrix}.$

   Write the Lagrangian function, take the derivatives $\partial Z/\partial x_j$ and $\partial Z/\partial \lambda_i$ and write out the expanded version of the Kuhn-Tucker minimum conditions (13.17).
5. Convert the minimization problem in Prob. 4 into a maximization problem, formulate the Lagrangian function, take the derivatives with respect to x_j and λ_i, and apply the Kuhn-Tucker maximum conditions (13.16). Are the results consistent with those obtained in Prob. 4?

13.2 The Constraint Qualification

The Kuhn-Tucker conditions are necessary conditions *only if* a particular proviso is satisfied. That proviso, called the *constraint qualification,* imposes a certain restriction on the constraint functions of a nonlinear programming problem, for the specific purpose of ruling out certain irregularities on the boundary of the feasible set, that would invalidate the Kuhn-Tucker conditions should the optimal solution occur there.

Irregularities at Boundary Points

Let us first illustrate the nature of such irregularities by means of some concrete examples.

Example 1

Maximize $\quad \pi = x_1$

subject to $\quad x_2 - (1 - x_1)^3 \leq 0$

and $\quad x_1, x_2 \geq 0$

As shown in Fig. 13.2, the feasible region is the set of points that lie in the first quadrant on or below the curve $x_2 = (1 - x_1)^3$. Since the objective function directs us to maximize x_1, the optimal solution is the point (1, 0). But the solution fails to satisfy the Kuhn-Tucker maximum conditions. To check this, we first write the Lagrangian function

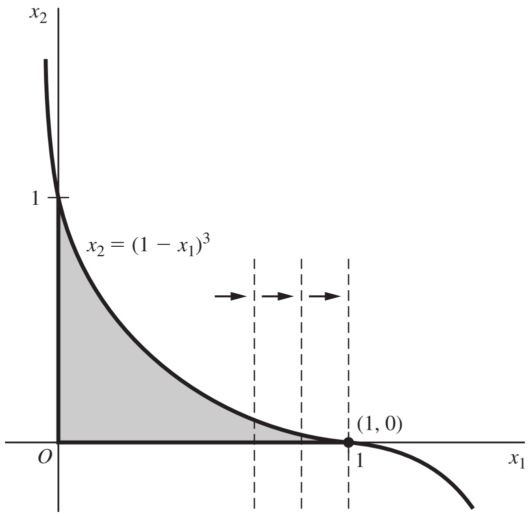
$$Z = x_1 + \lambda_1[-x_2 + (1 - x_1)^3]$$

As the first marginal condition, we should then have

$$\frac{\partial Z}{\partial x_1} = 1 - 3\lambda_1(1 - x_1)^2 \leq 0$$

In fact, since $x_1^* = 1$ is positive, complementary slackness requires that this derivative vanish when evaluated at the point (1, 0). However, the actual value we get happens to be $\partial Z/\partial x_1^* = 1$, thus violating the given marginal condition.

FIGURE 13.2



The reason for this anomaly is that the optimal solution (1, 0) occurs in this example at an outward-pointing *cusp,* which constitutes one type of irregularity that can invalidate the Kuhn-Tucker conditions at a boundary optimal solution. A *cusp* is a sharp point formed when a curve takes a sudden reversal in direction, such that the slope of the curve on one side of the point is the same as the slope of the curve on the other side of the point. Here, the boundary of the feasible region at first follows the constraint curve, but when the point (1, 0) is reached, it takes an abrupt turn westward and follows the trail of the horizontal axis thereafter. Since the slopes of both the curved side and the horizontal side of the boundary are zero at the point (1, 0), that point is a cusp.

Cusps are the most frequently cited culprits for the failure of the Kuhn-Tucker conditions, but the truth is that the presence of a cusp is neither necessary nor sufficient to cause those conditions to fail at an optimal solution. Examples 2 and 3 will confirm this.

Example 2

To the problem of Example 1, let us add a new constraint

$$2x_1 + x_2 \le 2$$

whose border, $x_2 = 2 - 2x_1$, plots as a straight line with slope -2 which passes through the optimal point in Fig. 13.2. Clearly, the feasible region remains the same as before, and so does the optimal solution at the cusp. But if we write the new Lagrangian function

$$Z = x_1 + \lambda_1[-x_2 + (1 - x_1)^3] + \lambda_2[2 - 2x_1 - x_2]$$

and the marginal conditions

$$\frac{\partial Z}{\partial x_1} = 1 - 3\lambda_1(1 - x_1)^2 - 2\lambda_2 \le 0$$

$$\frac{\partial Z}{\partial x_2} = -\lambda_1 - \lambda_2 \le 0$$

$$\frac{\partial Z}{\partial \lambda_1} = -x_2 + (1 - x_1)^3 \ge 0$$

$$\frac{\partial Z}{\partial \lambda_2} = 2 - 2x_1 - x_2 \ge 0$$

it turns out that the values $x_1^* = 1$, $x_2^* = 0$, $\lambda_1^* = 1$, and $\lambda_2^* = \frac{1}{2}$ do satisfy these four inequalities, as well as the nonnegativity and complementary-slackness conditions. As a matter of fact, λ_1^* can be assigned any nonnegative value (not just 1), and all the conditions can still be satisfied—which goes to show that the optimal value of a Lagrange multiplier is not necessarily unique. More importantly, however, this example shows that the Kuhn-Tucker conditions can remain valid despite the cusp.

Example 3 The feasible region of the problem

$$\begin{aligned} &\text{Maximize} && \pi = x_2 - x_1^2 \\ &\text{subject to} && -\left(10 - x_1^2 - x_2\right)^3 \le 0 \\ & && -x_1 \le -2 \\ &\text{and} && x_1, x_2 \ge 0 \end{aligned}$$

as shown in Fig. 13.3, contains no cusp anywhere. Yet, at the optimal solution, (2, 6), the Kuhn-Tucker conditions nonetheless fail to hold. For, with the Lagrangian function

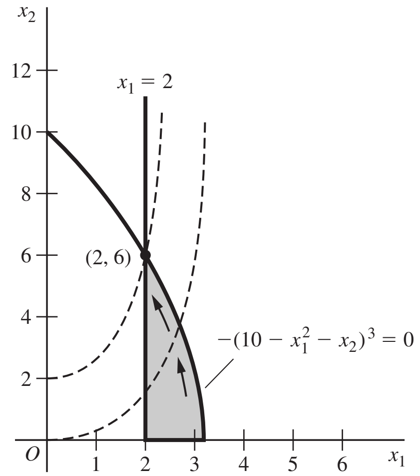
$$Z = x_2 - x_1^2 + \lambda_1\left(10 - x_1^2 - x_2\right)^3 + \lambda_2(-2 + x_1)$$

the second marginal condition would require that

$$\frac{\partial Z}{\partial x_2} = 1 - 3\lambda_1\left(10 - x_1^2 - x_2\right)^2 \le 0$$

Indeed, since x_2^* is positive, this derivative should vanish when evaluated at the point (2, 6). But actually we get $\partial Z/\partial x_2 = 1$, regardless of the value assigned to λ_1. Thus the Kuhn-Tucker conditions can fail even in the absence of a cusp—nay, even when the feasible region is a convex set as in Fig. 13.3. The fundamental reason why cusps are neither necessary nor sufficient for the failure of the Kuhn-Tucker conditions is that the preceding irregularities referred to before relate, not to the shape of the feasible region per se, but to the forms of the constraint functions themselves.

FIGURE 13.3



The Constraint Qualification

Boundary irregularities—cusp or no cusp—will not occur if a certain constraint qualification is satisfied.

To explain this, let $x^* \equiv (x_1^*, x_2^*, \ldots, x_n^*)$ be a boundary point of the feasible region and a possible candidate for a solution, and let $dx \equiv (dx_1, dx_2, \ldots, dx_n)$ represent a particular direction of movement from the said boundary point. The direction-of-movement interpretation of the vector dx is perfectly in line with our earlier interpretation of a vector as a directed line segment (an arrow), but here, the point of departure is the point x^* instead of the point of origin, and so the vector dx is *not* in the nature of a radius vector. We shall now impose two requirements on the vector dx. First, if the jth choice variable has a zero value at the point x^*, then we shall only permit a nonnegative change on the x_j axis, that is,

$$dx_j \geq 0 \qquad \text{if} \qquad x_j^* = 0 \tag{13.22}$$

Second, if the ith constraint is exactly satisfied at the point x^*, then we shall only allow values of $dx_1, \ldots, dx_n$ such that the value of the constraint function $g^i(x^*)$ will not increase (for a maximization problem) or will not decrease (for a minimization problem), that is,

$$dg^i(x^*) = g_1^i\,dx_1 + g_2^i\,dx_2 + \cdots + g_n^i\,dx_n \quad \begin{cases} \leq 0 \text{ (max.)} \\ \geq 0 \text{ (min.)} \end{cases} \qquad \text{if} \qquad g^i(x^*) = r_i \tag{13.23}$$

where all the partial derivatives of g_j^i are to be evaluated at x^*. If a vector dx satisfies (13.22) and (13.23), we shall refer to it as a *test vector*. Finally, if there exists a differentiable arc that (1) emanates from the point x^*, (2) is contained entirely in the feasible region, and (3) is tangent to a given test vector, we shall call it a *qualifying arc* for that test vector. With this background, the constraint qualification can be stated simply as follows:

> The constraint qualification is satisfied if, for any point x^* on the boundary of the feasible region, there exists a qualifying arc for every test vector dx.

Example 4

We shall show that the optimal point (1, 0) of Example 1 in Fig. 13.2, which fails the Kuhn-Tucker conditions, also fails the constraint qualification. At that point, $x_2^* = 0$; thus the test vector must satisfy

$$dx_2 \geq 0 \qquad \text{[by (13.22)]}$$

Moreover, since the (only) constraint, $g^1 = x_2 - (1 - x_1)^3 \leq 0$, is exactly satisfied at (1, 0), we must let [by (13.23)]

$$g_1^1\,dx_1 + g_2^1\,dx_2 = 3(1 - x_1^*)^2\,dx_1 + dx_2 = dx_2 \leq 0$$

These two requirements together imply that we must let $dx_2 = 0$. In contrast, we are free to choose dx_1. Thus, for instance, the vector $(dx_1, dx_2) = (2, 0)$ is an acceptable test vector, as is $(dx_1, dx_2) = (-1, 0)$. The latter test vector would plot in Fig. 13.2 as an arrow starting from (1, 0) and pointing in the due-west direction (not drawn), and it is clearly possible to draw a qualifying arc for it. (The curved boundary of the feasible region itself can serve as a qualifying arc.) On the other hand, the test vector $(dx_1, dx_2) = (2, 0)$ would plot as an arrow starting from (1, 0) and pointing in the due-east direction (not drawn). Since there is no way to draw a smooth arc tangent to this vector and lying entirely within the feasible region, no qualifying arcs exist for it. Hence the optimal solution point (1, 0) violates the constraint qualification.

Example 5

Referring to Example 2, let us illustrate that, after an additional constraint $2x_1 + x_2 \leq 2$ is added to Fig. 13.2, the point (1, 0) will satisfy the constraint qualification, thereby revalidating the Kuhn-Tucker conditions.

As in Example 4, we have to require $dx_2 \geq 0$ (because $x_2^* = 0$) and $dx_2 \leq 0$ (because the first constraint is exactly satisfied); thus, $dx_2 = 0$. But the second constraint is also exactly satisfied, thereby requiring

$$g_1^2\, dx_1 + g_2^2\, dx_2 = 2dx_1 + dx_2 = 2dx_1 \leq 0 \qquad \text{[by (13.23)]}$$

With nonpositive dx_1 and zero dx_2, the only admissible test vectors—aside from the null vector itself—are those pointing in the due-west direction in Fig. 13.2 from (1, 0). All of these lie along the horizontal axis in the feasible region, and it is certainly possible to draw a qualifying arc for each test vector. Hence, this time the constraint qualification indeed is satisfied.

Linear Constraints

Earlier, in Example 3, it was demonstrated that the convexity of the feasible set does not guarantee the validity of the Kuhn-Tucker conditions as necessary conditions. However, if the feasible region is a convex set formed by *linear* constraints only, then the constraint qualification will invariably be met, and the Kuhn-Tucker conditions will always hold at an optimal solution. This being the case, we need never worry about boundary irregularities when dealing with a nonlinear programming problem with linear constraints.

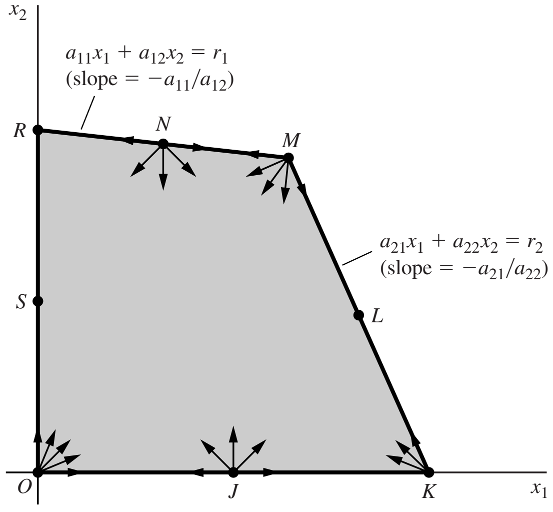
Example 6

Let us illustrate the linear-constraint result in the two-variable, two-constraint framework. For a maximization problem, the linear constraints can be written as

$$a_{11}x_1 + a_{12}x_2 \leq r_1$$
$$a_{21}x_1 + a_{22}x_2 \leq r_2$$

where we shall take all the parameters to be positive. Then, as indicated in Fig. 13.4, the first constraint border will have a slope of $-a_{11}/a_{12} < 0$, and the second, a slope of $-a_{21}/a_{22} < 0$. The boundary points of the shaded feasible region fall into the following five types: (1) the point of origin, where the two axes intersect, (2) points that lie on one axis segment, such

FIGURE 13.4



as *J* and *S*, (3) points at the intersection of one axis and one constraint border, namely, *K* and *R*, (4) points lying on a single constraint border, such as *L* and *N*, (5) the point of intersection of the two constraints, *M*. We may briefly examine each type in turn with reference to the satisfaction of the constraint qualification.

1. At the origin, no constraint is exactly satisfied, so we may ignore (13.23). But since $x_1 = x_2 = 0$, we must choose test vectors with $dx_1 \geq 0$ and $dx_2 \geq 0$, by (13.22). Hence all test vectors from the origin must point in the due-east, due-north, or northeast directions, as depicted in Fig. 13.4. These vectors all happen to fall within the feasible set, and a qualifying arc clearly can be found for each.
2. At a point like *J*, we can again ignore (13.23). The fact that $x_2 = 0$ means that we must choose $dx_2 \geq 0$, but our choice of dx_1 is free. Hence all vectors would be acceptable except those pointing southward ($dx_2 < 0$). Again all such vectors fall within the feasible region, and there exists a qualifying arc for each. The analysis of point *S* is similar.
3. At points *K* and *R*, both (13.22) and (13.23) must be considered. Specifically, at *K*, we have to choose $dx_2 \geq 0$ since $x_2 = 0$, so that we must rule out all southward arrows. The second constraint being exactly satisfied, moreover, the test vectors for point *K* must satisfy

   $$g_1^2\, dx_1 + g_2^2\, dx_2 = a_{21}\, dx_1 + a_{22}\, dx_2 \leq 0 \qquad \textbf{(13.24)}$$

   Since at *K* we also have $a_{21}x_1 + a_{22}x_2 = r_2$ (second constraint border), however, we may add this equality to (13.24) and modify the restriction on the test vector to the form

   $$a_{21}(x_1 + dx_1) + a_{22}(x_2 + dx_2) \leq r_2 \qquad \textbf{(13.24')}$$

   Interpreting $(x_j + dx_j)$ to be the new value of x_j attained at the arrowhead of a test vector, we may construe (13.24′) to mean that all test vectors must have their arrowheads located on or below the second constraint border. Consequently, all these vectors must again fall within the feasible region, and a qualifying arc can be found for each. The analysis of point *R* is analogous.
4. At points such as *L* and *N*, neither variable is zero and (13.22) can be ignored. However, for point *N*, (13.23) dictates that

   $$g_1^1\, dx_1 + g_2^1\, dx_2 = a_{11}\, dx_1 + a_{12}\, dx_2 \leq 0 \qquad \textbf{(13.25)}$$

   Since point *N* satisfies $a_{11}\, dx_1 + a_{12}\, dx_2 = r_1$ (first constraint border), we may add this equality to (13.25) and write

   $$a_{11}(x_1 + dx_1) + a_{12}(x_2 + dx_2) \leq r_1 \qquad \textbf{(13.25')}$$

   This would require the test vectors to have arrowheads located on or below the first constraint border in Fig. 13.4. Thus we obtain essentially the same kind of result encountered in the other cases. This analysis of point *L* is analogous.
5. At point *M*, we may again disregard (13.22), but this time (13.23) requires all test vectors to satisfy both (13.24) and (13.25). Since we may modify the latter conditions to the forms in (13.24′) and (13.25′), all test vectors must now have their arrowheads located on or below the first as well as the second constraint borders. The result thus again duplicates those of the previous cases.

In this example, it so happens that, for every type of boundary point considered, the test vectors all lie within the feasible region. While this locational feature makes the qualifying arcs easy to find, it is by no means a prerequisite for their existence. In a problem with a nonlinear constraint border, in particular, the constraint border itself may serve as a qualifying arc for some test vector that lies outside of the feasible region. An example of this can be found in one of the problems below.

EXERCISE 13.2

1. Check whether the solution point $(x_1^*, x_2^*) = (2, 6)$ in Example 3 satisfies the constraint qualification.
2. Maximize $\pi = x_1$
   subject to $x_1^2 + x_2^2 \leq 1$
   and $x_1, x_2 \geq 0$
   Solve graphically and check whether the optimal-solution point satisfies (*a*) the constraint qualification and (*b*) the Kuhn-Tucker conditions.
3. Minimize $C = x_1$
   subject to $x_1^2 - x_2 \geq 0$
   and $x_1, x_2 \geq 0$
   Solve graphically. Does the optimal solution occur at a cusp? Check whether the optimal solution satisfies (*a*) the constraint qualification and (*b*) the Kuhn-Tucker minimum conditions.
4. Minimize $C = x_1$
   subject to $-x_2 - (1 - x_1)^3 \geq 0$
   and $x_1, x_2 \geq 0$
   Show that (*a*) the optimal solution $(x_1^*, x_2^*) = (1, 0)$ does not satisfy the Kuhn-Tucker conditions, but (*b*) by introducing a new multiplier $\lambda_0 \geq 0$, and modifying the Lagrangian function (13.15) to the form
   $$Z_0 = \lambda_0 f(x_1, x_2, \ldots, x_n) + \sum_{i=1}^{m} \lambda_1 [r_i - g^i(x_1, x_2, \ldots, x_n)]$$
   the Kuhn-Tucker conditions can be satisfied at (1, 0). (*Note:* The Kuhn-Tucker conditions on the multipliers extend to only $\lambda_1, \ldots, \lambda_m$, but not to λ_0.)

13.3 Economic Applications

War-Time Rationing

Typically during times of war the civilian population is subject to some form of rationing of basic consumer goods. Usually, the method of rationing is through the use of redeemable coupons used by the government. The government will supply each consumer with an allotment of coupons each month. In turn, the consumer will have to redeem a certain number of coupons at the time of purchase of a rationed good. This effectively means the consumer pays *two* prices at the time of the purchase. He or she pays both the coupon price and the monetary price of the rationed good. This requires the consumer to have both sufficient funds and sufficient coupons in order to buy a unit of the rationed good.

Consider the case of a two-good world where both goods, x and y, are rationed. Let the consumer's utility function be $U = U(x, y)$. The consumer has a fixed money budget of B

and faces exogenous prices P_x and P_y. Further, the consumer has an allotment of coupons, denoted C, which can be used to purchase either x or y at a coupon price of c_x and c_y. Therefore the consumer's maximization problem is

$$\begin{aligned} &\text{Maximize} && U = U(x, y) \\ &\text{subject to} && P_x x + P_y y \le B \\ &&& c_x x + c_y y \le C \\ &\text{and} && x, y \ge 0 \end{aligned}$$

The Lagrangian for the problem is

$$Z = U(x, y) + \lambda_1(B - P_x x - P_y y) + \lambda_2(C - c_x x + c_y y)$$

where λ_1 and λ_2 are the Lagrange multipliers. Since both constraints are linear, the constraint qualification is satisfied and the Kuhn-Tucker conditions are necessary:

$$\begin{array}{lll} Z_x = U_x - \lambda_1 P_x - \lambda_2 c_x \le 0 & x \ge 0 & x Z_x = 0 \\ Z_y = U_y - \lambda_1 P_y - \lambda_2 c_y \le 0 & y \ge 0 & y Z_y = 0 \\ Z_{\lambda_1} = B - P_x x - P_y y \ge 0 & \lambda_1 \ge 0 & \lambda_1 Z_{\lambda_1} = 0 \\ Z_{\lambda_2} = C - c_x x - c_y y \ge 0 & \lambda_2 \ge 0 & \lambda_2 Z_{\lambda_2} = 0 \end{array}$$

Example 1 Suppose the utility function is of the form $U = xy^2$. Further, let $B = 100$ and $P_x = P_y = 1$ while $C = 120$, $c_x = 2$, and $c_y = 1$.

The Lagrangian takes the specific form

$$Z = xy^2 + \lambda_1(100 - x - y) + \lambda_2(120 - 2x - y)$$

The Kuhn-Tucker conditions are now

$$\begin{array}{lll} Z_x = y^2 - \lambda_1 - 2\lambda_2 \le 0 & x \ge 0 & xZ_x = 0 \\ Z_y = 2xy - \lambda_1 - \lambda_2 \le 0 & y \ge 0 & yZ_y = 0 \\ Z_{\lambda_1} = 100 - x - y \ge 0 & \lambda_1 \ge 0 & \lambda_1 Z_{\lambda_1} = 0 \\ Z_{\lambda_2} = 120 - 2x - y \ge 0 & \lambda_2 \ge 0 & \lambda_2 Z_{\lambda_2} = 0 \end{array}$$

Again, the solution procedure involves a certain amount of trial and error. We can first choose one of the constraints to be nonbinding and solve for x and y. Once found, use these values to test if the constraint chosen to be nonbinding is violated. If it is, then redo the procedure choosing another constraint to be nonbinding. If violation of the nonbinding constraint occurs again, then we can assume both constraints bind and the solution is determined only by the constraints.

Step 1: Assume that the second (ration) constraint is nonbinding in the solution, so that $\lambda_2 = 0$ by complementary slackness. But let x, y, and λ_1 be positive so that complementary slackness would give us the following three equations:

$$\begin{aligned} Z_x &= y^2 - \lambda_1 = 0 \\ Z_y &= 2xy - \lambda_1 = 0 \\ Z_{\lambda_1} &= 100 - x - y = 0 \end{aligned}$$

Solving for x and y yields a trial solution

$$x = 33^1/_3 \qquad y = 66^2/_3$$

However, when we substitute these solutions into the coupon constraint we find that

$$2(33^1/_3) + 66^2/_3 = 133^1/_3 > 120$$

This solution violates the coupon constraint, and must be rejected.

Step 2: Now let us reverse the assumptions on λ_1 and λ_2 so that $\lambda_1 = 0$, but let $\lambda_2, x, y > 0$. Then, from the marginal conditions, we have

$$Z_x = y^2 - 2\lambda_2 = 0$$
$$Z_y = 2xy - \lambda_2 = 0$$
$$Z_{\lambda_1} = 120 - 2x - y = 0$$

Solving this system of equations yields another trial solution

$$x = 20 \qquad y = 80$$

which implies that $\lambda_2 = 2xy = 3{,}200$. These solution values, together with $\lambda_1 = 0$, satisfy both the budget and ration constraints. Thus we can accept them as the final solution to the Kuhn-Tucker conditions.

This optimal solution, however, contains a curious abnormality. With the budget constraint binding in the solution, we would normally expect the related Lagrange multiplier to be positive, yet we actually have $\lambda_1 = 0$. Thus, in this example, while the budget constraint is *mathematically* binding (satisfied as a strict equality in the solution), it is *economically* nonbinding (not calling for a positive marginal utility of money).

Peak-Load Pricing

Peak and off-peak pricing and planning problems are commonplace for firms with capacity-constrained production processes. Usually the firm has invested in capacity in order to target a primary market. However there may exist a secondary market in which the firm can often sell its product. Once the capital equipment has been purchased to service the firm's primary market, it is freely available (up to capacity) to be used in the secondary market. Typical examples include schools and universities that build to meet daytime needs (peak), but may offer night-school classes (off-peak); theaters that offer shows in the evening (peak) and matinees (off-peak); and trucking companies that have dedicated routes but may choose to enter "back-haul" markets. Since the capacity cost is a factor in the profit-maximizing decision for the peak market and is already paid, it normally should not be a factor in calculating optimal price and quantity for the smaller, off-peak market. However, if the secondary market's demand is close to the same size as the primary market, capacity constraints may be an issue, especially since it is a common practice to price discriminate and charge lower prices in off-peak periods. Even though the secondary market is smaller than the primary, it is possible that, at the lower (profit-maximizing) price, off-peak demand exceeds capacity. In such cases capacity choices must be made taking both markets into account, making the problem a classic application of nonlinear programming.

Consider a profit-maximizing company that faces two average-revenue curves

$$P_1 = P^1(Q_1) \qquad \text{in the day time (peak period)}$$
$$P_2 = P^2(Q_2) \qquad \text{in the night time (off-peak period)}$$

To operate, the firm must pay b per unit of output, whether it is day or night. Furthermore, the firm must purchase capacity at a cost of c per unit of capacity. Let K denote total capacity measured in units of Q. The firm must pay for capacity, regardless of whether it operates in the off-peak period. Who should be charged for the capacity costs: peak, off-peak, or both sets of customers? The firm's maximization problem becomes

$$\begin{aligned}
&\underset{Q_1,Q_2,K}{\text{Maximize}} && \pi = P_1Q_1 + P_2Q_2 - b(Q_1+Q_2) - cK \\
&\text{subject to} && Q_1 \le K \\
& && Q_2 \le K \\
&\text{where} && P_1 = P^1(Q_1) \\
& && P_2 = P^2(Q_2) \\
&\text{and} && Q_1, Q_2, K \ge 0
\end{aligned}$$

In view that the total revenue for Q_i,

$$R_i \equiv P_iQ_i = P^i(Q_i)Q_i$$

is a function of Q_i alone, we can simplify the statement of the problem to

$$\begin{aligned}
&\text{Maximize} && \pi = R_1(Q_1) + R_2(Q_2) - b(Q_1+Q_2) - cK \\
&\text{subject to} && Q_1 \le K \\
& && Q_2 \le K \\
&\text{and} && Q_1, Q_2, K \ge 0
\end{aligned}$$

Note that both constraints are linear; thus the constraint qualification is satisfied and the Kuhn-Tucker conditions are necessary.

The Lagrangian function is

$$Z = R_1(Q_1) + R_2(Q_2) - b(Q_1+Q_2) - cK + \lambda_1(K - Q_1) + \lambda_2(K - Q_2)$$

and the Kuhn-Tucker conditions are

$$\begin{aligned}
Z_1 &= \text{MR}_1 - b - \lambda_1 \le 0 & Q_1 &\ge 0 & Q_1Z_1 &= 0 \\
Z_2 &= \text{MR}_2 - b - \lambda_2 \le 0 & Q_2 &\ge 0 & Q_2Z_2 &= 0 \\
Z_K &= -c + \lambda_1 + \lambda_2 \le 0 & K &\ge 0 & KZ_K &= 0 \\
Z_{\lambda_1} &= K - Q_1 \ge 0 & \lambda_1 &\ge 0 & \lambda_1 Z_{\lambda_1} &= 0 \\
Z_{\lambda_2} &= K - Q_2 \ge 0 & \lambda_2 &\ge 0 & \lambda_2 Z_{\lambda_2} &= 0
\end{aligned}$$

where MR_i is the marginal revenue of Q_i $(i = 1, 2)$.

The solution procedure again entails trial and error. Let us first assume that Q_1, Q_2, $K > 0$. Then, by complementary slackness, we have

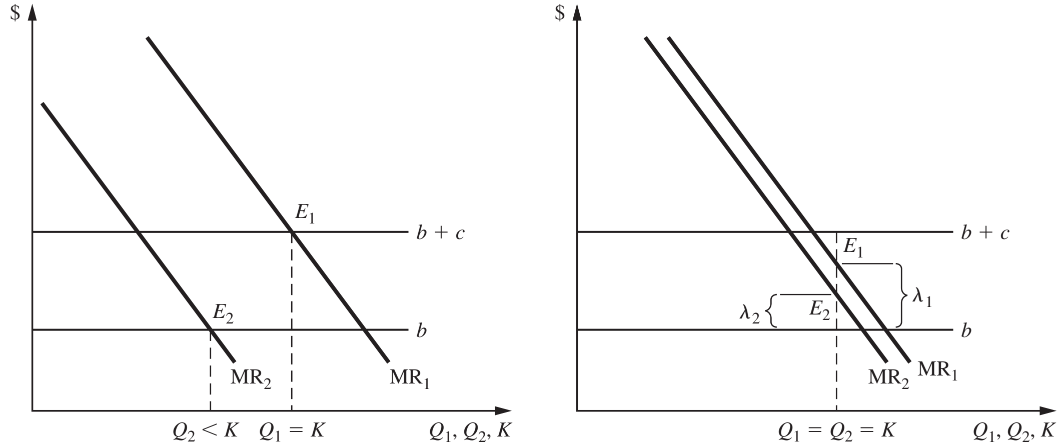
$$\begin{aligned}
\text{MR}_1 - b - \lambda_1 &= 0 \\
\text{MR}_2 - b - \lambda_2 &= 0 \\
-c + \lambda_1 + \lambda_2 &= 0 \quad (\lambda_1 = c - \lambda_2)
\end{aligned} \tag{13.26}$$

which can be condensed into two equations after eliminating λ_1:

$$\begin{aligned}
\text{MR}_1 &= b + c - \lambda_2 \\
\text{MR}_2 &= b + \lambda_2
\end{aligned} \tag{13.26'}$$

Then we proceed in two steps.

FIGURE 13.5



(*a*) Off-peak constraint nonbinding

(*b*) Off-peak constraint binding

Step 1: Since the off-peak market is a secondary market, its marginal-revenue function (MR_2) can be expected to lie below that of the primary market (MR_1) as illustrated in Fig. 13.5. Moreover, the capacity constraint is more likely to be nonbinding in the secondary market so that λ_2 is more likely to be zero. So we try $\lambda_2 = 0$. Then (13.26′) becomes

$$\begin{aligned} \text{MR}_1 &= b + c \\ \text{MR}_2 &= b \end{aligned} \qquad \textbf{(13.26″)}$$

The fact that the primary market absorbs the entire capacity cost c implies that $Q_1 = K$. However, we still need to check whether the constraint $Q_2 \le K$ is satisfied. If so, we have found a valid solution. Figure 13.5(a) illustrates the case where $Q_1 = K$ and $Q_2 < K$ in the solution. The MR_1 curve intersects the $b + c$ line at point E_1, and the MR_2 curve intersects the b line at point E_2.

What if the previous trial solution entails $Q_2 > K$, as would occur if the MR_2 curve is very close to MR_1, so as to intersect the b line at an output larger than K? Then, of course, the second constraint is violated, and we must reject the assumption of $\lambda_2 = 0$, and proceed to the next step.

Step 2: Now let us assume both Lagrange multipliers to be positive, and thus $Q_1 = Q_2 = K$. Then, unable to eliminate any variables from (13.26), we have

$$\begin{aligned} \text{MR}_1 &= b + \lambda_1 \\ \text{MR}_2 &= b + \lambda_2 \\ c &= \lambda_1 + \lambda_2 \end{aligned} \qquad \textbf{(13.26‴)}$$

This case is illustrated in Fig. 13.5(b), where points E_1 and E_2 satisfy the first two equations in (13.26‴). From the third equation, we see that the capacity cost c is the sum of the two Lagrange multipliers. This means λ_1 and λ_2 represent the portions of the capacity cost borne respectively by the two markets.

Example 2 Suppose the average-revenue function during peak hours is

$$P_1 = 22 - 10^{-5} Q_1$$

and that during off-peak hours it is

$$P_2 = 18 - 10^{-5} Q_2$$

To produce a unit of output per half-day requires a unit of capacity costing 8 cents per day. The cost of a unit of capacity is the same whether it is used at peak times only, or off-peak also. In addition to the costs of capacity, it costs 6 cents in operating costs (labor and fuel) to produce 1 unit per half-day (both day and evening).

If we assume that the capacity constraint is nonbinding in the secondary market ($\lambda_2 = 0$), then the given Kuhn-Tucker conditions become

$$\begin{aligned} &\lambda_1 = c = 8 \\ &\underbrace{\begin{matrix} 22 - 2 \times 10^{-5} Q_1 \\ 18 - 2 \times 10^{-5} Q_2 \end{matrix}}_{\text{MR}} \quad \begin{matrix} = b + c \\ = \underbrace{b \qquad}_{\text{MC}} \end{matrix} \quad \begin{matrix} = 14 \\ = 6 \end{matrix} \end{aligned}$$

Solving this system gives us

$$Q_1 = 400{,}000$$
$$Q_2 = 600{,}000$$

which violates the assumption that the second constraint is nonbinding because $Q_2 > Q_1 = K$.

Therefore, let us assume that both constraints are binding. Then $Q_1 = Q_2 = Q$ and the Kuhn-Tucker conditions become

$$\lambda_1 + \lambda_2 = 8$$
$$22 - 2 \times 10^{-5} Q = 6 + \lambda_1$$
$$18 - 2 \times 10^{-5} Q = 6 + \lambda_2$$

which yield the following solution

$$Q_1 = Q_2 = K = 500{,}000$$
$$\lambda_1 = 6 \qquad \lambda_2 = 2$$
$$P_1 = 17 \qquad P_2 = 13$$

Since the capacity constraint is binding in both markets, the primary market pays $\lambda_1 = 6$ of the capacity cost and the secondary market pays $\lambda_2 = 2$.

EXERCISE 13.3

1. Suppose in Example 2 a unit of capacity costs only 3 cents per day.
   (*a*) What would be the profit-maximizing peak and off-peak prices and quantities?
   (*b*) What would be the values of the Lagrange multipliers? What interpretation do you put on their values?
2. A consumer lives on an island where she produces two goods, x and y, according to the production possibility frontier $x^2 + y^2 \leq 200$, and she consumes all the goods herself. Her utility function is

   $U = xy^3$

The consumer also faces an environmental constraint on her total output of both goods. The environmental constraint is given by $x + y \le 20$.

(*a*) Write out the Kuhn-Tucker first-order conditions.

(*b*) Find the consumer's optimal x and y. Identify which constraints are binding.

3. An electric company is setting up a power plant in a foreign country, and it has to plan its capacity. The peak-period demand for power is given by $P_1 = 400 - Q_1$ and the off-peak demand is given by $P_2 = 380 - Q_2$. The variable cost is 20 per unit (paid in both markets) and capacity costs 10 per unit which is only paid once and is used in both periods.

(*a*) Write out the Lagrangian and Kuhn-Tucker conditions for this problem.

(*b*) Find the optimal outputs and capacity for this problem.

(*c*) How much of the capacity is paid for by each market (i.e., what are the values of λ_1 and λ_2)?

(*d*) Now suppose capacity cost is 30 cents per unit (paid only once). Find quantities, capacity, and how much of the capacity is paid for by each market (i.e., λ_1 and λ_2).

13.4 Sufficiency Theorems in Nonlinear Programming

In the previous sections, we have introduced the Kuhn-Tucker conditions and illustrated their applications *as necessary* conditions in optimization problems with inequality constraints. Under certain circumstances, the Kuhn-Tucker conditions can also be taken as sufficient conditions.

The Kuhn-Tucker Sufficiency Theorem: Concave Programming

In classical optimization problems, the sufficient conditions for maximum and minimum are traditionally expressed in terms of the signs of second-order derivatives or differentials. As we have shown in Sec. 11.5, however, these second-order conditions are closely related to the concepts of concavity and convexity of the objective function. Here, in nonlinear programming, the sufficient conditions can also be stated directly in terms of concavity and convexity. And, in fact, these concepts will be applied not only to the objective function $f(x)$ but to the constraint functions $g^i(x)$ as well.

For the *maximization* problem, Kuhn and Tucker offer the following statement of sufficient conditions (sufficiency theorem):

Given the nonlinear programming problem

$$\begin{aligned}
&\text{Maximize} && \pi = f(x) \\
&\text{subject to} && g^i(x) \le r_i \qquad (i = 1, 2, \ldots, m) \\
&\text{and} && x \ge 0
\end{aligned}$$

if the following conditions are satisfied:

(*a*) the objective function $f(x)$ is differentiable and *concave* in the nonnegative orthant

(*b*) each constraint function $g^i(x)$ is differentiable and *convex* in the nonnegative orthant

(*c*) the point x^* satisfies the Kuhn-Tucker maximum conditions

then x^* gives a global maximum of $\pi = f(x)$.

Note that, in this theorem, the constraint qualification is nowhere mentioned. This is because we have already assumed, in condition (*c*), that the Kuhn-Tucker conditions are satisfied at x^* and, consequently, the question of the constraint qualification is no longer an issue.

As it stands, the above theorem indicates that conditions (a), (b), and (c) are sufficient to establish x^* to be an optimal solution. Looking at it differently, however, we may also interpret it to mean that given (a) and (b), then the Kuhn-Tucker maximum conditions are sufficient for a maximum. In the preceding section, we learned that the Kuhn-Tucker conditions, though not necessary *per se,* become necessary when the constraint qualification is satisfied. Combining this information with the sufficiency theorem, we may now state that if the constraint qualification is satisfied and if conditions (a) and (b) are realized, then the Kuhn-Tucker maximum conditions will be *necessary-and-sufficient* for a maximum. This would be the case, for instance, when all the constraints are linear inequalities, which is sufficient for satisfying the constraint qualification.

The maximization problem dealt with in the sufficiency theorem above is often referred to as *concave programming*. This name arises because Kuhn and Tucker adopt the $\geq$ inequality instead of the $\leq$ inequality in every constraint, so that condition (b) would require the $g^i(x)$ functions to be *all concave,* like the $f(x)$ function. But we have modified the formulation in order to convey the idea that in a maximization problem, a constraint is imposed to "rein in" (hence, $\leq$) the attempt to ascend to higher points on the objective function. Though different in form, the two formulations are equivalent in substance. For brevity, we omit the proof.

As stated above, the sufficiency theorem deals only with maximization problems. But adaptation to *minimization* problems is by no means difficult. Aside from the appropriate changes in the theorem to reflect the reversal of the problem itself, all we have to do is to interchange the two words *concave* and *convex* in conditions (a) and (b) and to use the Kuhn-Tucker *minimum* conditions in condition (c). (See Exercise 13.4-1.)

The Arrow-Enthoven Sufficiency Theorem: Quasiconcave Programming

To apply the Kuhn-Tucker sufficiency theorem, certain concavity-convexity specifications must be met. These constitute quite stringent requirements. In another sufficiency theorem—the Arrow-Enthoven sufficiency theorem†—these specifications are relaxed to the extent of requiring only *quasiconcavity* and *quasiconvexity* in the objective and constraint functions. With the requirements thus weakened, the scope of applicability of the sufficient conditions is correspondingly widened.

In the original formulation of the Arrow-Enthoven paper, with a maximization problem and with constraints in the $\geq$ form, the $f(x)$ and $g^i(x)$ functions must uniformly be quasiconcave in order for their theorem to be applicable. This gives rise to the name *quasiconcave programming*. In our discussion here, however, we shall again use the $\leq$ inequality in the constraints of a maximization problem and the $\geq$ inequality in the minimization problem.

The theorem is as follows:

Given the nonlinear programming problem

$$\begin{aligned}
&\text{Maximize} && \pi = f(x) \\
&\text{subject to} && g^i(x) \leq r_i \qquad (i = 1, 2, \ldots, m) \\
&\text{and} && x \geq 0
\end{aligned}$$

† Kenneth J. Arrow and Alain C. Enthoven, "Quasi-concave Programming," *Econometrica,* October, 1961, pp. 779–800.

if the following conditions are satisfied:

(a) the objective function $f(x)$ is differentiable and *quasiconcave* in the nonnegative orthant

(b) each constraint function $g^i(x)$ is differentiable and *quasiconvex* in the nonnegative orthant

(c) the point x^* satisfies the Kuhn-Tucker maximum conditions

(d) any *one* of the following is satisfied:

- (d-i) $f_j(x^*) < 0$ for at least one variable x_j
- (d-ii) $f_j(x^*) > 0$ for some variable x_j that can take on a positive value without violating the constraints
- (d-iii) the n derivatives $f_j(x^*)$ are not all zero, and the function $f(x)$ is twice differentiable in the neighborhood of x^* [i.e., all the second-order partial derivatives of $f(x)$ exist at x^*]
- (d-iv) the function $f(x)$ is concave

then x^* gives a global maximum of $\pi = f(x)$.

Since the proof of this theorem is quite lengthy, we shall omit it here. However, we do want to call your attention to a few important features of this theorem. For one thing, while Arrow and Enthoven have succeeded in weakening the concavity-convexity specifications to their quasiconcavity-quasiconvexity counterparts, they find it necessary to append a new requirement, (d). Note, though, that only *one* of the four alternatives listed under (d) is required to form a complete set of sufficient conditions. In effect, therefore, the above theorem contains as many as *four* different sets of sufficient conditions for a maximum. In the case of (d-iv), with $f(x)$ concave, it would apear that the Arrow-Enthoven sufficiency theorem becomes identical with the Kuhn-Tucker sufficiency theorem. But this is not true. Inasmuch as Arrow and Enthoven only require the constraint functions $g^i(x)$ to be *quasiconvex,* their sufficient conditions are still weaker.

As stated, the theorem lumps together the conditions (a) through (d) as a set of sufficient conditions. But it is also possible to interpret it to mean that, when (a), (b), and (d) are satisfied, then the Kuhn-Tucker maximum conditions become sufficient conditions for a maximum. Furthermore, if the constraint qualification is also satisfied, then the Kuhn-Tucker conditions will become necessary-and-sufficient for a maximum.

Like the Kuhn-Tucker theorem, the Arrow-Enthoven theorem can be adapted with ease to the *minimization* framework. Aside from the obvious changes that are needed to reverse the direction of optimization, we simply have to interchange the words *quasiconcave* and *quasiconvex* in conditions (a) and (b), replace the Kuhn-Tucker maximum conditions by the minimum conditions, reverse the inequalities in (d-i) and (d-ii), and change the word *concave* to *convex* in (d-iv).

A Constraint-Qualification Test

It was mentioned in Sec. 13.2 that if all constraint functions are linear, then the constraint qualification is satisfied. In case the $g^i(x)$ functions are nonlinear, the following test offered by Arrow and Enthoven may prove useful in determining whether the constraint qualification is satisfied:

For a maximization problem, if

(a) every constraint function $g^i(x)$ is differentiable and quasiconvex

(b) there exists a point x^0 in the nonnegative orthant such that all the constraints are satisfied as strict inequalities at x^0

(c) *one* of the following is true:

(c-i) every $g^i(x)$ function is convex

(c-ii) the partial derivatives of every $g^i(x)$ are not all zero when evaluated at every point x in the feasible region

then the constraint qualification is satisfied.

Again, this test can be adapted to the minimization problem with ease. To do so, just change the word *quasiconvex* to *quasiconcave* in condition (a), and change the word *convex* to *concave* in (c-i).

EXERCISE 13.4

1. Given: Minimize $C = F(x)$
   subject to $G^i(x) \geq r_i \qquad (i = 1, 2, \ldots, m)$
   and $x > 0$

   (a) Convert it into a maximization problem.

   (b) What in the present problem are the equivalents of the f and g^i functions in the Kuhn-Tucker sufficiency theorem?

   (c) Hence, what concavity-convexity conditions should be placed on the F and G^i functions to make the sufficient conditions for a maximum applicable here?

   (d) On the basis of the above, how would you state the Kuhn-Tucker sufficient conditions for a *minimum?*

2. Is the Kuhn-Tucker sufficiency theorem applicable to:

   (a) Maximize $\pi = x_1$
   subject to $x_1^2 + x_3^2 \leq 1$
   and $x_1, x_2 \geq 0$

   (b) Minimize $C = (x_1 - 3)^2 + (x_2 - 4)^2$
   subject to $x_1 + x_2 \geq 4$
   and $x_1, x_2 \geq 0$

   (c) Minimize $C = 2x_1 + x_2$
   subject to $x_1^2 - 4x_1 + x_2 \geq 0$
   and $x_1, x_2 \geq 0$

3. Which of the following functions are mathematically acceptable as the objective function of a *maximization* problem which qualifies for the application of the Arrow-Enthoven sufficiency theorem?

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

   (b) $f(x_1, x_2) = 6x_1 - 9x_2$

   (c) $f(x_1, x_2) = x_2 - \ln x_1$ (*Note:* See Exercise 12.4-4.)

4. Is the Arrow-Enthoven constraint qualification satisfied, given that the constraints of a *maximization* problem are:
   (*a*) $x_1^2 + (x_2 - 5)^2 \le 4$ and $5x_1 + x_2 < 10$
   (*b*) $x_1 + x_2 \le 8$ and $-x_1 x_2 \le -8$ (*Note:* $-x_1 x_2$ is not convex.)

13.5 Maximum-Value Functions and the Envelope Theorem†

A maximum-value function is an objective function where the choice variables have been assigned their optimal values. These optimal values of the choice variables are, in turn, functions of the exogenous variables and parameters of the problem. Once the optimal values of the choice variables have been substituted into the original objective function, the function indirectly becomes a function of the parameters only (through the parameters' influence on the optimal values of the choice variables). Thus the maximum-value function is also referred to as the *indirect objective function*.

The Envelope Theorem for Unconstrained Optimization

What is the significance of the indirect objective function? Consider that in any optimization problem the direct objective function is maximized (or minimized) for a given set of parameters. The indirect objective function traces out all the maximum values of the objective function as these parameters vary. Hence the indirect objective function is an "envelope" of the set of optimized objective functions generated by varying the parameters of the model. For most students of economics the first illustration of this notion of an envelope arises in the comparison of short-run and long-run cost curves. Students are typically taught that the long-run average cost curve is an envelope of all the short-run average cost curves (what parameter is varying along the envelope in this case?). A formal derivation of this concept is one of the exercises we will be doing in this section.

To illustrate, consider the following unconstrained maximization problem with two choice variables x and y and one parameter ϕ:

$$\text{Maximize} \qquad U = f(x, y, \phi) \tag{13.27}$$

The first-order necessary condition is

$$f_x(x, y, \phi) = f_y(x, y, \phi) = 0 \tag{13.28}$$

If second-order conditions are met, these two equations implicitly define the solutions

$$x^* = x^*(\phi) \qquad y^* = y^*(\phi) \tag{13.29}$$

If we substitute these solutions into the objective function, we obtain a new function

$$V(\phi) = f(x^*(\phi), \quad y^*(\phi), \phi) \tag{13.30}$$

where this function is the value of f when the values of x and y are those that maximize $f(x, y, \phi)$. Therefore, $V(\phi)$ is *the maximum-value function* (or indirect objective function).

† This section of the chapter presents an overview of the envelope theorem. A richer treatment of this topic can be found in Chap. 7 of *The Structure of Economics: A Mathematical Analysis* (3rd ed.) by Eugene Silberberg and Wing Suen (McGraw-Hill, 2001) on which parts of this section are based.

If we differentiate V with respect to ϕ, its only argument, we get

$$\frac{dV}{d\phi} = f_x \frac{\partial x^*}{\partial \phi} + f_y \frac{\partial y^*}{\partial \phi} + f_\phi \tag{13.31}$$

However, from the first-order conditions we know $f_x = f_y = 0$. Therefore, the first two terms disappear and the result becomes

$$\frac{dV}{d\phi} = f_\phi \tag{13.31'}$$

This result says that, at the optimum, as ϕ varies, with x^* and y^* allowed to adjust, the derivative $dV/d\phi$ gives the same result as if x^* and y^* are treated as constants. Note that ϕ enters the maximum-value function (13.30) in three places: one direct and two indirect (through x^* and y^*). Equation (13.31′) shows that, at the optimum, only the direct effect of ϕ on the objective function matters. This is the essence of the envelope theorem. The envelope theorem says that only the direct effects of a change in an exogenous variable need be considered, even though the exogenous variable may also enter the maximum-value function indirectly as part of the solution to the endogenous choice variables.

The Profit Function

Let us now apply the notion of the maximum-value function to derive the profit function of a competitive firm. Consider the case where a firm uses two inputs: capital K and labor L. The profit function is

$$\pi = Pf(K, L) - wL - rK \tag{13.32}$$

where P is the output price and w and r are the wage rate and rental rate, respectively.

The first-order conditions are

$$\begin{aligned} \pi_L &= Pf_L(K, L) - w = 0 \\ \pi_K &= Pf_K(K, L) - r = 0 \end{aligned} \tag{13.33}$$

which respectively define the input-demand equations

$$\begin{aligned} L^* &= L^*(w, r, P) \\ K^* &= K^*(w, r, P) \end{aligned} \tag{13.34}$$

Substituting the solutions K^* and L^* into the objective function gives us

$$\pi^*(w, r, P) = Pf(K^*, L^*) - wL^* - rK^* \tag{13.35}$$

where $\pi^*(w, r, P)$ is the *profit function* (an indirect objective function). The profit function gives the maximum profit as a function of the exogenous variables w, r, and P.

Now consider the effect of a change in w on the firm's profits. If we differentiate the original profit function (13.32) with respect to w, holding all other variables constant, we get

$$\frac{\partial \pi}{\partial w} = -L \tag{13.36}$$

However, this result does not take into account the profit-maximizing firm's ability to make a substitution of capital for labor and adjust the level of output in accordance with profit-maximizing behavior.

In contrast, since $\pi^*(w, r, P)$ is the maximum value of profits for any values of w, r, and P, changes in π^* from a change in w takes all capital-for-labor substitutions into account. To evaluate a change in the maximum profit function caused by a change in w, we differentiate $\pi^*(w, r, P)$ with respect to w to obtain

$$\frac{\partial \pi^*}{\partial w} = (Pf_L - w)\frac{\partial L^*}{\partial w} + (Pf_K - r)\frac{\partial K^*}{\partial w} - L^* \tag{13.37}$$

From the first-order conditions (13.33), the two terms in parentheses are equal to zero. Therefore, the equation becomes

$$\frac{\partial \pi^*}{\partial w} = -L^*(w, r, P) \tag{13.38}$$

This result says that, at the profit-maximizing position, a change in profits with respect to a change in the wage rate is the same whether or not the factors are held constant or allowed to vary as the factor price changes. In this case, (13.38) shows that the derivative of the profit function with respect to w is the negative of the factor demand function $L^*(w, r, P)$. Following the preceding procedure, we can also show the additional comparative-static results:

$$\frac{\partial \pi^*(w, r, P)}{\partial r} = -K^*(w, r, P) \tag{13.39}$$

and

$$\frac{\partial \pi^*(w, r, P)}{\partial P} = f(K^*, L^*) \tag{13.40}$$

Equations (13.38), (13.39), and (13.40) are collectively known as *Hotelling's lemma*. We have obtained these comparative-static derivatives from the profit function by allowing K^* and L^* to adjust to any parameter change. But it is easy to see that the same results will emerge if we differentiate the profit function (13.35) with respect to each parameter while holding K^* and L^* constant. Thus Hotelling's lemma is simply another manifestation of the envelope theorem that we encountered earlier in (13.31′).

Reciprocity Condition

Consider again our two-variable unconstrained maximization problem

$$\text{Maximize} \qquad U = f(x, y, \phi) \qquad \text{[from (13.27)]}$$

where x and y are the choice variables and ϕ is a parameter. The first-order conditions are $f_x = f_y = 0$, which imply $x^* = x^*(\phi)$ and $y^* = y^*(\phi)$.

We are interested in the comparative statics regarding the directions of change in $x^*(\phi)$ and $y^*(\phi)$ as ϕ changes and the effects on the value function. The maximum-value function is

$$V(\phi) = f(x^*(\phi), y^*(\phi), \phi) \tag{13.41}$$

By definition, $V(\phi)$ gives the maximum value of f for any given ϕ.

Now consider a new function that depicts the difference between the actual value and the maximum value of U:

$$\Omega(x, y, \phi) = f(x, y, \phi) - V(\phi) \tag{13.42}$$

This new function Ω has a maximum value of zero when $x = x^*$ and $y = y^*$; for any $x \neq x^*, y \neq y^*$ we have $f \leq V$. In this framework $\Omega(x, y, \phi)$ can be considered a function of three independent variables, x, y, and ϕ. The maximum of $\Omega(x, y, \phi) = f(x, y, \phi) - V(\phi)$ can be determined by the first- and second-order conditions.

The first-order conditions are

$$\begin{aligned} \Omega_x(x, y, \phi) &= f_x = 0 \\ \Omega_y(x, y, \phi) &= f_y = 0 \end{aligned} \tag{13.43}$$

and

$$\Omega_\phi(x, y, \phi) = f_\phi - V_\phi = 0 \tag{13.44}$$

We can see that the first-order conditions of our new function Ω in (13.43) are nothing but the original maximum conditions for $f(x, y, \phi)$ in (13.28), whereas the condition in (13.44) really restates the envelope theorem (13.31′). These first-order conditions hold whenever $x = x^*(\phi)$ and $y = y^*(\phi)$. The second-order sufficient conditions are satisfied if the Hessian of Ω

$$H = \begin{vmatrix} f_{xx} & f_{xy} & f_{x\phi} \\ f_{yx} & f_{yy} & f_{y\phi} \\ f_{\phi x} & f_{\phi y} & f_{\phi\phi} - V_{\phi\phi} \end{vmatrix}$$

is characterized by

$$f_{xx} < 0 \qquad f_{xx} f_{yy} - f_{xy}^2 > 0 \qquad H < 0$$

In deriving the Hessian above, we listed the variables in the order (x, y, ϕ) and, consequently, the first entry in the second-order conditions, $(\Omega_{xx} =)\ f_{xx} < 0$ relates to the variable x. Had we adopted an alternative listing order, then the first entry could have been $\Omega_{yy} = f_{yy} < 0$, or

$$\Omega_{\phi\phi} = f_{\phi\phi} - V_{\phi\phi} < 0 \tag{13.45}$$

It turns out that (13.45) can lead us to a result that provides a quick way to reach a comparative-static conclusion. First, we know from (13.41) that

$$V_\phi(\phi) = f_\phi(x^*(\phi), y^*(\phi), \phi)$$

Differentiating both sides with respect to ϕ yields

$$V_{\phi\phi} = f_{\phi x} \frac{\partial x^*}{\partial \phi} + f_{\phi y} \frac{\partial y^*}{\partial \phi} + f_{\phi\phi} \tag{13.46}$$

Using (13.45) and Young's theorem, we can write

$$V_{\phi\phi} - f_{\phi\phi} = f_{x\phi} \frac{\partial x^*}{\partial \phi} + f_{y\phi} \frac{\partial y^*}{\partial \phi} > 0 \tag{13.47}$$

Suppose that ϕ enters only in the first-order condition for x, such that $f_{y\phi} = 0$. Then (13.47) reduces to

$$f_{x\phi} \frac{\partial x^*}{\partial \phi} > 0 \tag{13.48}$$

which implies that $f_{x\phi}$ and $\partial x^*/\partial \phi$ will have the same sign. Thus, whenever we see the parameter ϕ appearing only in the first-order condition relating to x, and once we have determined the sign of the derivative $f_{x\phi}$ from the objective function $U = f(x, y, \phi)$, we can immediately tell the sign of the comparative-static derivative $\partial x^*/\partial \phi$ without further ado.

For example, in the profit-maximization model:

$$\pi = Pf(K, L) - wL - rK$$

where the first-order conditions are

$$\pi_L = Pf_L - w = 0$$
$$\pi_K = Pf_K - r = 0$$

the exogenous variable w enters only the first-order condition $Pf_L - w = 0$, with

$$\frac{\partial \pi_L}{\partial w} = -1$$

Therefore, by (13.48), we can conclude that $\partial L^*/\partial w$ will also be negative.

Further, if we combine the envelope theorem with Young's theorem, we can derive a relation known as the *reciprocity condition:* $\partial L^*/\partial r = \partial K^*/\partial w$. From the indirect profit function $\pi^*(w, r, P)$, Hotelling's lemma gives us

$$\pi_w^* = \frac{\partial \pi^*}{\partial w} = -L^*(w, r, P)$$
$$\pi_r^* = \frac{\partial \pi^*}{\partial r} = -K^*(w, r, P)$$

Differentiating again and applying Young' theorem, we have

$$\pi_{wr}^* = -\frac{\partial L^*}{\partial r} = -\frac{\partial K^*}{\partial w} = \pi_{rw}^*$$

or

$$\frac{\partial L^*}{\partial r} = \frac{\partial K^*}{\partial w} \qquad \textbf{(13.49)}$$

This result is referred to as the *reciprocity condition* because it shows the symmetry between the comparative-static cross effect produced by the price of one input on the demand for the "other" input. Specifically, in the comparative-static sense, the effect of r (the rental rate for capital K) on the optimal demand for labor L is the same as the effect of w (the wage rate for labor L) on the optimal demand for capital K.

The Envelope Theorem for Constrained Optimization

The envelope theorem can also be derived for the case of constrained optimization. Again we will have an objective function (U), two choice variables (x and y) and one parameter (ϕ); except now we introduce the following constraint:

$$g(x, y; \phi) = 0$$

The problem becomes:

$$\begin{aligned} &\text{Maximize} \quad U = f(x, y; \phi) \\ &\text{subject to} \quad g(x, y; \phi) = 0 \end{aligned} \qquad \textbf{(13.50)}$$

The Lagrangian for this problem is

$$Z = f(x, y; \phi) + \lambda[0 - g(x, y; \phi)] \qquad \textbf{(13.51)}$$

with first-order conditions

$$Z_x = f_x - \lambda g_x = 0$$
$$Z_y = f_y - \lambda g_y = 0$$
$$Z_\lambda = -g(x, y; \phi) = 0$$

Solving this system of equations gives us

$$x = x^*(\phi) \qquad y = y^*(\phi) \qquad \lambda = \lambda^*(\phi)$$

Substituting the solutions into the objective function, we get

$$U^* = f(x^*(\phi), y^*(\phi), \phi) = V(\phi) \tag{13.52}$$

where $V(\phi)$ is the indirect objective function, a maximum-value function. This is the maximum value of y for any ϕ and x_i's that satisfy the constraint.

How does $V(\phi)$ change as ϕ changes? First, we differentiate V with respect to ϕ:

$$\frac{dV}{d\phi} = f_x \frac{\partial x^*}{\partial \phi} + f_y \frac{\partial y^*}{\partial \phi} + f_\phi \tag{13.53}$$

In this case, however, (13.53) will not simplify to $dV/d\phi = f_\phi$ since in constrained optimization, it is not necessary to have $f_x = f_y = 0$ (see Table 12.1). But if we substitute the solutions to x and y into the constraint (producing an identity), we get

$$g(x^*(\phi), y^*(\phi), \phi) \equiv 0$$

and differentiating this with respect to ϕ yields

$$g_x \frac{\partial x^*}{\partial \phi} + g_y \frac{\partial y^*}{\partial \phi} + g_\phi \equiv 0 \tag{13.54}$$

If we multiply (13.54) by λ, combine the result with (13.53), and rearrange terms, we get

$$\frac{dV}{d\phi} = (f_x - \lambda g_x)\frac{\partial x^*}{\partial \phi} + (f_y - \lambda g_y)\frac{\partial y^*}{\partial \phi} + f_\phi - \lambda g_\phi = Z_\phi \tag{13.55}$$

where Z_ϕ is the partial derivative of the Lagrangian function with respect to ϕ, holding all other variables constant. This result is in the same spirit as (13.31), and by virtue of the first-order conditions, it reduces to

$$\frac{dV}{d\phi} = Z_\phi \tag{13.55'}$$

which represents the envelope theorem in the framework of constrained optimization. Note, however, in the present case, the Lagrangian function replaces the objective function in deriving the indirect objective function.

While the results in (13.55) nicely parallel the unconstrained case, it is important to note that some of the comparative-static results depend critically on whether the parameters enter only the objective function, or only the constraints, or enter both. If a parameter enters only in the objective function, then the comparative-static results are the same as for the unconstrained case. However, if the parameter enters the constraint, the relation

$$V_{\phi\phi} \geq f_{\phi\phi}$$

will no longer hold.

Interpretation of the Lagrange Multiplier

In the consumer choice problem in Chap. 12 we derived the result that the Lagrange multiplier λ represented the change in the value of the Lagrange function when the consumer's budget changed. We interpreted λ as the marginal utility of income. Now let us derive a more general interpretation of the Lagrange multiplier with the assistance of the envelope theorem. Consider the problem

$$\begin{aligned} &\text{Maximize} && U = f(x, y) \\ &\text{subject to} && g(x, y) = c \end{aligned}$$

where c is a constant. The Lagrangian for this problem is

$$Z = f(x, y) + \lambda[c - g(x, y)] \tag{13.56}$$

The first-order conditions are

$$\begin{aligned} Z_x &= f_x(x, y) - \lambda g_x(x, y) = 0 \\ Z_y &= f_y(x, y) - \lambda g_y(x, y) = 0 \\ Z_\lambda &= c - g(x, y) = 0 \end{aligned} \tag{13.57}$$

From the first two equations in (13.57), we get

$$\lambda = \frac{f_x}{g_x} = \frac{f_y}{g_y} \tag{13.58}$$

which gives us the condition that the slope of the level curve (indifference curve) of the objective function must equal the slope of the constraint at the optimum.

Equations (13.57) implicitly define the solutions

$$x^* = x^*(c) \qquad y^* = y^*(c) \qquad \lambda^* = \lambda^*(c) \tag{13.59}$$

Substituting (13.59) back into the Lagrangian yields the maximum-value function,

$$V(c) = Z^*(c) = f(x^*(c), y^*(c)) + \lambda^*(c)[c - g(x_1^*(c), y^*(c))] \tag{13.60}$$

Differentiating with respect to c yields

$$\begin{aligned} \frac{dV}{dc} = \frac{dZ^*}{dc} = f_x\frac{\partial x^*}{\partial c} + f_y\frac{\partial y^*}{\partial c} + [c - g(x^*(c), y^*(c))]\frac{\partial \lambda^*}{\partial c} \\ - \lambda^*(c)g_x\frac{\partial x^*}{\partial c} - \lambda^*(c)g_y\frac{\partial y^*}{\partial c} + \lambda^*(c)\frac{dc}{dc} \end{aligned}$$

By rearranging we get

$$\frac{dZ^*}{dc} = [f_x - \lambda^* g_x]\frac{\partial x^*}{\partial c} + [f_y - \lambda^* g_y]\frac{\partial y^*}{\partial c} + [c - g(x^*, y^*)]\frac{\partial \lambda^*}{\partial c} + \lambda^*$$

By (13.57), the three terms in brackets are all equal to zero. Therefore this expression simplifies to

$$\frac{dV}{dc} = \frac{dZ^*}{dc} = \lambda^* \tag{13.61}$$

which shows that the optimal value λ^* measures the rate of change of the maximum value of the objective function when c changes, and is for this reason referred to as the "shadow price" of c. Note that, in this case, c enters the problem only through the constraint; it is not an argument of the original objective function.

13.6 Duality and the Envelope Theorem

A consumer's expenditure function and his or her indirect utility function exemplify the minimum- and maximum-value functions for *dual problems*.† An expenditure function specifies the minimum expenditure required to obtain a fixed level of utility given the utility function and the prices of consumption goods. An indirect utility function specifies the maximum utility that can be obtained given prices, income, and the utility function.

The Primal Problem

Let $U(x, y)$ be a utility function where x and y are consumption goods. The consumer has a budget B and faces market prices P_x and P_y for goods x and y, respectively. This problem will be considered the *primal problem:*

$$\begin{aligned} &\text{Maximize} && U = U(x, y) \\ &\text{subject to} && P_x x + P_y y = B \end{aligned} \qquad \text{[Primal]} \qquad \textbf{(13.62)}$$

For this problem, we have the familiar Lagrangian

$$Z = U(x, y) + \lambda(B - P_x x - P_y y)$$

The first-order conditions are

$$\begin{aligned} Z_x &= U_x - \lambda P_x = 0 \\ Z_y &= U_y - \lambda P_y = 0 \\ Z_\lambda &= B - P_x x - P_y y = 0 \end{aligned} \qquad \textbf{(13.63)}$$

This system of equations implicitly defines a solution for x^m, y^m, and λ^m as a function of the exogenous variables B, P_x, P_y:

$$\begin{aligned} x^m &= x^m(P_x, P_y, B) \\ y^m &= y^m(P_x, P_y, B) \\ \lambda^m &= \lambda^m(P_x, P_y, B) \end{aligned}$$

The solutions x^m and y^m are the consumer's ordinary demand functions, sometimes called the "Marshallian" demand functions, hence the superscript m.

Substituting the solutions x^m and y^m into the utility function yields

$$U^* = U^*(x^m(P_x, P_y, B), y^m(P_x, P_y, B)) \equiv V(P_x, P_y, B) \qquad \textbf{(13.64)}$$

where V is the indirect utility function—a maximum-value function showing the maximum attainable utility in problem (13.62). We shall return to this function later.

† Duality in economic theory is the relationship between two constrained optimization problems. If one of the problems requires constrained maximization, the other problem will require constrained minimization. The structure and solution of either problem can provide information about the structure and solution of the other problem.

The Dual Problem

Now consider a related *dual problem* for the consumer with the objective of minimizing the expenditure on x and y while maintaining a fixed utility level U^* derived from (13.64) of the primal problem:

$$\begin{aligned} &\text{Minimize} \quad && E = P_x x + P_y y \\ &\text{subject to} \quad && U(x, y) = U^* \end{aligned} \qquad \text{[Dual]} \tag{13.65}$$

Its Lagrangian is

$$Z^d = P_x x + P_y y + \mu\,[U^* - U(x, y)]$$

and the first-order conditions are

$$\begin{aligned} Z^d_x &= P_x - \mu U_x = 0 \\ Z^d_y &= P_y - \mu U_y = 0 \\ Z^d_\lambda &= U^* - U(x, y) = 0 \end{aligned} \tag{13.66}$$

This system of equations implicitly defines a set of solution values to be labeled x^h, y^h, and λ^h:

$$\begin{aligned} x^h &= x^h(P_x, P_y, U^*) \\ y^h &= y^h(P_x, P_y, U^*) \\ \mu^h &= \mu^h(P_x, P_y, U^*) \end{aligned}$$

Here x^h and y^h are the compensated ("real income" held constant) demand functions. They are commonly referred to as "Hicksian" demand functions, hence the h superscript.

Substituting x^h and y^h into the objective function of the dual problem yields

$$P_x x^h(P_x, P_y, U^*) + P_y y^h(P_x, P_y, U^*) \equiv E(P_x, P_y, U^*) \tag{13.67}$$

where E is the expenditure function—a minimum-value function showing the minimum expenditure needed to attain the utility level U^*.

Duality

If we take the first two equations in (13.63) and in (13.64), and eliminate the Lagrange multipliers, we can write

$$\frac{Px}{Py} = \frac{Ux}{Uy} \tag{13.68}$$

This is the tangency condition in which the consumer chooses the optimal bundle where the slope of the indifference curve equals the slope of the budget constraint. The tangency condition is identical for both problems. Thus, when the target level of utility in the minimization problem is set equal to the value U^* obtained from the maximization problem, we get

$$\begin{aligned} x^m(P_x, P_y, B) &= x^h(P_x, P_y, U^*) \\ y^m(P_x, P_y, B) &= y^h(P_x, P_y, U^*) \end{aligned} \tag{13.69}$$

i.e., the solutions to both the maximization problem and the minimization problem produce identical values for x and y. However, the solutions are functions of different exogenous variables, so comparative-static exercises will generally produce different results.

The fact that the solution values for x and y in the primal and dual problems are determined by the tangency point of the same indifference curve and budget-constraint line means that the minimized expenditure in the dual problem is equal to the given budget B of the primal problem:

$$E(P_x, P_y, U^*) = B \tag{13.70}$$

This result is parallel to the result in (13.64), which reveals that the maximized value of utility V in the primal problem is equal to the given target level of utility U^* in the dual problem.

While the solution values of x and y are identical in the two problems, the same cannot be said about the Lagrange multipliers. From the first equation in (13.63) and in (13.66), we can calculate $\lambda = U_x/P_x$, but $\mu = P_x/U_x$. Thus, the solution values of λ and μ are reciprocal to each other:

$$\lambda = \frac{1}{\mu} \qquad \text{or} \qquad \lambda^m = \frac{1}{\mu^h} \tag{13.71}$$

Roy's Identity

One application of the envelope theorem is the derivation of Roy's identity. Roy's identity states that the individual consumer's Marshallian demand function is equal to negative of the ratio of two partial derivatives of the maximum-value function.

Substituting the optimal values x^m, y^m, and λ^m into the Lagrangian of (13.62) gives us

$$V(P_x, P_y, B) = U(x^m, y^m) + \lambda^m\left(B - P_x x^m - P_y y^m\right) \tag{13.72}$$

When we differentiate (13.72) with respect to P_x we find

$$\begin{aligned}\frac{\partial V}{\partial P_x} &= (U_x - \lambda^m P_x)\frac{\partial x^m}{\partial P_x} + (U_y - \lambda^m P_y)\frac{\partial y^m}{\partial P_x} \\ &\quad + (B - P_x x^m - P_y y^m)\frac{\partial \lambda^m}{\partial P_x} - \lambda^m x^m\end{aligned}$$

At the optimum, the first-order conditions (13.63) enable us to simplify this to

$$\frac{\partial V}{\partial P_x} = -\lambda^m x^m$$

Next, differentiate the value function with respect to B to get

$$\begin{aligned}\frac{\partial V}{\partial B} &= (U_x - \lambda^m P_x)\frac{\partial x^m}{\partial B} + (U_y - \lambda^m P_y)\frac{\partial y^m}{\partial B} \\ &\quad + (B - P_x x^m - P_y y^m)\frac{\partial \lambda^m}{\partial B} + \lambda^m\end{aligned}$$

Again, at the optimum, (13.63) enables us to simplify this to

$$\frac{\partial V}{\partial B} = \lambda^m$$

By taking the ratio of these two partial derivatives, we find that

$$\frac{\partial V/\partial P_x}{\partial V/\partial B} = -x^m \tag{13.73}$$

This result, known as *Roy's identity,* shows that the Marshallian demand for commodity x is the negative of the ratio of two partial derivatives of the maximum-value function V with respect to P_x and B, respectively. In view of the symmetry between x and y in the problem, a result similar to (13.73) can also be written for y^m, the Marshallian demand for y. Of course, this result could be arrived at directly by applying the envelope theorem.

Shephard's Lemma

In Sec. 13.5, we derived Hotelling's lemma, which states that the partial derivatives of the maximum value of the profit function yields the firm's input-demand functions and the supply functions. A similar approach applied to the expenditure function yields Shephard's lemma.

Consider the consumer's minimization problem (13.65). The Lagrangian is

$$Z^d = P_x x + P_y y + \mu[U^* - U(x, y)]$$

From the first-order conditions, the following solutions are implicitly defined

$$\begin{aligned} x^h &= x^h(P_x, P_y, U^*) \\ y^h &= y^h(P_x, P_y, U^*) \\ \mu^h &= \mu^h(P_x, P_y, U^*) \end{aligned}$$

Substituting these solutions into the Lagrangian yields the expenditure function:

$$E(P_x, P_y, U^*) = P_x x^h + P_y y^h + \mu^h[U^* - U(x^h, y^h)]$$

Taking the partial derivatives of this function with respect to P_x and P_y and evaluating them at the optimum, we find that $\partial E/\partial P_x$ and $\partial E/\partial P_y$ represent the consumer's Hicksian demands:

$$\begin{aligned} \frac{\partial E}{\partial P_x} &= (P_x - \mu^h U_x)\frac{\partial x^h}{\partial P_x} + (P_y - \mu^h U_y)\frac{\partial y^h}{\partial P_x} + [U^* - U(x^h, y^h)]\frac{\partial \mu^h}{\partial P_x} + x^h \\ &= (0)\frac{\partial x^h}{\partial P_x} + (0)\frac{\partial y^h}{\partial P_x} + (0)\frac{\partial \mu^h}{\partial P_x} + x^h = x^h \end{aligned} \qquad \textbf{(13.74)}$$

and

$$\begin{aligned} \frac{\partial E}{\partial P_y} &= (P_x - \mu^h U_x)\frac{\partial x^h}{\partial P_y} + (P_y - \mu^h U_y)\frac{\partial y^h}{\partial P_y} + [U^* - U(x^h, y^h)]\frac{\partial \mu^h}{\partial P_y} + y^h \\ &= (0)\frac{\partial x^h}{\partial P_y} + (0)\frac{\partial y^h}{\partial P_y} + (0)\frac{\partial \mu^h}{\partial P_y} + y^h = y^h \end{aligned} \qquad \textbf{(13.74')}$$

Finally, differentiating E with respect to the constraint U^* yields μ^h, the marginal cost of the constraint

$$\begin{aligned} \frac{\partial E}{\partial U^*} &= (P_x - \mu^h U_x)\frac{\partial x^h}{\partial U^*} + (P_y - \mu^h U_y)\frac{\partial y^h}{\partial P_y} \\ &\quad + [U^* - U(x^h, y^h)]\frac{\partial \mu^h}{\partial U^*} + \mu^h \\ &= (0)\frac{\partial x^h}{\partial U^*} + (0)\frac{\partial y^h}{\partial U^*} + (0)\frac{\partial \mu^h}{\partial U^*} + \mu^h = \mu^h \end{aligned} \qquad \textbf{(13.74'')}$$

Together, the three partial derivatives (13.74), (13.74′), and (13.74″) are referred to as Shephard's lemma.

Example 1

Consider a consumer with the utility function $U = xy$, who faces a budget constraint of B and is given prices P_x and P_y.

The choice problem is

$$\text{Maximize} \quad U = xy$$
$$\text{subject to} \quad P_x x + P_y y = B$$

The Lagrangian for this problem is

$$Z = xy + \lambda(B - P_x x - P_y y)$$

The first-order conditions are

$$Z_x = y - \lambda P_x = 0$$
$$Z_y = x - \lambda P_y = 0$$
$$Z_\lambda = B - P_x x - P_y y = 0$$

Solving the first-order conditions yields the following solutions:

$$x^m = \frac{B}{2P_x} \qquad y^m = \frac{B}{2P_y} \qquad \lambda^m = \frac{B}{2P_x P_y}$$

where x^m and y^m are the consumer's Marshallian demand functions. For the second-order condition, since the bordered Hessian is

$$|\bar{H}| = \begin{vmatrix} 0 & 1 & -P_x \\ 1 & 0 & -P_y \\ -P_x & -P_y & 0 \end{vmatrix} = 2P_x P_y > 0$$

the solution does represent a maximum.†

We can now derive the indirect utility function for this problem by substituting x^m and y^m into the utility function:

$$V(P_x, P_y, B) = \left(\frac{B}{2P_x}\right)\left(\frac{B}{2P_y}\right) = \frac{B^2}{4P_x P_y} \qquad \textbf{(13.75)}$$

where V denotes the maximized utility. Since V represents the maximized utility, we can set $V = U^*$ in (13.75) to get $B^2/4P_x P_y = U^*$, and then rearrange terms to express B as

$$B = (4P_x P_y U^*)^{1/2} = 2P_x^{1/2} P_y^{1/2} U^{*1/2}$$

Now, think of the consumer's dual problem of expenditure minimization. In the dual problem, the minimum-expenditure function E should be equal to the given budget amount B of the primal problem. Therefore, we can immediately conclude from the preceding equation that

$$E(P_x, P_y, U^*) = B = 2P_x^{1/2} P_y^{1/2} U^{*1/2} \qquad \textbf{(13.76)}$$

† Note that the bordered Hessian is written here (and in Example 2 on page 440) with the borders in the third row and column, instead of in the first row and column as in (12.19). This is the result of listing the Lagrange multiplier as the last rather than the first variable as we did in previous chapters. Exercise 12.3-3 shows that the two alternative expressions for the bordered Hessian are transformable into each other by elementary row operations without affecting its value. However, when more than two choice variables appear in a problem, it is preferable to use the (12.19) format because that makes it easier to write out the bordered leading principal minors.

Let's now use this example to verify Roy's identity (13.73)

$$x^m = -\frac{\partial V/\partial P_x}{\partial V/\partial B}$$

Taking the relevant partial derivatives of V, we find

$$\frac{\partial V}{\partial P_x} = -\frac{B^2}{4P_x^2 P_y}$$

and

$$\frac{\partial V}{\partial B} = \frac{B}{2P_x P_y}$$

The negative of the ratio of these two partials is

$$-\frac{\dfrac{\partial V}{\partial P_x}}{\dfrac{\partial V}{\partial B}} = -\frac{\left(\dfrac{B^2}{4P_x^2 P_y}\right)}{\left(\dfrac{B}{2P_x P_y}\right)} = \frac{B}{2P_x} = x^m$$

Thus we find that Roy's identity does hold.

Example 2

Now consider the dual problem of cost minimization given a fixed level of utility related to Example 1. Letting U^* denote the target level of utility, the problem is:

$$\begin{aligned} &\text{Minimize} && P_x x + P_y y \\ &\text{subject to} && xy = U^* \end{aligned}$$

The Lagrangian for the problem is

$$Z^d = P_x x + P_y y + \mu(U^* - xy)$$

The first-order conditions are

$$\begin{aligned} Z_x^d &= P_x - \mu y = 0 \\ Z_y^d &= P_y - \mu x = 0 \\ Z_\mu^d &= U^* - xy = 0 \end{aligned}$$

Solving the system of equations for x, y, and μ, we get

$$\begin{aligned} x^h &= \left(\frac{P_y U^*}{P_x}\right)^{\frac{1}{2}} \\ y^h &= \left(\frac{P_x U^*}{P_y}\right)^{\frac{1}{2}} \qquad \textbf{(13.77)} \\ \mu^h &= \left(\frac{P_x P_y}{U^*}\right)^{\frac{1}{2}} \end{aligned}$$

where x^h and y^h are the consumer's compensated (Hicksian) demand functions. Checking the second-order condition for a minimum, we find

$$|\overline{H}| = \begin{vmatrix} 0 & -\mu & -y \\ -\mu & 0 & -x \\ -y & -x & 0 \end{vmatrix} = -2xy\mu < 0$$

Thus the sufficient condition for a minimum is satisfied.

Substituting x^h and y^h into the original objective function gives us the minimum-value function, or expenditure function

$$E = P_x x^h + P_y y^h = P_x \left(\frac{P_y U^*}{P_x}\right)^{1/2} + P_y \left(\frac{P_x U^*}{P_y}\right)^{1/2}$$
$$= (P_x P_y U^*)^{1/2} + (P_x P_y U^*)^{1/2}$$
$$= 2P_x^{1/2} P_y^{1/2} U^{*1/2} \qquad \textbf{(13.76')}$$

Note that this result is identical with (13.76) in Example 1. The only difference lies in the process used to derive the result. Equation (13.76′) is obtained directly from an expenditure-minimization problem, whereas (13.76) is indirectly deduced, via the duality relationship, from a utility-maximization problem.

We shall now use this example to test the validity of Shephard's lemma (13.74), (13.74′), and (13.74″). Differentiating the expenditure function in (13.76′) with respect to P_x, P_y, and U^*, respectively, and relating the resulting partial derivatives to (13.77), we find

$$\frac{\partial E(P_x, P_y, U^*)}{\partial P_x} = \frac{P_y^{1/2} U^{*1/2}}{P_x^{1/2}} = x^h$$
$$\frac{\partial E(P_x, P_y, U^*)}{\partial P_y} = \frac{P_x^{1/2} U^{*1/2}}{P_y^{1/2}} = y^h$$
$$\frac{\partial E(P_x, P_y, U^*)}{\partial U^*} = \frac{P_x^{1/2} P_y^{1/2}}{U^{*1/2}} = \mu^h$$

Thus, Shephard's Lemma holds in this example.

EXERCISE 13.6

1. A consumer has the following utility function: $U(x, y) = x(y + 1)$, where x and y are quantities of two consumption goods whose prices are P_x and P_y, respectively. The consumer also has a budget of B. Therefore, the consumer's Lagrangian is

   $x(y + 1) + \lambda(B - P_x x - P_y y)$

   (*a*) From the first-order conditions find expressions for the demand functions. What kind of good is y? In particular what happens when $P_y > B$?

   (*b*) Verify that this is a maximum by checking the second-order conditions. By substituting x^* and y^* into the utility function, find an expression for the indirect utility function

   $U^* = U(P_x, P_y, B)$

   and derive an expression for the expenditure function

   $E = E(P_x, P_y, U^*)$

   (*c*) This problem could be recast as the following dual problem

   Minimize $\quad P_x x + P_y y$

   subject to $\quad x(y + 1) = U^*$

   Find the values of x and y that solve this minimization problem and show that the values of x and y are equal to the partial derivatives of the expenditure function, $\partial E / \partial P_x$ and $\partial E / \partial P_y$, respectively.

13.7 Some Concluding Remarks

In the present part of the book, we have covered the basic techniques of optimization. The somewhat arduous journey has taken us (1) from the case of a single choice variable to the more general *n*-variable case, (2) from the polynomial objective function to the exponential and logarithmic, and (3) from the unconstrained to the constrained variety of extremum.

Most of this discussion consists of the "classical" methods of optimization, with differential calculus as the mainstay, and derivatives of various orders as the primary tools. One weakness of the calculus approach to optimization is its essentially myopic nature. While the first- and second-order conditions in terms of derivatives or differentials can normally locate relative or local extrema without difficulty, additional information or further investigation is often required for identification of absolute or global extrema. Our detailed discussion of concavity, convexity, quasiconcavity, and quasiconvexity is intended as a useful stepping-stone from the realm of relative extrema to that of absolute ones.

A more serious limitation of the calculus approach is its inability to cope with constraints in the inequality form. For this reason, the budget constraint in the utility-maximization model, for instance, is stated in the form that the total expenditure be exactly *equal to* (and not "less than or equal to") a specified sum. In other words, the limitation of the calculus approach makes it necessary to deny the consumer the option of saving part of the available funds. And, for the same reason, the classical approach does not allow us to specify explicitly that the choice variables must be nonnegative as is appropriate in most economic analysis.

Fortunately, we are liberated from these limitations when we introduce the modern optimization technique known as nonlinear programming. Here we can openly admit inequality constraints, including nonnegativity restrictions on the choice variables, into the problem. This obviously represents a giant step forward in the development of optimization methodology.

Still, even in nonlinear programming, the analytical framework remains static. The problem and its solution relate only to the optimal state at one point of time and cannot address the question of how an optimizing agent should, under given circumstances, behave over a period of time. The latter question pertains to the realm of *dynamic optimization,* which we are unable to handle until we have learned the basics of dynamic analysis—the analysis of movements of variables over time. In fact, aside from its application to dynamic optimization, dynamic analysis is, in itself, an important branch of economic analysis. For this reason, we shall now turn our attention to the subject of dynamic analysis in Part 5.