12. Optimization with Equality Constraints

Chapter 11 presented a general method for finding the relative extrema of an objective function of two or more choice variables. One important feature of that discussion is that all the choice variables are independent of one another, in the sense that the decision made regarding one variable does not impinge upon the choices of the remaining variables.

If the said firm is somehow required to observe a restriction such as a production quota in the form of $Q_1 + Q_2 = 950$, however, the independence between the choice variables will be lost. The new optimum satisfying the production quota constitutes a *constrained optimum*, which, in general, may be expected to differ the *free optimum* discussed in Chapter 11.

12.1 Effects of a Constraint

Let us consider a consumer with the simple utility function

$$U = x_1 x_2 + 2x_1 \tag{1}$$

Since the marginal utilities are positive for all positive levels of x_1 and x_2 here, to have U maximized without any constraint, the consumer should purchase an infinite amount of both goods, a solution that obviously has little relevance. To render the optimization problem meaningful, the purchasing power of the consumer must also be taken into account. I f the consumer intends to spend a given sum, \$60, on the two goods and if the current prices are $P_{10} = 4$ and $P_{20} = 2$, then the *budget constraint* can be expressed by the linear equation

$$4x_1 + 2x_2 = 60. \tag{2}$$

The problem now is to maximize (1) subject to the constraint stated in (2). What the constraint does is to narrow the domain, and hence the range of the objective function. Graphically, the domain of (1) is represented by the nonnegative quadrant of the $x_1 x_2$ plane in Figure 12.1 (a). After the budget constraint (2) is added, the domain is immediately reduced to the set of points lying on the budget line. This will affect the range of the objective function too. Only that subsets of the utility surface lying directly above the budget-constraint line will now be relevant. The said subset may look like the curve in Figure 12.1 (b).

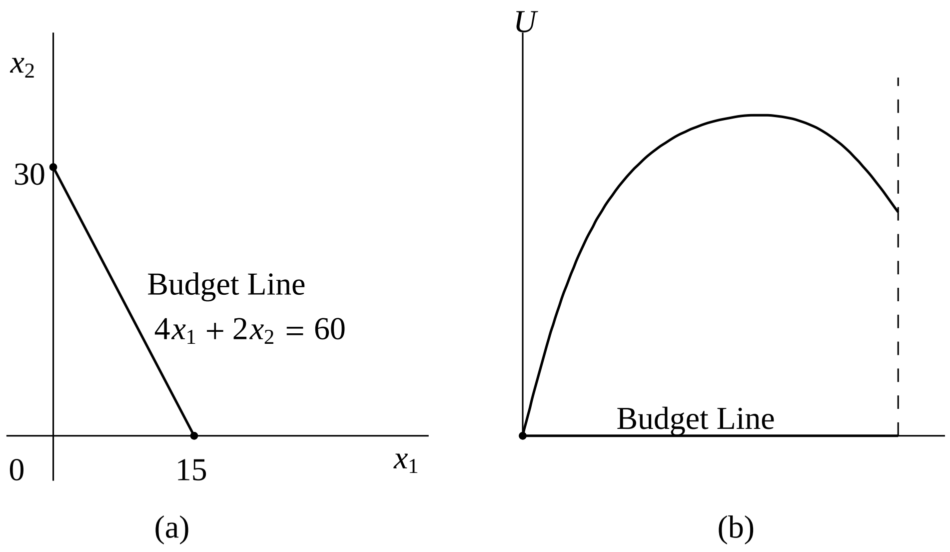

Figure 12.1

In general, for a function $z = f(x, y)$, the difference between a constrained extremum and a free extremum may be illustrated in the three-dimensional graph of Figure 12.2

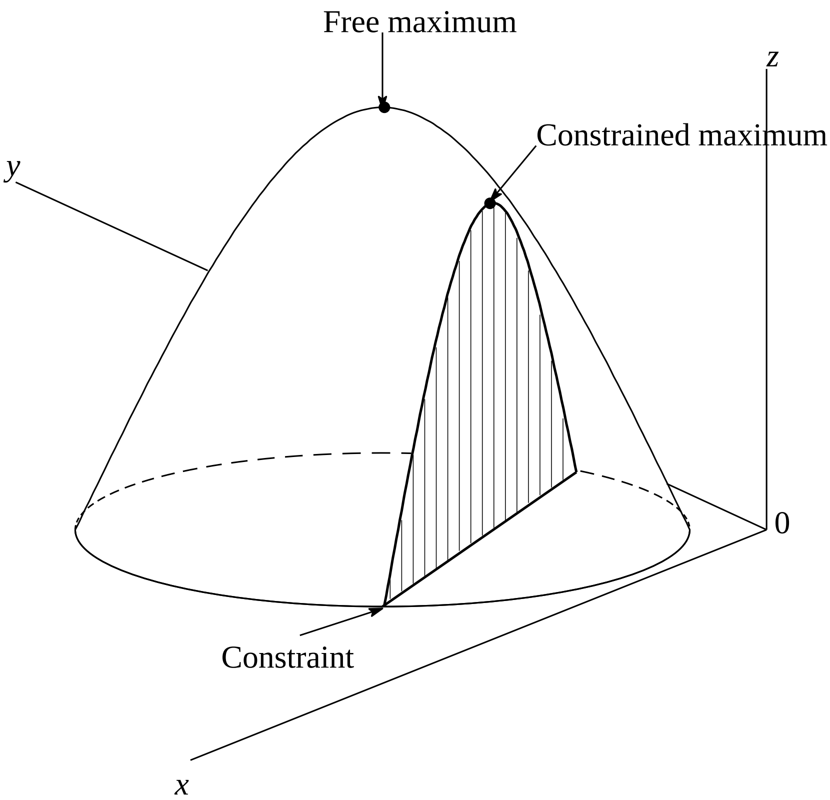

Figure 12.2

12.2 Finding the Stationary Values

Even without any new technique of solution, the constrained maximum in the simple example defined by (1) and (2) can be easily found. The constraint (2) implies

$$x_2 = 30 - 2x_1. \tag{3}$$

We can combine the constraint with the objective function by substituting the above equation into (1).

$$U = x_1(30 - 2x_1) + 2x_1 = 32x_1 - 2{x_1}^2 \tag{4}$$

By setting the first derivative equal to zero

$$\frac{dU}{dx_1} = 32 - 4x_1 = 0, \tag{5}$$

we get the solution $x_1 = 8$, which by virtue of (3) leads to $x_2 = 14$. Then, we can find the stationary value $U^* = 128$. Since the second derivative is

$$\frac{d^2U}{d{x_1}^2} = -4 < 0, \tag{6}$$

that stationary value constitutes a maximum of U.

• Lagrange-Multiplier Method

Given the problem of maximizing (1) subject to the constraint (2), let us write what is referred to as the *Lagrangian function*, which is a modified version of the objective function that incorporates the constraint.

$$Z = x_1 x_2 + 2x_1 + \lambda(60 - 4x_1 - 2x_2) \tag{7}$$

The symbol λ is called a *Lagrange multiplier*. If the constraint (2) is satisfied, then the last term in (7) will vanish regardless of the value of λ. So, Z will be identical with U. Moreover, we only have to seek the free maximum of Z, instead of the constrained maximum of U.

We simply treat λ as an additional choice variable in (7). The first-order conditions for free maximum of Z are given by

$$\begin{aligned} Z_\lambda &= 60 - 4x_1 - 2x_2 = 0, \\ Z_{x_1} &= x_2 + 2 - 4\lambda = 0, \\ Z_{x_2} &= x_1 - 2\lambda = 0. \end{aligned} \tag{8}$$

The first equation guarantees the satisfaction of the constraint. Solving (8) for the critical values of the variables, we find $x_1^* = 8$ and $x_2^* = 14$. Furthermore, it is clear from (7) that $Z^* = 128$. This is identical with the value of U^* found earlier.

Example 1

Find the extremum of

$$z = xy \quad \text{subject to} \quad x + y = 6. \tag{9}$$

The Lagrangian function is

$$Z = xy + \lambda(6 - x - y). \tag{10}$$

For a stationary value of Z, it is necessary that

$$Z_\lambda = 6 - x - y = 0, \tag{11}$$

$$Z_x = y - \lambda = 0, \tag{12}$$

$$Z_y = x - \lambda = 0. \tag{13}$$

Thus, we can find

$$\lambda^* = 3 \quad x^* = 3 \quad y^* = 3. \tag{14}$$

The stationary value is $Z^* = z^* = 9$, which needs to be tested against a second-order condition before we can tell whether it is a maximum or minimum.

Example 2

Find the extremum of

$$z = x_1{}^2 + x_2{}^2 \quad \text{subject to} \quad x_1 + 4x_2 = 2 \tag{15}$$

The Lagrangian function is

$$Z = x_1{}^2 + x_2{}^2 + \lambda(2 - x_1 - 4x_2). \tag{16}$$

The first-order necessary condition for a stationary value is

$$Z_\lambda = 6 - x_1 - 4x_2 = 0, \tag{17}$$
$$Z_1 = 2x_1 - \lambda = 0, \tag{18}$$
$$Z_2 = 2x_2 - 4\lambda = 0. \tag{19}$$

The stationary value of Z, defined by the solution

$$\lambda^* = \frac{4}{17} \qquad x_1^* = \frac{2}{17} \qquad x_2^* = \frac{8}{17} \tag{20}$$

is therefore $Z^* = z^* = \frac{4}{17}$

In general, given an objective function

$$z = f(x, y) \tag{21}$$

subject to the constraint

$$g(x, y) = c \tag{22}$$

where c is a constant, we can write the Lagrangian function as

$$Z = f(x, y) + \lambda[c - g(x, y)]. \tag{23}$$

For stationary value of Z, the first-order necessary condition is

$$\begin{aligned} Z_\lambda &= c - g(x, y) = 0, \\ Z_x &= f_x(x, y) - \lambda g_x(x, y) = 0, \\ Z_y &= f_y(x, y) - \lambda g_y(x, y) = 0. \end{aligned} \tag{24}$$

Since the first equation in (24) is simply a restatement of (22), the stationary values of the Lagrangian function Z will satisfy the constraint of the original function z. And since the expression $\lambda[c - g(x, y)]$ is assuredly zero, the stationary values of Z in (23) must be identical with those of (21) subject to (22).

• **Total-Differential Approach**

In the discussion of the free extremum of $z = f(x, y)$, it was learned that the first-order necessary condition may be stated in terms of the total differential dz:

$$dz = f_x dx + f_y dy = 0. \tag{25}$$

This statement remains valid after a constraint $g(x, y) = c$ is added. When $g(x, y) = c$ is added, however, then dg must be equal to dc, which is zero since c is a constant. Hence,

$$(dg =) g_x dx + g_y dy = 0. \tag{26}$$

By solving (26) for dy and substituting the result into (25), it should be clear that in order to satisfy this necessary condition, we must have

$$\frac{f_x}{g_x} = \frac{f_y}{g_y}. \tag{27}$$

The condition (27), together with the constraint $g(x, y) = c$, will provide the two equations from which to find the critical values of x and y.

Does the total-differential approach yield the same first-order condition as the Lagrangian multiplier method? Let us compare (24) with the result just obtained. The last two equations in (24) can be written as

$$\frac{f_x}{g_x} = \lambda \quad \text{and} \quad \frac{f_y}{g_y} = \lambda. \tag{28}$$

These conditions convey precisely the same information as (27). Note that whereas the total-differential approach yields only the value of x^* and y^*, the Lagrangian-multiplier method also gives the value of λ^*.

We can rewrite the condition (28) as

$$Df = \lambda Dg \tag{29}$$

where

$$Df = \begin{bmatrix} f_x \\ f_y \end{bmatrix} \quad \text{and} \quad Dg = \begin{bmatrix} g_x \\ g_y \end{bmatrix}. \tag{30}$$

Df and Dg are the vectors of partial derivatives of f and g with respect to each of their arguments. The vectors Df and Dg are called the *gradients* of f and g, respectively. The gradient of a function $f(\cdot)$ evaluated at a point (x^*, y^*) is a vector perpendicular to the tangent line of the function at that point (see Figure 12.3). The condition (29) says that the first-order necessary condition for (x^*, y^*) to be a constrained extremum is for the gradient of the constraint to be proportional to the gradient of the objective function at that point.

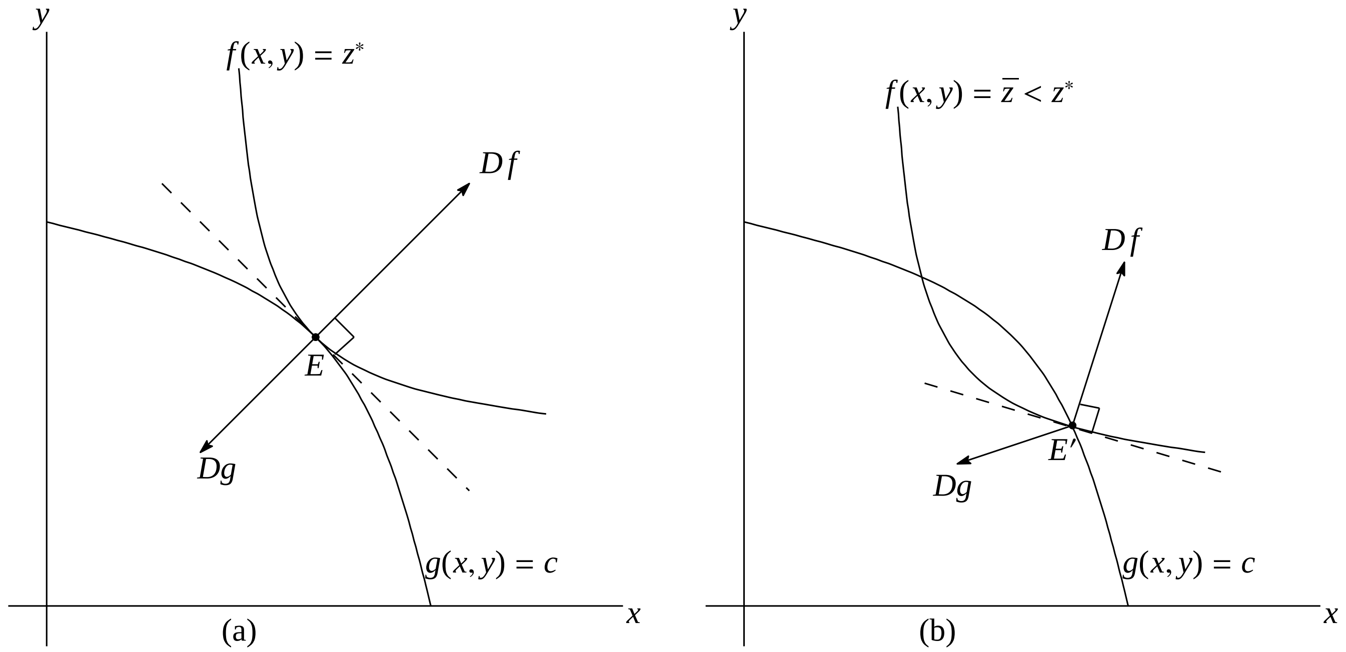

Figure 12.3

In figure 12.3, we plot the two contour lines $f(x, y) = z^*$ and $f(x, y) = \bar{z} < z^*$ together with the constraint $g(x, y) = c$. In figure 12.3 (a) where the gradient of the constraint to be proportional to the gradient of the objective function at point E, the stationary value z^* is reached. At point E' in figure 12.3 (b), however, $\bar{z}$ is not the stationary value.

• An Interpretation of the Lagrange Multiplier

As we shall presently demonstrate, λ^* provides a measure of the sensitivity Z^* (and z^*) to a shift of the constant c.

We again resort to the implicit function theorem. Taking the three equations in (24) to be in the form of $F^j(\lambda, x, y; c) = 0$ (with $j = 1, 2, 3$), and assuming them to have continuous partial derivatives. And also we assume

$$|J| = \begin{vmatrix} \frac{\partial F^1}{\partial \lambda} & \frac{\partial F^1}{\partial x} & \frac{\partial F^1}{\partial y} \\ \frac{\partial F^2}{\partial \lambda} & \frac{\partial F^2}{\partial x} & \frac{\partial F^2}{\partial y} \\ \frac{\partial F^3}{\partial \lambda} & \frac{\partial F^3}{\partial x} & \frac{\partial F^3}{\partial y} \end{vmatrix} = \begin{vmatrix} 0 & -g_x & -g_y \\ -g_x & f_{xx} - \lambda g_{xx} & f_{xy} - \lambda g_{xy} \\ -g_y & f_{xy} - \lambda g_{xy} & f_{yy} - \lambda g_{yy} \end{vmatrix} \neq 0. \tag{31}$$

Then, we can express λ^*, x^* and y^* as implicit functions of the parameter c:

$$\lambda^* = \lambda^*(c) \qquad x^* = x^*(c) \qquad \text{and} \qquad y^* = y^*(c) \tag{32}$$

of all which will have continuous derivatives. Also, we have the equilibrium identities

$$\begin{aligned} c - g(x^*, y^*) &\equiv 0, \\ f_x(x^*, y^*) - \lambda^* g_x(x^*, y^*) &\equiv 0, \\ f_y(x^*, y^*) - \lambda^* g_y(x^*, y^*) &\equiv 0. \end{aligned} \tag{33}$$

Now since the optimal value of Z depends on λ^*, x^* and y^*, that is,

$$Z^* = f(x^*, y^*) + \lambda^*[c - g(x^*, y^*)], \tag{34}$$

we may, in view of (32), consider Z^* to be a function of c alone. Differentiating Z^* totally with respect to c, we find

$$\begin{aligned} \frac{dZ^*}{dc} &= f_x \frac{dx^*}{dc} + f_y \frac{dy^*}{dc} + [c - g(x^*, y^*)] \frac{d\lambda^*}{dc} + \lambda^* \left(1 - g_x \frac{dx^*}{dc} - g_y \frac{dy^*}{dc}\right) \\ &= (f_x - \lambda^* g_x) \frac{dx^*}{dc} + (f_y - \lambda^* g_y) \frac{dy^*}{dc} + [c - g(x^*, y^*)] \frac{d\lambda^*}{dc} + \lambda^* \\ &= \lambda^*. \end{aligned} \tag{35}$$

In the last line, we use the identities (33). The above equation shows that the solution value of the Lagrangian multiplier constitutes a measure of the effect of a change in the constraint via the parameter c on the optimal value of the objective function.

12.3 Second-Order Conditions

• Second-Order Total Differential

Since the constraint $g(x, y) = c$ means $dg = g_x dx + g_y dy$ as in (26), dx and dy no longer are both arbitrary. If we take dx as an arbitrary change, then dy must satisfy

$$dy = -\frac{g_x}{g_y} dx. \tag{36}$$

To find an appropriate new expression for d^2z, we must treat dy as a variable dependent on x and y during differentiation.

$$\begin{aligned} d^2z = d(dz) &= \frac{\partial(dz)}{\partial x}dx + \frac{\partial(dz)}{\partial y}dy \\ &= \frac{\partial}{\partial x}(f_x dx + f_y dy)dx + \frac{\partial}{\partial y}(f_x dx + f_y dy)dy \\ &= \left[f_{xx}dx + \left(f_{xy}dy + f_y\frac{\partial(dy)}{\partial x}\right)\right]dx + \left[f_{yx}dx + \left(f_{yy}dy + f_y\frac{\partial(dy)}{\partial y}\right)\right]dy \\ &= f_{xx}dx^2 + f_{xy}dydx + f_y\frac{\partial(dy)}{\partial x}dx + f_{yx}dxdy + f_{yy}dy^2 + f_y\frac{\partial(dy)}{\partial y}dy. \end{aligned}$$

Since the next equation holds

$$f_y\left[\frac{\partial(dy)}{\partial x}dx + \frac{\partial(dy)}{\partial y}dy\right] = f_y d(dy) = f_y d^2y,$$

then the desired expression for d^2z is

$$d^2z = f_{xx}dx^2 + 2f_{xy}dxdy + f_{yy}dy^2 + f_y d^2y. \tag{37}$$

This differs from the second-order total differential without a constraint only by the last term, $f_y d^2y$. It should be noted that this last term is in the first degree.

By totally differentiating the constraint twice, we can get

$$(d^2g =)g_{xx}dx^2 + 2g_{xy} + g_{yy}dy^2 + g_y d^2y = 0. \tag{38}$$

Solving this last equation for d^2y and substituting the result in (37), we can write d^2z as the following quadratic form:

$$d^2z = \left(f_{xx} - \frac{f_y}{g_y}g_{xx}\right)dx^2 + 2\left(f_{xy} - \frac{f_y}{f_y}g_{xy}\right)dxdy + \left(f_{yy} - \frac{f_y}{g_y}g_{yy}\right)dy^2 \tag{39}$$

By partially differentiating the derivatives in (24) and using (28), we find the following second derivatives

$$\begin{aligned} Z_{xx} &= f_{xx} - \lambda g_{xx}, \\ Z_{xy} &= f_{xy} - \lambda g_{xy} \quad = Z_{yx}, \\ Z_{yy} &= f_{yy} - \lambda g_{yy}. \end{aligned} \tag{40}$$

Then, we can finally express d^2z more neatly as follows:

$$\begin{aligned} d^2z &= Z_{xx}dx^2 + 2Z_{xy}dxdy + Z_{yy}dy^2 \\ &= \quad Z_{xx}dx^2 \quad + Z_{xy}dxdy \\ &\quad + Z_{yx}dydx \quad + Z_{yy}dy^2 \end{aligned} \tag{41}$$

• Second-Order Conditions

For a constrained extremum of $z = f(x, y)$ subject to $g(x, y) = c$, the second-order sufficient conditions revolve around the algebraic sign of the second-order total differential d^2z, evaluated at a stationary point. However, there is one important change. In the present context, we are concerned with the sign definiteness of d^2z, not for all possible values of dx and dy, but only for those dx and dy values satisfying the linear constraint (26), $dg = g_x dx + g_y dy = 0$.

Thus the second-order sufficient conditions are

$$\text{For maximum of } z: \quad d^2z \text{ negative definite, subject to } dg = 0$$

$$\text{For minimum of } z: \quad d^2z \text{ positive definite, subject to } dg = 0$$

• The Bordered Hessian

In place of the Hessian determinant $|H|$, however, in the constrained-extremum case we shall encounter what is known as a *bordered Hessian*.

Let us first analyze the conditions for the sign definiteness of a two-variable quadratic form subject to a linear constraint:

$$q = au^2 + 2huv + bv^2 \quad \text{subject to} \quad \alpha u + \beta v = 0. \tag{42}$$

Since the constraint implies $v = -(\alpha/\beta)u$, we can rewrite q as a function of one variable only:

$$\begin{aligned} q &= au^2 - 2h\frac{\alpha}{\beta}u^2 + b\frac{\alpha^2}{\beta^2}u^2 \\ &= (a\beta^2 - 2h\alpha\beta + b\alpha^2)\frac{u^2}{\beta^2} \end{aligned} \tag{43}$$

It is obvious that q is positive (negative) definite if and only if the expression in parentheses is positive (negative). Now, it so happens that the following symmetric determinant

$$\begin{vmatrix} 0 & \alpha & \beta \\ \alpha & a & h \\ \beta & h & b \end{vmatrix} = 2h\alpha\beta - a\beta^2 - b\alpha^2 \tag{44}$$

is exactly the *negative* of the said parenthetical expression. Consequently, we can state that

$$q \text{ is } \left\{ \begin{matrix} \text{positive definite} \\ \text{negative definite} \end{matrix} \right\} \text{subject to } \alpha u + \beta v = 0 \quad \text{iff} \quad \begin{vmatrix} 0 & \alpha & \beta \\ \alpha & a & h \\ \beta & h & b \end{vmatrix} \left\{ \begin{matrix} < 0 \\ > 0 \end{matrix} \right. \tag{45}$$

When applied to the quadratic form d^2z in (41), the variables u and v become dx and dy, and the Hessian $\begin{vmatrix} Z_{xx} & Z_{xy} \\ Z_{yx} & Z_{yy} \end{vmatrix}$. More, we have $\alpha = g_x$ and $\beta = g_y$. We have the following determinantal criterion for the sign definiteness of d^2z :

$$d^2z \text{ is } \left\{ \begin{matrix} \text{positive definite} \\ \text{negative definite} \end{matrix} \right\} \text{subject to } dg = 0 \quad \text{iff} \quad \begin{vmatrix} 0 & g_x & g_y \\ g_x & Z_{xx} & Z_{xy} \\ g_y & Z_{yx} & Z_{yy} \end{vmatrix} \left\{ \begin{matrix} < 0 \\ > 0 \end{matrix} \right. \tag{46}$$

The determinant to the right, often referred to as a bordered Hessian, shall be denoted by $|\bar{H}|$.

We may conclude that, given a stationary value of $z = f(x, y)$ or of $Z = f(x, y) + \lambda[c - g(x, y)]$, a positive $|\bar{H}|$ is sufficient to establish it as a relative maximum of z; similarly, a negative $|\bar{H}|$ is sufficient to establish it as a minimum.

Example 1

Let us return to Example 1 of Section 12.2 and ascertain whether the stationary value found there gives a maximum or a minimum. Since $Z_x = y - \lambda$ and $Z_y = x - \lambda$, the second-order partial derivatives are $Z_{xx} = 0$, $Z_{xy} = Z_{yx} = 1$, and $Z_{yy} = 0$. The border elements we need are $g_x = 1$ and $g_y = 1$. Thus we find that

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

which establishes the value $z^* = 9$ as a maximum.

Example 2

Continuing on to Example 2 of Section 12.2, we see that $Z_1 = 2x_1 - \lambda$ and $Z_2 = 2x_2 - 4\lambda$. These yield $Z_{11} = 2$, $Z_{12} = Z_{21} = 0$ and $Z_{22} = 2$. From the constraint $x_1 + 4x_2 = 2$, we obtain $g_1 = 1$ and $g_2 = 4$. The bordered Hessian is

$$|\bar{H}| = \begin{vmatrix} 0 & 1 & 4 \\ 1 & 2 & 0 \\ 4 & 0 & 2 \end{vmatrix} = -34 < 0.$$

Hence, the value $z^* = \frac{4}{17}$ is a minimum.

Example 3

Consider a simple two-period model where a consumer's utility is a function of consumption in both periods.

$$U(x_1, x_2) = x_1 x_2 \tag{47}$$

where x_1 is consumption in period 1 and x_2 is consumption in period 2. The consumer's intertemporal budget constraint is given by

$$x_1 + \frac{x_2}{1+r} = B. \tag{48}$$

The Lagrangian function for this utility maximization problem is

$$Z = x_1 x_2 + \lambda \left(B - x_1 - \frac{x_2}{1+r} \right). \tag{49}$$

The first-order necessary conditions are

$$\frac{\partial Z}{\partial \lambda} = B - x_1 - \frac{x_2}{1+r} = 0, \tag{50}$$

$$\frac{\partial Z}{\partial x_1} = x_2 - \lambda = 0, \tag{51}$$

$$\frac{\partial Z}{\partial x_2} = x_1 - \frac{\lambda}{1+r} = 0. \tag{52}$$

The last two equation implies

$$\frac{x_2}{x_1} = \frac{\lambda}{\lambda/(1+r)} = 1 + r. \tag{53}$$

Substituting this equation into the budget constraint yields the solution

$$x_1^* = \frac{B}{2} \quad \text{and} \quad x_2^* = \frac{B(1+r)}{2}. \tag{54}$$

The bordered Hessian for this problem is

$$|\bar{H}| = \begin{vmatrix} 0 & -1 & -\frac{1}{1+r} \\ -1 & 0 & 1 \\ -\frac{1}{1+r} & 1 & 0 \end{vmatrix} = \frac{2}{1+r} > 0. \tag{55}$$

Thus the second-order sufficient condition is satisfied for a maximum U.

12.5 Utility Maximization and Consumer Demand

Let us reexamine the maximization of a utility function cited in Section 12.1.

The problem is to maximize a smooth utility function

$$U = U(x, y) \quad (U_x, U_y > 0) \tag{56}$$

subject to

$$xP_x + yP_y = B. \tag{57}$$

• **First-Order Condition**

The Lagrangian function of this problem is

$$Z = U(x, y) + \lambda(B - xP_x - yP_y). \tag{58}$$

As the first-order condition, we have the following set of simultaneous equations:

$$\begin{aligned} Z_\lambda &= B - xP_x - yP_y = 0 \\ Z_x &= U_x - \lambda P_x = 0 \\ Z_y &= U_y - \lambda P_y = 0 \end{aligned} \tag{59}$$

The last two equations are equivalent to

$$\frac{U_x}{P_x} = \frac{U_y}{P_y} = \lambda. \tag{60}$$

The first-order condition calls for the satisfaction of (60) subject to the budget constraint. What (60) states that, in order to maximize utility, consumers must allocate their budgets so as to equalize the ratio of marginal utility to price for every commodity. In the optimum, these ratios should have the common value λ^*. The optimum value of the Lagrangian multiplier can be interpreted as the *marginal utility of money* when the consumer's utility is maximized.

If we restate the condition in (60) in the form

$$\frac{U_x}{U_y} = \frac{P_x}{P_y} \tag{61}$$

the first-order condition can be given an alternative interpretation, in terms of indifference curves.

An *indifference curve* is defined as the locus of the combinations of x and y that will yield a constant level of U. On an indifference curve, we must have

$$dU = U_x dx + U_y dy = 0 \tag{62}$$

with the implication that

$$dy/dx = -U_x/U_y. \tag{63}$$

An indifference curve has a slope which is equal to the negative of the marginal-utility ratio U_x/U_y. Note that U_x/U_y, the negative of the indifference-curve slope, is called the *marginal rate of substitution* between the two goods.

What about the meaning of P_x/P_y? The budget constraint, $xP_x + yP_y = B$, can be written alternatively as

$$y = \frac{B}{P_y} - \frac{P_x}{P_y}x. \tag{64}$$

So, the negative of P_x/P_y is the slope of the budget constraint.

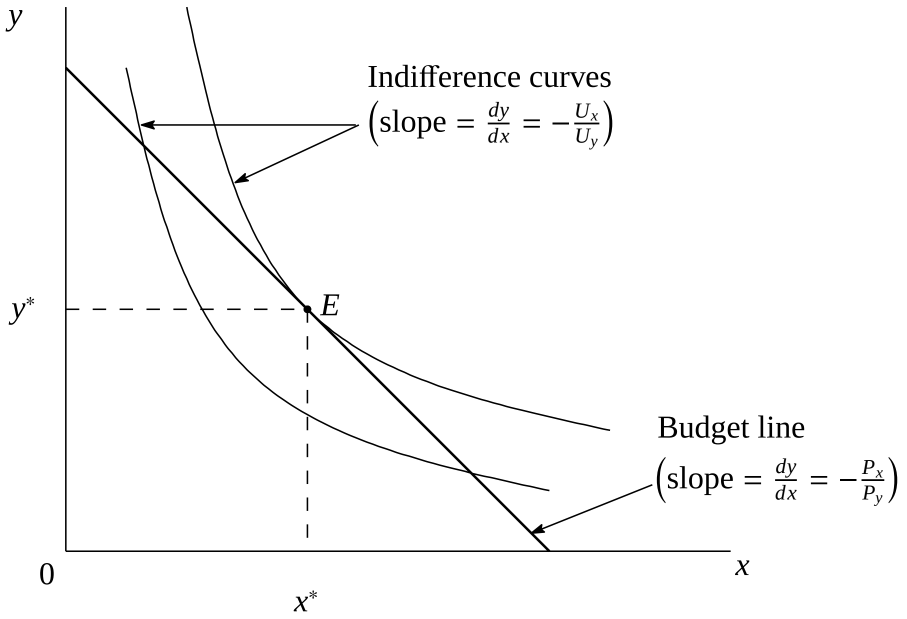

Figure 12.4

To maximize utility, a consumer must allocate the budget such that the slope of the budget line is equal to the slope of some indifference curve. This condition is met at point E in Figure 12.4.

• Second-Order Condition

If the bordered Hessian in the present problem is positive,

$$|\bar{H}| = \begin{vmatrix} 0 & P_x & P_y \\ P_x & U_{xx} & U_{xy} \\ P_y & U_{yx} & U_{yy} \end{vmatrix} = 2P_xP_yU_{xy} - {P_y}^2U_{xx} - {P_x}^2U_{yy} > 0, \tag{65}$$

then the stationary value of U will assuredly be a maximum. A positive $|\bar{H}|$ means the strict convexity of the indifference curve at the point of tangency E. The strict convexity would be ensured by a positive d^2y/dx^2. By totally differentiating (63), we can get the expression for d^2y/dx^2

$$\frac{d^2y}{dx^2} = \frac{d}{dx}\left(-\frac{U_x}{U_y}\right) = -\frac{1}{{U_y}^2}\left(U_y\frac{dU_x}{dx} - U_x\frac{dU_y}{dx}\right) \tag{66}$$

As shown in Figure 12.4, y is itself a function of x along an indifference curve. Then, we have

$$\frac{dU_x}{dx} = U_{xx} + U_{yx}\frac{dy}{dx} \qquad \frac{dU_y}{dx} = U_{xy} + U_{yy}\frac{dy}{dx} \tag{67}$$

where dy/dx refers to the slope of the indifference curve. At the point of tangency E, this slope is identical with that of the budget constraint. Thus we can rewrite (67) as

$$\frac{dU_x}{dx} = U_{xx} - U_{yx}\frac{P_x}{P_y} \qquad \frac{dU_y}{dx} = U_{xy} - U_{yy}\frac{P_x}{P_y} \tag{68}$$

Substituting (68) into (66) and utilizing the information that

$$U_x = \frac{U_y P_x}{P_y} \tag{69}$$

we can finally transform (66) into

$$\frac{d^2y}{dx^2} = \frac{2P_xP_yU_{xy} - P_y{}^2U_{xx} - P_x{}^2U_{yy}}{U_yP_y{}^2} = \frac{|\bar{H}|}{U_yP_y{}^2} \tag{70}$$

It is clear that when the second-order sufficient condition (65) is satisfied, the second derivative in (70) is positive, and the relevant indifferent curve is strictly convex at the point of tangency.

12.6 Homogeneous Functions

Definition
A function $f(x)$ is homogeneous of degree r, if

$$f(jx_1,\ \dots,\ jx_n) = j^r f(x_1,\ \dots,\ x_n) \tag{71}$$

Example 1

$$f(x, y, z) = x/y + 2z/3x \tag{72}$$

$$f(jx,\ jy,\ jz) = \frac{jx}{jy} + \frac{2jz}{3jx} = \frac{x}{y} + \frac{2z}{3x} = j^0 f(x, y, z) \tag{73}$$

The function f is homogeneous of degree 0.

Example 2

$$g(x, y, z) = x^2/y + 2z^2/3x \tag{74}$$

$$g(jx,\ jy,\ jz) = \frac{(jx)^2}{jy} + \frac{2(jz)^2}{3jx} = j\left(\frac{x}{y} + \frac{2z}{3x}\right) = jg(x, y, z) \tag{75}$$

The function g is homogeneous of degree 1.

• Linear Homogeneity
In the discussion of production functions, wide use is made of homogeneous functions of the first degree. These functions are often referred to as *linearly homogeneous* functions. Let us adopt as the framework of our discussion a production function in the form, say,

$$Q = f(K, L) \tag{76}$$

The mathematical assumption of linear homogeneity would amount to the economic assumption of constant returns to scale. What unique properties characterize this linearly homogeneous production function?

▷ **Property I** The average physical product of labor (APP_L) and of capital (APP_K) can be expressed as functions of the capital-labor ratio, $k \equiv K/L$, alone.

Proof

Since $f(K, L)$ is linearly homogeneous, we have $f(jK, jL) = jf(K, L) = jQ$. Substituting $1/L$ into j, we obtain

$$f\left(\frac{K}{L}, \frac{L}{L}\right) = f\left(\frac{K}{L}, 1\right) = f(k, 1) = \frac{Q}{L}. \tag{77}$$

Let $\phi(k)$ denote $f(k, 1)$.

$$APP_L \equiv \frac{Q}{L} = \phi(k) \tag{78}$$

$$APP_K \equiv \frac{Q}{K} = \frac{Q}{L}\frac{L}{K} = \frac{\phi(k)}{k} \tag{79}$$

Since both average products depend on k alone, both APP_L and APP_K are homogeneous of degree zero in the variables K and L.

▷ **Property II** The marginal physical products MPP_L and MPP_K can be expressed as functions of k alone.

Proof
We first differentiate k with respect to K and L.

$$\frac{\partial k}{\partial K} = \frac{\partial}{\partial K}\left(\frac{K}{L}\right) = \frac{1}{L} \qquad \frac{\partial k}{\partial L} = \frac{\partial}{\partial L}\left(\frac{K}{L}\right) = -\frac{K}{L^2} \tag{80}$$

Next, we rewrite (77) as

$$Q = L\phi(k) \tag{81}$$

and then differentiate Q with respect to K and L using (80).

$$\begin{aligned} MPP_K \equiv \frac{\partial Q}{\partial K} &= \frac{\partial L\phi(k)}{\partial K} \\ &= L\frac{\partial \phi(k)}{\partial K} = L\frac{d\phi(k)}{dk}\frac{\partial k}{\partial K} \\ &= L\phi'(k)\left(\frac{1}{L}\right) = \phi'(k) \end{aligned} \tag{82}$$

$$\begin{aligned} MPP_L \equiv \frac{\partial Q}{\partial L} &= \frac{\partial L\phi(k)}{\partial L} \\ &= \phi(k) + L\frac{\partial \phi(k)}{\partial L} \\ &= \phi(k) + L\frac{d\phi(k)}{dk}\frac{\partial k}{\partial L} \\ &= \phi(k) + L\phi'(k)\frac{\partial k}{\partial L} \\ &= \phi(k) + L\phi'(k)\frac{-K}{L^2} \\ &= \phi(k) - k\phi'(k) \end{aligned} \tag{83}$$

Since both marginal products depend on k alone, both MPP_L and MPP_K are homogeneous of degree zero in the variables K and L.

▷ **Property III(Euler's Theorem)** If $Q = f(K, L)$ is linearly homogeneous, then

$$K\frac{\partial Q}{\partial K} + L\frac{\partial Q}{\partial L} \equiv Q \tag{84}$$

Proof

$$\begin{aligned} K\frac{\partial Q}{\partial K} + L\frac{\partial Q}{\partial L} &= K\phi'(k) + L[\phi(k) - k\phi'(k)] \\ &= K\phi'(k) + L\phi(k) - K\phi'(k) \\ &= L\phi(k) \\ &= Q \end{aligned} \tag{85}$$

Note that this result is valid for any values K and L.

• Cobb-Douglas Production Function

One specific production function widely used in economic analysis is the *Cobb-Douglas production function*:

$$Q = AK^{\alpha}L^{1-\alpha}. \tag{86}$$

where A is a positive constant, and $\alpha \in [0, 1]$. What we shall consider here first is a generalized version of this function.

$$Q = AK^{\alpha}L^{\beta}. \tag{87}$$

Some of the major features of this function are:

1. it is homogeneous of degree $\alpha + \beta$;

$$A(jK)^{\alpha}(jL)^{\beta} = j^{\alpha+\beta}AK^{\alpha}L^{\beta} = j^{\alpha+\beta}Q. \tag{88}$$

2. in the special case of $\alpha + \beta = 1$, it is linearly homogeneous;

$$A(jK)^{\alpha}(jL)^{\beta} = j^{\alpha+\beta}Q = jQ. \tag{89}$$

3. its isoquants are negatively sloped throughout and strictly convex for positive values of K and L;
   For any positive output Q_0, (87) can be written as

$$AK^{\alpha}L^{\beta} = Q_0. \tag{90}$$

   Taking the natural log of both sides , we find that

$$\ln A + \alpha \ln K + \beta \ln L - \ln Q_0 = 0. \tag{91}$$

   By the implicit-function rule and the log rule, we have

$$\frac{dK}{dL} = -\frac{\beta K}{\alpha L} < 0. \tag{92}$$

   Then,

$$\frac{d^2K}{dL^2} = \frac{d}{dL}\left(-\frac{\beta K}{\alpha L}\right) = -\frac{\beta}{\alpha}\frac{1}{L^2}\left(L\frac{dK}{dL} - K\right) > 0. \tag{93}$$

   Thus, the isoquant is downward-slopping throughout and strictly convex in the LK plane for positive values of K and L.

• $\alpha + \beta = 1$ case
The total product in this special case is expressed as

$$Q = AK^{\alpha}L^{1-\alpha} = A\left(\frac{K}{L}\right)^{\alpha} L = LAk^{\alpha}. \tag{94}$$

Therefore, the average products are

$$APP_L = \frac{Q}{L} = Ak^{\alpha} \tag{95}$$

$$APP_K = \frac{Q}{K} = \frac{LAk^{\alpha}}{K} = \frac{Ak^{\alpha}}{k} = Ak^{\alpha-1} \tag{96}$$

And the marginal products are

$$\frac{\partial Q}{\partial K} = A\alpha K^{\alpha-1}L^{-(\alpha-1)} = A\alpha\left(\frac{K}{L}\right)^{\alpha-1} = A\alpha k^{\alpha-1} \tag{97}$$

$$\frac{\partial Q}{\partial L} = AK^{\alpha}(1-\alpha)L^{-\alpha} = A(1-\alpha)\left(\frac{K}{L}\right)^{\alpha} = A(1-\alpha)k^{\alpha} \tag{98}$$

We can verify Euler's theorem.

$$\begin{aligned} K\frac{\partial Q}{\partial K} + L\frac{\partial Q}{\partial L} &= KA\alpha k^{\alpha-1} + LA(1-\alpha)k^{\alpha} \\ &= LAk^{\alpha}\left[\frac{K\alpha}{Lk} + 1 - \alpha\right] \\ &= LAk^{\alpha}[\alpha + 1 - \alpha] = LAk^{\alpha} = Q \end{aligned} \tag{99}$$

If each input is assumed to be paid by the amount of its marginal product, the relative share of total product accruing to capital will be

$$\frac{K(\partial Q/\partial K)}{Q} = \frac{KA\alpha k^{\alpha-1}}{LAk^{\alpha}} = \alpha. \tag{100}$$

Similarly, labor's relative share will be

$$\frac{L(\partial Q/\partial L)}{Q} = \frac{LA(1-\alpha)k^{\alpha}}{LAk^{\alpha}} = 1 - \alpha. \tag{101}$$

Thus the exponent of each input variable indicates the relative share of that input in the total product.