2 Lagrange's Method

Statement of the Problem

Let us begin the development of the general theory of optimization subject to constraints, using a setting very close to the consumer choice model of Chapter 1. Suppose the choice variables are x_1 and x_2. I shall write them more compactly as a vector x arranged in a column,

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Initially, I shall use vectors only to abbreviate lists of components; actual operations with vectors and matrices will appear gradually.

We will need notation that distinguishes between the general vector x and some particular value of x such as the optimum, while remembering the family resemblance between the two: both are vectors of choice variables. I shall generally use the symbol $\bar{x}$ to denote the optimum value of the general variable x. The components of $\bar{x}$ will be written $\bar{x}_1$ etc.

The function to be maximized, called the objective function, is $F(x)$. The constraint is a general non-linear one,

$$G(x) = c, \tag{2.1}$$

where G is a function and c a given constant. The model of Chapter 1 was a special case of this: F was the utility function, G was a linear function showing the expenditure,

$$G(x) = p_1\, x_1 + p_2\, x_2,$$

and c was income. If it helps, you can now think of G as a more general non-linear expenditure or cost function, such as would arise if the consumer faced a quantity discount or premium price schedule.

With this notation, the problem in this chapter is to find the value $\bar{x}$ that maximizes $F(x)$ subject to $G(x) = c$.

The Arbitrage Argument

As we did in Chapter 1, start with a trial point and an infinitesimal change. Let the particular trial point be $\bar{x}$, and the infinitesimal change

$$dx = \begin{pmatrix} dx_1 \\ dx_2 \end{pmatrix}.$$

We want $(x + dx)$ to satisfy the constraint, and see if it yields a higher value of the objective function.

Since the change in x is infinitesimal, we can approximate the changes in values in functions of x by the first-order linear terms in their Taylor series. Let subscripts on a function denote its partial derivatives with respect to the indicated argument; for example F_1 is $\partial F/\partial x_1$. We must remember that in general each partial derivative is itself a function of the whole vector x. Only in the special case where F is linear will the partial derivatives be constants; only in the special case where F is additively separable (a function of x_1 plus a function of x_2) will F_1 be independent of x_2 and vice versa. Thus we should write the partial derivatives as functions $F_1(x)$ and $F_2(x)$. When these are evaluated at our initial trial point x, their values will be written $F_1(\bar{x})$ and $F_2(\bar{x})$.

The first-order Taylor approximation of the change in $F(x)$ as a result of the infinitesimal move from the general point x to $x + dx$ is

$$\begin{aligned} dF(x) &= F(x + dx) - F(x) \\ &= F_1(x)\, dx_1 + F_2(x)\, dx_2. \end{aligned} \tag{2.2}$$

Observe the similarity with the expression for the change in utility (1.3) of Chapter 1; the marginal utilities that were motivated there by economic intuition are simply the partial derivatives of the utility function with respect to the amounts of the two goods consumed.

There is a similar expression for the change in $G(x)$:

$$dG(x) = G_1(x)\, dx_1 + G_2(x)\, dx_2. \tag{2.3}$$

In Chapter 1, the partial derivatives of G were simply the prices. If we now think of G as a more general non-linear outlay or expenditure function, the partial derivatives are the marginal prices of the respective commodities.

Now we can modify the arbitrage argument of Chapter 1 to apply to the new more general setting. Start at a point $\bar{x}$ where the constraint (2.1) holds, and consider a change dx such that $(x + dx)$ also satisfies the constraint. Then $dG(\bar{x}) = 0$. Using (2.3) with the particular initial point, we have

$$G_1(\bar{x})\, dx_1 = -G_2(\bar{x})\, dx_2.$$

Call the common value of these two sides dc. Then our arbitrage consists of reallocating an amount dc in the value of the function $G(x)$ away from x_2 and toward x_1.

First suppose $G_1(\bar{x})$ and $G_2(\bar{x})$ are both non-zero. Then

$$dx_1 = dc/G_1(\bar{x}), \qquad \text{and} \qquad dx_2 = dc/G_2(\bar{x}).$$

The resulting change in the value of the objective function is found by substituting into (2.2) as

$$dF(\bar{x}) = [\, F_1(\bar{x})/G_1(\bar{x}) - F_2(\bar{x})/G_2(\bar{x})\,]\; dc. \tag{2.4}$$

If the bracketed expression is non-zero, then $F(x)$ can be increased by choosing dc to have the same sign as that of the bracketed expression. As before, we can turn this around to find a no-arbitrage condition that holds when $\bar{x}$ is the optimum choice. So long as neither $\bar{x}_1$ nor $\bar{x}_2$ has hit some natural boundary such as zero, then changes dc of either sign are possible. If x is optimum, then no such change should be capable of increasing $F(x)$. Therefore the bracketed expression in (2.4) should be zero, or

$$F_1(\bar{x})/G_1(\bar{x}) = F_2(\bar{x})/G_2(\bar{x}). \tag{2.5}$$

This is the analog of the condition (1.6) of the previous chapter.

Note the exact statement: *if* the optimum choice is $\bar{x}$, *then* it satisfies (2.5). I have not established any implication the other way round, so there is no guarantee that a solution to (2.5) is the optimum. This is the difference between necessary and sufficient conditions, and I will discuss it in more detail later in this chapter.

If $\bar{x}$ lies at some natural boundary, for example if one of $\bar{x}_1$ and $\bar{x}_2$ is zero when both must be non-negative, then only one-sided changes dc are meaningful, and we get an inequality that corresponds to (1.4) and (1.5). I shall not consider this case in this chapter, but shall return to it in the next.

As in Chapter 1, let us define λ to be the common value of the two sides in (2.5). Then we can write that equation as a set of two equations

$$F_j(\bar{x}) = \lambda\; G_j(\bar{x}), \qquad j = 1, 2. \tag{2.6}$$

Remember that the λ of Chapter 1 could be interpreted as the marginal utility of income. In the same way, the λ just introduced turns out to be the rate at which the optimum value of $F(x)$ responds to a change in c. I shall develop this interpretation and its implications in Chapter 4. The main topic in the rest of this chapter involves writing (2.6) in a way that easily extends to more general settings, and provides a method for finding the optimum x. But first a couple of necessary digressions.

Constraint Qualification

What happens if say $G_1(\bar{x})$ is zero? Now $\bar{x}_1$ can be changed slightly without affecting the constraint. If $F_1(\bar{x})$ is not zero, it is desirable to do so. For example, if $F_1(x)$ is positive, then $F(x)$ can be increased by raising x_1. This goes on until either $F_1(x)$ drops to zero, or $G_1(x)$ becomes non-zero. In the consumption interpretation, a consumer will go on using more of a free good either until he is satiated, or until the marginal unit of the good is no longer free. Therefore if $G_1(\bar{x})$ is zero and $\bar{x}$ is optimum, then $F_1(\bar{x})$ must be zero, too. We can define the ratio of these two zeros as we please, and there is no harm in defining it so that (2.5) is satisfied.

What if $G_1(\bar{x})$ and $G_2(\bar{x})$ are both zero? This might mean that both goods are free, and should be consumed to the point of satiation. But there is a more ominous possibility arising from the quirks of algebra and calculus. Take the budget line (1.1) of the previous chapter, and write its equation as

$$(\, p_1\, x_1 + p_2\, x_2 - I\,)^3 = 0.$$

This is an unnecessarily complicated way of writing (1.1), but the two are mathematically fully equivalent, and we should see if the change makes any difference. Let $G(x)$ be the function on the left-hand side of this. Then

$$G_1(x) = 3\; p_1\; (\, p_1\, x_1 + p_2\, x_2 - I\,)^2,$$

which is always zero when x satisfies the budget constraint. The same is true for $G_2(x)$. Goods are not free at the margin, and yet their quantities have zero effect on the constraint function. In such a case, our method runs into trouble.

The formal mathematical theory cops out and simply refuses to deal with such cases by assuming a condition called a Constraint Qualification. In the present instance that simply amounts to assuming that at least one of $G_1(\bar{x})$ and $G_2(\bar{x})$ is non-zero. If this is not true in a particular application, then conditions like (2.6) may be invalid there. Luckily, failure of the Constraint Qualification is rarely a problem in practice. In the rare cases where it arises, it can usually be circumvented by writing the algebraic form of the constraint differently, as with the budget constraint above. But students of the subject should not forget the problem altogether; it is a favorite source of trick questions in examinations. If something strange seems to be going wrong when you try the standard methods, you should check to see if the problem violates the Constraint Qualification. But I shall omit further mention of this complication except in the formal statements and mathematical proofs.

The Tangency Argument

The second digression relates the arbitrage argument to the tangency condition more familiar from elementary economics texts. In Chapter 1 we saw an alternative way to obtain the condition (1.6), based on the tangency of the budget line and an indifference curve. The same can be done for (2.5). Figure 2.1 shows the story. Along the curve $G(x) = c$, we have $dG(x) = 0$, and from (2.3) we can calculate the slope of the tangent to the curve at x as

$$dx_2/dx_1 = -G_1(x)/G_2(x). \tag{2.7}$$

Note the reversal of the subscripts, exactly as in Chapter 1. Note also that if $G_2(x) = 0$ the curve is vertical; this is not a serious problem. If both $G_1(x)$ and $G_2(x)$ are zero, the slope is not well defined, and the method may run into a problem as was explained just above. In most economic applications, G is an increasing function of both arguments. Then $G_1(x)$ and $G_2(x)$ are both positive, dx_2/dx_1 along the curve is negative, and we have the usual downward-sloping transformation frontier for the constraint.

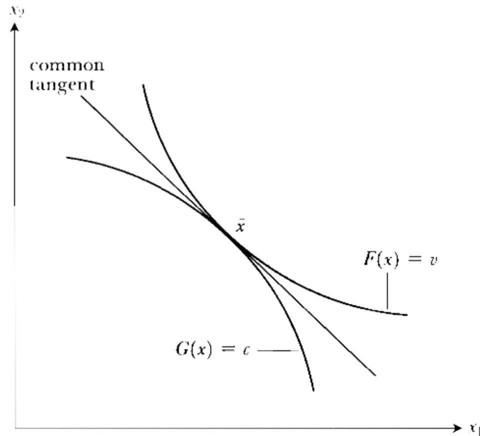

Fig. 2.1 – The tangency solution

A contour of the objective function F, that is, a curve of equal values $F(x)$, runs through the point x. The slope of the tangent to the contour at this point is similarly calculated:

$$dx_2/dx_1 = -F_1(x)/F_2(x). \tag{2.8}$$

Once again, in most economic applications, F is an increasing function and the contour (indifference curve) is downward-sloping.

If $\bar{x}$ is optimum, the two curves must be mutually tangential, that is, have the same slope, at this point. Equating the two expressions, we have

$$F_1(\bar{x})/F_2(\bar{x}) = G_1(\bar{x})/G_2(\bar{x}), \tag{2.9}$$

which is equivalent to (2.5).

Necessary vs. Sufficient Conditions

Recall what the above argument established: *if* $\bar{x}$ maximizes $F(x)$ subject to $G(x) = c$, *then* (2.5) holds. In other words, the condition

(2.5) is a logical consequence of the optimality of $\bar{x}$. Therefore it is called a necessary condition for optimality. To be more precise, since it involves the first-order derivatives of the functions F and G, it is called the ***first-order necessary condition***.

In searching for an optimum, the first-order necessary condition helps to narrow down the search. The necessary conditions were established starting with an assumed known optimum $\bar{x}$. But we turn the story around by treating the components $\bar{x}_1$ and $\bar{x}_2$ as unknowns, and the constraint (2.1) and the necessary condition (2.5) as two equations that will determine them. Typically, there is a whole continuous range of values of $\bar{x}$ satisfying the constraint (2.1). But there are only a few values of $\bar{x}$, and if we are lucky, just one, that also satisfies the condition (2.5).

If we know from separate reasoning that our problem indeed has a solution, and we find that there is a unique $\bar{x}$ satisfying the constraint and the first-order necessary condition, then it must be the solution we seek. If there are multiple solutions to (2.1) and (2.5) taken together, then all are candidates for optimality as far as the present analysis is concerned, and some other method must be used to find the correct solution. Even then, the first-order necessary condition (2.5) will have cut down quite drastically the number of candidate points we need to examine.

The main reason that the first-order necessary condition does not always lead us to the right solution is that the same first-order condition is also necessary for the problem of *minimizing* the same function $F(x)$ subject to the same constraint $G(x) = c$. Minimizing $F(x)$ is the same as maximizing $-F(x)$. By the same reasoning as that leading to (2.8), we can find the slope of a contour of this function:

$$dx_2/dx_1 = -[-F_1(x)]/[-F_2(x)] = -F_1(x)/F_2(x).$$

Equating the value of this at $\bar{x}$ to (2.7), we get (2.9) again.

There is also the point that to obtain the condition, we asked if the value $F(x)$ could be improved by making *small* changes in x. If not, then x is better than the comparison points in a small neighborhood, or it yields a local peak of $F(x)$. Now a function can in general have several such local peaks, and several local troughs too. The same first-order necessary condition will be true at all these points. Only one will give a true or global maximum.

Finally, either a maximum or a minimum implies (2.5) but not vice versa. Therefore the condition might be satisfied at a point that is neither a maximum nor a minimum, even in a small neighborhood. As a simple example, consider $F(x) = x^3$ where x is a scalar. We have $F'(0) = 0$, but $x = 0$ gives neither a maximum nor a minimum of $F(x)$.

To distinguish such cases, any point satisfying the first-order necessary conditions is called a ***stationary point***. The true optimum is one of the stationary points. To locate it among these candidates, we need some other test. Such tests typically rely on the curvature, or the second-order derivatives, of the functions. In Figure 2.1 the curvatures of the contours of F and G have been chosen correctly for a maximum. Thus the curve $G(x) = c$ gets flatter as x_1 decreases and x_2 increases along it; the economic interpretation is that the marginal rate of transformation of x_1 into x_2 diminishes as more and more of such a transformation is carried out. Similarly, the contour of F shows a diminishing marginal rate of substitution.

Tests involving curvatures or second-order derivatives are the subject of Chapters 6–8. These tests differ from the first-order conditions of this chapter in another way: *if* such a condition holds, *then* the point in question is a maximum, at least in comparison with neighboring points; the condition ensures optimality. Therefore such a condition is called a ***second-order sufficient*** condition.

Lagrange's Method

Now let us express the first-order necessary condition (2.6) in a way that is easy to remember and use. This is called Lagrange's Method after its inventor. Note that we want to use the condition to solve for the optimum $\bar{x}$. We introduced λ as the common value of the two sides in (2.5), so it is just as much unknown as the optimum $\bar{x}$. That is, we have to determine it as an integral part of the solution. In the meantime, call it an undetermined Lagrange multiplier. Define a new function, called the Lagrangian,

$$L(x, \lambda) = F(x) + \lambda\,[\,c - G(x)\,]. \tag{2.10}$$

Note that L is also a function of c, and of any other parameters that appear in the functional forms of F and G. Such arguments of L will be shown explicitly only when they are important in the context.

Denote the partial derivatives of L by

$$L_j \equiv \partial L/\partial x_j, \qquad L_\lambda \equiv \partial L/\partial \lambda.$$

Then

$$L_j(x, \lambda) = F_j(x) - \lambda\, G_j(x),$$

and

$$L_\lambda(x, \lambda) = c - G(x).$$

The first-order necessary condition (2.6) becomes just $L_j = 0$ for $j = 1$ and 2, and the constraint (2.1) simply $L_\lambda = 0$. Then we can state the result of the whole argument so far into a simple statement:

Lagrange's Theorem: Suppose x is a two-dimensional vector, c is a scalar, and F and G functions taking scalar values. Define the function L as in (2.10). If $\bar{x}$ maximizes $F(x)$ subject to $G(x) = c$, with no other constraints (such as non-negativity), and if $G_j(\bar{x}) \neq 0$ for at least one j, then there is a value of λ such that

$$L_j(\bar{x}, \lambda) = 0 \ \text{ for } \ j = 1, 2 \qquad L_\lambda(\bar{x}, \lambda) = 0. \tag{2.11}$$

Remember that the theorem provides necessary conditions for optimality. In other words, it starts with a known optimum $\bar{x}$, and establishes that it must satisfy (2.11). But in practice, much of the use of the theorem is in helping us narrow down the search for an initially unknown optimum. We regard (2.11) as three equations for the three unknowns $\bar{x}_1$, $\bar{x}_2$, and λ. The equations are generally non-linear and neither existence nor uniqueness of the solution is guaranteed. If the conditions have no solution, the reason may be either that the maximization problem itself has no solution, or that the Constraint Qualification fails and the first-order conditions are inapplicable. If the conditions have multiple solutions, we need the second-order conditions to arbitrate between the candidate solutions. But in most of our applications, the problems will be well posed enough that the first-order necessary conditions take us to the unique solution. I shall now develop some examples that use Lagrange's method, and offer some exercises for you to attempt similar solutions. After you have gained some experience of problems with two variables and one constraint, you will be ready for the extensions considered in the next chapter.

While the notation $\bar{x}$ keeps the theoretical developments clear by distinguishing the general point from the particular optimum, it becomes cumbersome in applications where we are searching for an unknown optimum. Therefore we often drop the bar on x in conditions like (2.11) when using them in particular contexts, such as the examples below.

Examples

Example 2.1: Preferences that Imply Constant Budget Shares

Consider a consumer choosing between two goods x and y, with prices p and q respectively. (The notation x_1, x_2 etc. was used in the theoretical part because it generalizes more easily to several goods and constraints, but the x, y notation is simpler in examples with just two goods.) His income is I, so the budget constraint is

$$p\,x + q\,y = I.$$

Suppose the utility function is

$$U(x, y) = \alpha\ \ln(x) + \beta\ \ln(y), \tag{2.12}$$

where α, β are positive constants and ln denotes natural logarithms.

Write the Lagrangian

$$L(x, y, \lambda) = \alpha\ \ln(x) + \beta\ \ln(y) + \lambda\,[I - p\,x - q\,y].$$

Recall that $d\ln(x)/dx = 1/x$. Therefore the first-order necessary conditions (2.11) become

$$\partial L/\partial x \equiv \alpha/x - \lambda\, p = 0,$$

$$\partial L/\partial y \equiv \beta/y - \lambda\, q = 0,$$

and

$$\partial L/\partial \lambda \equiv I - p\,x - q\,y = 0.$$

To solve these, substitute for x, y from the first two into the third. This gives

$$\lambda = (\alpha + \beta)/I, \tag{2.13}$$

and then

$$x = \frac{\alpha\, I}{(\alpha+\beta)\, p}, \qquad y = \frac{\beta\, I}{(\alpha+\beta)\, q}. \tag{2.14}$$

These are the demand functions, namely the solutions for the optimum quantities in terms of the prices, income and the given parameters α, β.

We can write them alternatively as

$$\frac{p\,x}{I} = \frac{\alpha}{\alpha+\beta}, \qquad \frac{q\,y}{I} = \frac{\beta}{\alpha+\beta}. \tag{2.15}$$

In other words, for the utility function specified, the shares of income spent on the two goods are constants. This is a convenient property, and one that is sometimes close enough to reality. In the initial exploration of theoretical models, this specification is often the crucial simplification that yields concrete results that suggest the directions for further analysis and testing. Therefore this function is a favorite of economists.

Note that in (2.13) the marginal utility of income is inversely proportional to the income. This might seem a natural consequence of the intuitively appealing idea of diminishing marginal utility. But that is a treacherous concept; see Exercise 2.1 below.

Example 2.2: Guns vs. Butter

Consider an economy with 100 units of labor. It can produce guns x or butter y. To produce x guns, it takes x^2 units of labor; likewise y^2 units of labor are needed to produce y guns. Therefore the economy's resource constraint is

$$x^2 + y^2 = 100.$$

Geometrically, you can easily see that the production possibility frontier is a quarter-circle.

The objective function to be maximized is

$$F(x,y) = a\,x + b\,y.$$

where a, b are given positive constants.

To solve this problem, form the Lagrangian

$$L(x,y,\lambda) = a\,x + b\,y + \lambda\,[\,100 - x^2 - y^2\,].$$

The first-order conditions are

$$\partial L/\partial x \equiv a - 2\,\lambda\,x = 0,$$

$$\partial L/\partial y \equiv b - 2\,\lambda\,y = 0,$$

and

$$\partial L/\partial \lambda \equiv 100 - x^2 - y^2 = 0.$$

Substitute from the first two into the third to get

$$100 = (a^2 + b^2)/(4\,\lambda^2),$$

or

$$\lambda = (a^2 + b^2)^{1/2}/20.$$

Then

$$x = 10\;a/(a^2 + b^2)^{1/2}, \qquad y = 10\;b/(a^2 + b^2)^{1/2}. \tag{2.16}$$

You can think of a, b as the weights or social values attached to the two goods, and then (2.16) gives the economy's optimal supplies as functions of these weights. If both weights are increased in equal proportions, say doubled, then the optimum quantities x and y are unchanged. The supplies are homogeneous of degree zero in the values, so only the *relative* values matter. The supply of each good increases as its relative value increases. In later work, especially the chapter on comparative statics, we shall see how generally valid such properties are.

Exercises

Exercise 2.1: The Cobb-Douglas Utility Function

Consider the consumer's problem as in Example 1, but with a different utility function $\widetilde{U}$ defined by

$$\widetilde{U} = x^{\alpha}\; y^{\beta}.$$

Show that it yields the same constant-budget-share demand functions (2.14) as above. (Hint to simplify the solution process: eliminate the Lagrange multiplier between the first-order conditions for the two goods. This gives a relation between x and y. Simplify this as far as possible, and then use it and the budget constraint to solve for the quantities.)

Note that the two utility functions are linked

$$U(x,y) = \ln[\widetilde{U}(x,y)], \qquad \text{or} \qquad \widetilde{U}(x,y) = \exp[U(x,y)].$$

This illustrates that changing the utility function by any increasing transformation does not affect the consumer's optimum choice. If observed demand behavior is all that matters, then the form of the utility function is indeterminate (and irrelevant) to within such a transformation. Any properties that depend on the choice of a particular form are meaningless.

One such is diminishing marginal utility of income. If we write the multiplier for this problem as $\widetilde{\lambda}$ to distinguish it from the λ of Example 2.1, then you should verify that

$$\widetilde{\lambda} = \frac{\alpha^{\alpha}\,\beta^{\beta}}{(\alpha+\beta)^{\alpha+\beta}}\,\frac{I^{\alpha+\beta-1}}{p^{\alpha}\,q^{\beta}}. \tag{2.17}$$

If $(\alpha+\beta) > 1$, then $\widetilde{\lambda}$ increases with income.

In some circumstances, specific forms of the utility function play special roles. This happens when some assumptions about interpersonal comparability are made, or when functions that are additively separable across time-periods or states of the world are used for representing preferences in situations involving time or uncertainty. But in all these cases, the primacy of the special functional forms arises from those other considerations, not from the underlying mechanism of individual choice.

When these other considerations are absent, we are free to transform the utility function for computational convenience. Note that changing both α and β in the same proportions leaves demand unaffected in Example 2.1 as well as in Exercise 2.1. A glance at equations (2.14) or (2.17) shows that it is convenient to choose these proportions so that $\alpha+\beta = 1$.

Exercise 2.2: The Linear Expenditure System

Return once again to the consumer of Example 2.1, but let the utility function be modified to $\widehat{U}$, where

$$\widehat{U}(x,y) = \alpha\,\ln(x - x_0) + \beta\,\ln(y - y_0),$$

where x_0 and y_0 are given constants, and $\alpha+\beta = 1$. Show that the optimal expenditures on the two goods are linear functions of income and prices:

$$p\,x = \alpha\,I + \beta\,p\,x_0 - \alpha\,q\,y_0,$$
$$q\,y = \beta\,I - \beta\,p\,x_0 + \alpha\,q\,y_0.$$

This slight modification of the utility function brings with it a much richer range of possible optimum choice. The budget shares of the two goods can now vary systematically with income and prices. One good can be a necessity and the other a luxury (but neither good can be inferior since α and β must be positive to keep the marginal utilities positive). But the expenditures still have a simple functional form. For these reasons, this specification was popular in the early empirical work on consumer demand.

Exercise 2.3: Production and Cost-Minimization

Consider a producer who rents machines K at r per year and hires labor L at wage w per year to produce output Q, where

$$Q = \sqrt{K} + \sqrt{L}.$$

Suppose he wishes to produce a fixed quantity Q at minimum cost. Find his factor demand functions. Show that the Lagrange multiplier is given by

$$\lambda = 2\,w\,r\,Q/(w+r).$$

Suggest an economic interpretation for λ.

Now let p denote the price of output. Suppose the producer can vary the quantity of output, and seeks to maximize profit. Show that his optimum output supply is

$$Q = p\,(w+r)/(2\,w\,r).$$

Relate this to your interpretation of λ.

Further Reading

For supplementary treatments of Lagrange's method, see Varian (op. cit.), appendices to chs. 5 and 20, and Smith (op. cit.) ch. 2 (sects. 1–4), ch. 4 (sects. 1–3).

The development of the theory in this book is relatively intuitive and heuristic. There are several textbooks that are mathematically more rigorous; I mention just one:

MICHAEL D. INTRILIGATOR, ***Mathematical Optimization and Economic Theory***, Englewood Cliffs, NJ: Prentice-Hall, 1971; ch. 3 is about Lagrange's method.

3 Extensions and Generalizations

More Variables and Constraints

If there are n choice variables $(x_1, x_2, \dots x_n)$, we simply let the vector x have n components. Then (2.11) is extended to $j = 1, 2, \dots n$, and we have $(n+1)$ equations in the $(n+1)$ unknowns, namely the n components of x and the number λ.

If there are m constraints, write them as

$$G^i(x) = c_i, \qquad i = 1, 2, \dots m,$$

where the functions are identified by superscripts to avoid confusion with partial derivatives, which are being denoted by subscripts. For the moment, continue to ignore any other restrictions such as non-negativity on the variables. We need $m < n$, for n constraints on n variables will generally reduce the choice to a discrete set of points, while more constraints will, in general, be mutually inconsistent.

Lagrange's method extends to this situation very easily. We define a multiplier λ_i for each constraint, and define the Lagrangian

$$L(x_1, \dots x_n, \lambda_1, \dots \lambda_m) = F(x_1, \dots x_n) + \sum_{i=1}^{m} \lambda_i [c_i - G^i(x_1, \dots x_n)]. \tag{3.1}$$

The first-order necessary conditions satisfied at the optimum $\bar{x}$ are then

$$\partial L / \partial x_j = 0, \qquad j = 1, 2, \dots n, \tag{3.2}$$

and

$$\partial L / \partial \lambda_i = 0 \qquad i = 1, 2, \dots m. \tag{3.3}$$

When using these to search for the optimum, we treat them as $(m+n)$ equations in the $(m+n)$ unknowns $\bar{x}_1, \bar{x}_2, \dots \bar{x}_n, \lambda_1, \lambda_2, \dots \lambda_m$.

These can be put more neatly in vector-matrix form. Let c be a column vector with components c_i, and G a column vector-valued function with component functions G^i. Then all the constraints can be written together as a vector equation $G(x) = c$. Next, the partial derivatives $F_j(x)$ should be formed into a vector which I shall write as $F_x(x)$, the subscript x indicating the vector argument with respect to which the derivatives have been taken. I shall make the convention that when the argument of a function is a column vector, the vector of partial derivatives is a row vector, and vice versa. There is a good mathematical reason for this, but the main advantage here is that it will save us from having to transpose matrices all the time. Each G^i will have a row vector of partial derivatives $G^i_x(x)$ with components $G^i_j(x)$, and these row vectors will be stacked vertically to form an m-by-n matrix, written $G_x(x)$. The multipliers will form a row vector λ.

With this notation, (3.1) can be written more simply as

$$L(x, \lambda) = F(x) + \lambda [c - G(x)]. \tag{3.4}$$

To verify this, note that λ is an m-dimensional row vector, that is, a 1-by-m matrix, while the term in the square brackets is an m-dimensional column vector, that is, an m-by-1 matrix. The product of these two matrices is found by multiplying the ith element of the row vector by the corresponding element of the column vector, and adding all these products. The result is a 1-by-1 matrix, that is, a scalar. It is exactly the expression in (3.1).

In the same way, the conditions (3.2) and (3.3) become more compact:

$$L_x(\bar{x}, \lambda) = 0, \tag{3.5}$$

$$L_\lambda(\bar{x}, \lambda) = 0. \tag{3.6}$$

What happens to the Constraint Qualification? Note that in the case of Chapter 2, we had two variables $(n = 2)$ and one constraint $(m = 1)$. The matrix $G_x(x)$ was 1-by-2, that is, simply the row vector $(G_1(x), G_2(x))$. The Constraint Qualification was the assumption that this vector was not zero at $\bar{x}$. The condition for the general case is that the matrix should not have any singularity.

Since $m < n$, this amounts to requiring that it should have m independent rows, that is, its rank should be the maximum possible, namely m.

Let us sum up these results into a compact statement:

Lagrange's Theorem: Suppose x is an n-dimensional vector, c an m-dimensional vector, F a function taking scalar values, G a function taking m-dimensional vector values, with $m < n$. Define

$$L(x, \lambda) = F(x) + \lambda\,[\,c - G(x)\,], \tag{3.4}$$

where λ is an m-dimensional row vector. If x maximizes $F(x)$ subject to $G(x) = c$ and no other restrictions, and rank $G_x(\bar{x}) = m$, then there is a value of λ such that

$$L_x(\bar{x}, \lambda) = 0, \tag{3.5}$$

$$L_\lambda(\bar{x}, \lambda) = 0. \tag{3.6}$$

I have stated all these generalizations without proof. Most readers will probably accept the intuition of the case of two variables and one constraint, and merely test the general result in applications. Even they will find the precise statements like the ones above useful. For the more mathematically orientated readers, I give a formal proof of the most general result of this chapter, the Kuhn–Tucker Theorem, in the Appendix.

Non-Negative Variables

Next suppose that the variables x_j must be non-negative to make economic sense. If the optimum $\bar{x}$ happens to be such that these requirements are not binding, that is, all the x_j are in fact strictly positive, then the above conditions (3.2) and (3.3) continue to apply. If say x_1 is zero, then the arbitrage argument that leads to first-order conditions is more limited. We can consider only those infinitesimal changes dx for which $dx_1 > 0$. Generalization of the reasoning in Chapter 1 that led to the inequality condition (1.10) now gives us

$$L_1(\bar{x}) \equiv F_1(\bar{x}) - \sum_{i=1}^{m} \lambda_i\, G^i_1(\bar{x}) \leq 0.$$

Once again I omit the formal proof.

More generally, some components of $\bar{x}$ may be positive and others zero. Then an equation like (3.2) should hold for the partial derivative of the Lagrangian with respect to every component that is positive, and an inequality like the one just above with respect to every component that is zero, at the initial point. In other words, for every j, we should have

$$L_j(x) \leq 0, \qquad \bar{x}_j \geq 0, \tag{3.7}$$

with at least one of these holding as an equation. In exceptional cases, both might hold as equations. But the logical possibility that both inequalities may be strict is ruled out by the requirements of optimality.

The requirement that at least one inequality in (3.7) should hold as an equation is sometimes stated more compactly as

$$\bar{x}_j\, L_j(\bar{x}) = 0.$$

The point is that the product can be zero only if at least one of the factors is zero.

A pair of matched inequalities like (3.7), not both of which can be strict, is said to show ***complementary slackness***. A single inequality, say $\bar{x}_j \geq 0$, is ***binding*** if it holds as an equation, that is, if x_j is at the extreme limit of its permitted range; the inequality is said to be ***slack*** if $\bar{x}_j$ is positive, and so has some room to maneuver before hitting an extreme. Each one of the pair of inequalities in (3.7) therefore complements the slackness in the other: if one is slack, the other must be binding.

We can collect all the component inequalities in (3.7) into vectors. Here I shall use the following notation: $x \geq 0$ means that $x_j \geq 0$ for every j; $x > 0$ means that at least one of these component inequalities is strict; $x \gg 0$ means that all the component inequalities are strict. Then (3.7) becomes

$$L_x(\bar{x}) \leq 0, \quad \bar{x} \geq 0, \qquad \text{with complementary slackness}, \tag{3.7}$$

it being understood that complementary slackness holds for each component pair in the vector inequalities.

Once again, I summarize the result for future reference:

Lagrange's Theorem with Non-Negative Variables: Suppose x is an n-dimensional vector, c an m-dimensional vector, F a function taking scalar values, G a function taking m-dimensional vector values, and $m < n$. Define

$$L(x, \lambda) = F(x) + \lambda [c - G(x)], \tag{3.4}$$

where λ is an m-dimensional row vector. If $\bar{x}$ maximizes $F(x)$ subject to $G(x) = c$ and $x \geq 0$, and the constraint qualification $\operatorname{rank} G_x(\bar{x}) = m$ holds, then there is a value of λ such that

$$L_x(\bar{x}, \lambda) \leq 0, \quad \bar{x} \geq 0, \quad \text{with complementary slackness,} \tag{3.7}$$

and

$$L_\lambda(\bar{x}, \lambda) = 0. \tag{3.6}$$

First-order necessary conditions are supposed to narrow down our search for a solution. How does that work in this instance? We start without knowledge of which components of the solution are going to be positive and which ones zero. If we assume a particular pattern, say $\bar{x}_1 > 0$, $\bar{x}_2 = 0$, $\bar{x}_3 > 0$, ..., then from (3.7) we get $L_1(\bar{x}) = 0$, $\bar{x}_2 = 0$, $L_3(\bar{x}) = 0$, ... In other words, an assumed pattern leads to a set of n equations from (3.7). These may not have a solution at all, and even if they do, it may not satisfy the other inequality conditions required from the pattern. But if a solution satisfying the requirements exists, then it becomes a candidate for the optimum choice.

There are 2^n patterns of positive and zero components in the n-dimensional vector x. We can then repeat the same exercise for every one of these patterns and find other candidates for optimality. Then our search narrows down to all these candidates.

In some applications the search is relatively easy to perform; the simplex method for solving linear programming problems is essentially a systematic algorithm for such a search. But in general the search is too exhaustive and exhausting. If we had to carry it out every time, the prospects for solving constrained optimization problems would be very poor. Luckily many problems of practical interest offer short cuts or systems for searching among the patterns. In most basic contexts of economic theory, we can make good guesses about the likely pattern of equations and inequalities,

proceed on that basis, and use second-order sufficient conditions to verify that the resulting solution is indeed the optimum.

Inequality Constraints

Now we can consider more general inequality constraints. This is of considerable economic importance, because usually there is no compulsion to use all of available income or some resource, and we should determine in the process of solution whether it is optimal to use it fully.

Suppose the first component in the constraint need only hold as an inequality

$$G^1(x) \leq c_1.$$

We can fit this into the context discussed before using the same trick as the introduction of the 'unspent income' variable in the consumer's problem of Chapter 1. Let us define a new variable x_{n+1} as

$$x_{n+1} = c_1 - G^1(x). \tag{3.8}$$

In terms of the enlarged set of variables, the constraint has become an exact equation. The new variable x_{n+1} is restricted to be non-negative, but we know how to handle that.

Write $\widehat{L}$ as the Lagrangian for the new problem, to distinguish it from the L of the old one. Then

$$\begin{aligned}
\widehat{L}(x_1, \dots\, x_n, x_{n+1}, \lambda_1, \dots\, \lambda_m) \\
&= F(x_1, \dots\, x_n, \lambda_1, \dots\, \lambda_m) \\
&\qquad + \lambda_1 \left[c_1 - G^1(x_1, \dots x_n) - x_{n+1} \right] \\
&\qquad + \sum_{i=2}^{n} \lambda_i \left[c_i - G^i(x_1, \dots x_n) \right] \\
&= L(x_1, \dots x_n, \lambda_1, \dots \lambda_m) - \lambda_1\, x_{n+1}.
\end{aligned}$$

All the other first-order conditions are as before, but we have a new one with respect to x_{n+1}:

$$\partial \widehat{L} / \partial x_{n+1} \equiv -\lambda_1 \leq 0, \qquad x_{n+1} \geq 0,$$

with complementary slackness. Recall that

$$x_{n+1} = c_1 - G^1(x) = \partial L / \partial \lambda_1.$$

Therefore the condition can be written

$$\partial L/\partial\lambda_1 \geq 0, \qquad \lambda_1 \geq 0, \qquad (3.9$$

with complementary slackness: a form that is nicely symmetri
with the condition (3.6) for non-negative variables.

We can similarly allow all constraints to be inequalities. If w do that, there is no reason in general to maintain the restrictio $m < n$; any number of inequality constraints can still leave a non trivial range of variation for the choice variables x. The first-ord conditions are the exact analogs of (3.9) for the other component and we can stack them into a pair of vector inequalities with com plementary slackness in each component pair.

The Constraint Qualification is altered. Specifically, suppos k of the constraints are binding, that is, hold as equalities. Tak the rows corresponding to these k from the matrix of derivativ $G_x(x)$. The resulting k-by-n submatrix should have rank k.

Once again we have the formal statement:

Lagrange's Theorem with Inequality Constraints: Suppose x i an n-dimensional vector, c an m-dimensional vector, F a functio taking scalar values, and G a function taking m-dimensional vect values. Define

$$L(x,\lambda) = F(x) + \lambda\,[\,c - G(x)\,], \qquad (3.4$$

where λ is an m-dimensional row vector. If $\bar{x}$ maximizes $F($ subject to $G(x) \leq c$, and the constraint qualification rank hold namely that the submatrix of $G_x(x)$ formed by taking those row i for which $G^i(\bar{x}) = c_i$ has the maximum possible rank, then ther is a value of λ such that

$$L_x(x,\lambda) = 0, \qquad (3.5$$

and

$$L_\lambda(x,\lambda) \geq 0, \lambda \geq 0 \quad \text{with complementary slackness.} \qquad (3.10$$

Finally, we can combine the cases of non-negative variabl and inequality constraints to get the most general result of thi kind:

Kuhn Tucker Theorem: Suppose x is an n-dimensional vector, c an m dimensional vector, F a function taking scalar values, G a function taking m-dimensional vector values. Define

$$L(x,\lambda) = F(x) + \lambda\,[\,c - G(x)\,], \qquad (3.4)$$

where λ is an m-dimensional row vector. Suppose $\bar{x}$ maximizes $F(x)$ subject to $G(x) \leq c$ and $x \geq 0$, and the constraint qualification holds, namely the submatrix of $G_x(\bar{x})$ formed by taking those rows i for which $G^i(\bar{x}) = c_i$ has the maximum possible rank. Then there is a value of λ such that

$$L_x(\bar{x},\lambda) \leq 0, \;\; \bar{x} \geq 0, \quad \text{with complementary slackness,} \qquad (3.7)$$

and

$$L_\lambda(\bar{x},\lambda) \geq 0, \;\; \lambda \geq 0, \quad \text{with complementary slackness.} \qquad (3.10)$$

Once again, the exhaustive procedure for finding a solution using this theorem involves searching among all $2^{(m+n)}$ patterns that are possible from the $(m+n)$ complementary slackness conditions. Once again, short cuts are usually available. We shall soon see some examples that illustrate and apply the theorem.

Examples

Example 3.1: Quasi-Linear Preferences

Suppose there are two goods x and y, whose quantities must be non negative, and whose prices are p and q respectively, both being positive. Consider a consumer with income I and the utility function

$$U(x,y) = y + a\;\ln(x),$$

where a is a given positive constant. Such preferences are called quasi-linear, because the utility function can be chosen linear in the quantity of one of the goods.

To find this consumer's demand functions, we can use the Kuhn–Tucker Theorem. Form the Lagrangian

$$L(x,y,\lambda) = y + a\;\ln(x) + \lambda\,[\,I - p\,x - q\,y\,].$$

The first-order conditions (3.7) and (3.10) become

$$a/x - \lambda p \le 0, \qquad x \ge 0, \tag{3.11}$$

$$1 - \lambda q \le 0, \qquad y \ge 0, \tag{3.12}$$

$$I - p x - q y \ge 0, \qquad \lambda \ge 0, \tag{3.13}$$

in each case with complementary slackness.

Let us solve this mechanically to develop an intuition for the technique. With two non-negative variables and one inequality constraint, there are $2^3 = 8$ possible patterns of equations and inequalities. Let us see which ones offer candidates for optimality.

First note that the budget constraint cannot be slack. The economics of this is simple: the budget cannot go unspent because the goods have positive marginal utilities. Formally, if the budget constraint were slack, (3.13) would give $\lambda = 0$. Then (3.11) would require $a/x \le 0$ which is impossible; similarly (3.12) would require $1 \le 0$.

This reduces the number of cases we need look at to four, namely the patterns of positive and zero x and y that can arise in (3.11) and (3.12).

Both x and y being zero does not satisfy the budget equation, which must hold, as we saw above. Therefore this case is logically impossible. The economics of the matter is again that the goods have positive marginal utilities.

If $x = 0$ and $y = I/q > 0$, then (3.12) gives $\lambda = 1/q$, and then (3.11) becomes $p/q \ge \infty$, which is not true. Therefore this case cannot arise either. The economic reason is that the first small unit of x has infinite marginal utility, so it cannot be optimal to consume zero x.

Next consider $x > 0$ and $y = 0$. Then from the budget constraint $x = I/p$, and from (3.11), $\lambda = a/I$. Using this in (3.12), we have $1 \le a\,q/I$, or $I \le a\,q$. This is a condition on the given parameters of the problem, and they may or may not satisfy it. If they do, the premises of the case are mutually consistent and we have a candidate for optimality.

Finally, if both x and y are positive, (3.11) and (3.12) give $a/(p\,x) - \lambda = 1/q$, so $x = a\,q/p$. Then the budget constraint gives $y = I/q - a$. This is logically consistent if $I > a\,q$. Once again this may or may not be satisfied; if it is, we have a candidate for optimality.

This completes the discussion of the cases. Now let us infer some useful ideas from the procedure and the results. The first thing to note is that six of the eight possible patterns could be ruled out using economic sense; after some experience it is possible to cut down the amount of formal reasoning quite a lot in this way.

Next, observe that the space of all possible values of p, q, and I is split into two exhaustive and mutually exclusive sets by the conditions on the parameters that make each of the last two cases internally consistent. One requires $I \le a\,q$, while the other requires $I > a\,q$. In many applications, a similar neat classification will emerge.

Third, when $I = a\,q$, we have $\lambda = 1/q$. Therefore

$$1 - \lambda q = 0 \qquad \text{and} \qquad y = 0;$$

both inequalities in (3.12) hold as exact equations. When I discussed complementary slackness following (3.6), I mentioned this possibility, and called it an exceptional case. Now we see why: it arises at the special configuration of parameters where the solution is just at the point of switching from one pattern of equations and inequalities in the complementary slackness conditions to another pattern.

Finally, let us restate the optimum choice rule:

$$\text{If} \quad I \le a\,q, \quad x = I/p \quad \text{and} \quad y = 0,$$

$$\text{if} \quad I > a\,q, \quad x = a\,q/p \quad \text{and} \quad y = I/q - a.$$

To see the solution more clearly, carry out the thought experiment of starting at a very low level of income and raising it gradually. At first, all of income is spent on good x and none on y. After a point, the expenditure on x is kept constant, and all additional income is spent on y. We can think of good x as an exemplar of a necessity: it has an absolute first claim on income, but once its needs are satisfied, all extra income can go toward other goods.

Quasi-linear preferences are useful when we want to isolate one sector or industry and wish to avoid the feedback of income effects on the demand for its goods. This is often called 'partial equilibrium'; a better name would be 'industry analysis'. The assumption that changes in income do not affect the demand for the good in question is obviously not meant to be taken literally, but often proves an acceptable approximation or simplification for the purpose.

Example 3.2: Technological Unemployment

Suppose an economy has 300 units of labor and 450 units of land. These can be used in the production of wheat and beef. Each unit of wheat requires 2 of labor and 1 of land; each unit of beef requires 1 of labor and 2 of land.

A plan to produce x units of wheat and y units of beef is feasible if its requirements of each factor of production are less than the available amount of that factor:

$$2\,x + y \leq 300, \tag{3.14}$$

$$x + 2\,y \leq 450. \tag{3.15}$$

Each of the inequalities allows all points on or below a straight line. The set that is feasible given both constraints is the quadrilateral OABC of Figure 3.1. The north-east frontier of the feasible set, or the production possibility frontier, is ABC.

Along AB, (3.14) holds as an equation, while along BC, (3.15) does. Only at B do both hold as equations. Everywhere else, there is unemployment of one factor or the other.

You might be tempted to assume that it will be optimum to have full employment, and achieve production at B with $x = 50$ and $y = 200$. However, that is not necessarily so.

Suppose the society has an objective or social welfare function defined over the quantities of the two goods of the simple form we have used before:

$$W(x,y) = \alpha\ \ln x + \beta\ \ln y, \tag{3.16}$$

where α and β are given positive constants, and $(\alpha + \beta) = 1$.

We know from the intuition developed in the previous example that non-negativity constraints on x and y are not going to be binding. So let us leave them out from the start. Write the Lagrange multiplier for the constraint (3.14) as λ and that for (3.15) as μ. Form the Lagrangian

$$L(x,y,\lambda,\mu) = \alpha\ \ln x + \beta\ \ln y + \lambda\,[300 - 2\,x - y] + \mu\,[450 - x - 2\,y]$$

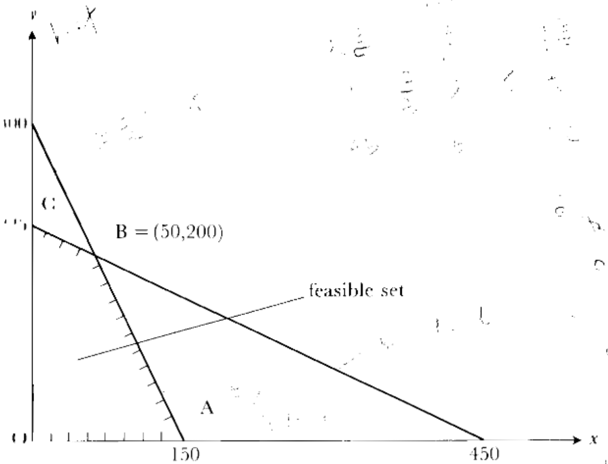

Fig. 3.1 – Production and unemployment

The first-order conditions are

$$\alpha/x - 2\,\lambda - \mu = 0, \tag{3.17}$$

$$\beta/y - \lambda - 2\,\mu = 0, \tag{3.18}$$

and

$$300 - 2\,x - y \geq 0, \quad \lambda \geq 0 \quad \text{with complementary slackness,} \tag{3.19}$$

$$450 - x - 2\,y \geq 0, \quad \mu \geq 0 \quad \text{with complementary slackness.} \tag{3.20}$$

Between (3.19) and (3.20) we have four possible patterns of equations and inequalities. We should suspect that it is not going to be sensible to keep *both* factors less than fully employed, that is, to have $\lambda = 0$ *and* $\mu = 0$. Let us check this out: using $\lambda = 0 = \mu$ in (3.17) and (3.18) would give $\alpha = 0 = \beta$, and that is not so. Therefore this case is ruled out, and we are left with three.

second line, λ is an m-dimensional row vector, $G_x(\bar{x})$ is an m-by-n matrix, and $d\bar{x}$ is an n-dimensional column vector. The final result is the product of the row vector λ and the column vector dc of equal dimensions m; therefore it is a scalar. In fact it is the inner product of the two vectors:

$$\lambda\, dc = \Sigma_i\, \lambda_i\, dc_i.$$

The result is important enough to be stated separately for reference:

Interpretation of Lagrange Multipliers: If v is the maximum of $F(x)$ subject to a vector of constraints $G(x) = c$, and λ is the row vector of multipliers for the constraints, then change dv that results from an infinitesimal change dc is given by

$$dv = \lambda\, dc. \tag{4.2}$$

It should be stressed that (4.2) gives only the first-order or linear approximation to the change in v if the change in c is more than infinitesimal. For such changes, we can carry the Taylor expansion to higher orders and find a closer approximation. This will be done, although for a different purpose, in Chapter 8.

Shadow Prices

To illustrate and explain (4.2), consider a planned economy for which a production plan $\bar{x}$ is to be chosen to maximize a social welfare function $F(x)$. The vector of the plan's resource requirements is $G(x)$, and the vector of the available amounts of these resources is c. Suppose the problem has been solved, and the vector of the Lagrange multipliers λ is known. Now suppose some power outside the economy puts a small additional amount dc_1 of the first resource (say labor) at its disposal. The optimization problem can be solved afresh with the new labor constraint to determine the new pattern of production. But we know the resultant increase in social welfare without having to do this calculation: it is simply $\lambda_1\, dc_1$. We can then say that the multiplier λ_1 is the marginal product of labor in this economy, measured in units of its social welfare. This is clearly a vital piece of economic information, and that is why Lagrange's method and his multipliers are so important in economics.

If there is only one scarce input, then a paraphrase of the argument of Chapter 1 yields another very instructive way of looking at this result. Suppose we use the additional labor input to raise the quantity of a particular good, say good j, leaving the outputs of all the other goods unchanged. Since we are assuming full employment of labor in both situations, the increase $d\bar{x}_j$ in the output of the chosen good must satisfy

$$G^1_j(\bar{x})\, d\bar{x}_j = dc_1, \qquad \text{or} \qquad d\bar{x}_j = dc_1\,/G^1_j(\bar{x}).$$

The resultant increase in social welfare is

$$F_j(\bar{x})\, dx_j = [\,F_j(\bar{x})/G^1_j(\bar{x})\,]\; dc_1.$$

The condition of optimality (2.5) says that the ratio in the square brackets should be the same for all j. Therefore the effect of the marginal increase in labor supply on social welfare is independent of how the extra labor is used. That is why we can speak unambiguously of the marginal product of labor.

Now suppose the additional labor can only be used at some cost. The maximum the economy is willing to pay in terms of its own social welfare units is clearly λ_1 per marginal unit of c_1. Any smaller payment leaves it with a positive net benefit from using the extra labor; for any larger payment the cost exceeds the benefit. In this natural sense, the Lagrange multiplier is the ***demand price*** the planner places on labor services. A price expressed in units of social welfare may seem strange, but a minor modification brings it into familiar light. Consider some other resource, say land, and number it 2. Now suppose the economy is offered the services of an extra dc_1 of labor, but asked to give in return the services of dc_2 of land. The net gain in social welfare from this transaction is $(\lambda_1\, dc_1 - \lambda_2\, dc_2)$. Therefore the most land the planner is willing to give up is $(\lambda_1/\lambda_2)\, dc_1$. Then it is equally natural to call the ratio (λ_1/λ_2) the demand price of a unit of labor measured in units of land. You know from microeconomic theory that ***relative*** prices rather than ***absolute*** ones govern market exchange; similarly the relative magnitudes of the Lagrange multipliers for different resources govern the planner's willingness to exchange one resource for another.

If a neighboring economy has a different trade-off between the two resources on account of differences in their relative availability

First consider the case, $\lambda = 0$ and $\mu > 0$, which I shall label Case (i). Here (3.17) gives $x = \alpha/\mu$ and (3.18) gives $y = \beta/2\mu$. Since $\mu > 0$, (3.20) then becomes

$$450 = x + 2y = (\alpha + \beta)/\mu, \qquad \text{or} \qquad \mu = 1/450.$$

Then $x = 450\,\alpha$ and $y = 225\,\beta$.

It remains to check out if the feasibility condition in (3.19) is true. We need

$$300 \geq 2x + y = 900\,\alpha + 225\,\beta = 900 - 675\,\beta,$$

or $\beta \geq 8/9$.

The other cases can be checked out in the same way, and I shall merely state the results:

If $\lambda > 0$ and $\mu = 0$ (Case (ii)), we get $x = 150\,\alpha$, $y = 300\,\beta$. The case is internally consistent if $\beta \leq 2/3$.

If $\lambda > 0$ and $\mu > 0$ (Case (iii)), we get the full employment point $x = 50$ and $y = 200$. This case is internally consistent if $2/3 < \beta < 8/9$.

The solution gives several useful insights. Once again, the range of parameters splits nicely into exhaustive and mutually exclusive regions, in each of which just one of the cases yields a candidate for optimality. For low values of β, the solution lies along the line AB. Then there is a middle range where the solution stays at the point B. Finally, for high values of β the solution lies along the line BC.

The social indifference curves for the objective function (3.16) are like hyperbolas. The higher is the weight β attached to y (relative to the weight of x), the more willing is society to sacrifice x for y, that is, the flatter are the hyperbolas. Therefore for low β we get a tangency of a social welfare contour and the production possibility set along the segment AB, for medium values we have a corner solution at B, and for high values a tangency along BC.

Next note that at any point along AB (except B), it is optimal to keep some land unemployed. To see why, note that the goods have fixed coefficients of input requirements, and wheat requires relatively more labor. If we wish to use the unemployed land, we must do so by producing less wheat and more beef. To try it the other way round would increase the labor requirement, but labor is already fully employed. But this is a situation with a relatively low β; beef is not highly valued relative to wheat, and the required sacrifice of wheat is not worth while.

If enough substitution in production were possible, then the difficulty would not arise and both factors could be fully employed. The unemployment in this setting is a consequence of the rigid technology, not of any effective demand failures or coordination failures.

Finally, look again at the complementary slackness conditions (3.19) and (3.20). If one factor is not fully employed at the optimum, then the Lagrange multiplier for its constraint is zero. In Chapter 1, the multiplier on the consumer's budget constraint was the marginal utility of money. In the same way, each multiplier in this problem gives the effect on social welfare of having another marginal unit of that factor. Then complementary slackness becomes economically quite intuitive: if it is optimal not to employ the available amount fully, then an increment must be worthless. In the next chapter I shall develop this idea in more detail.

Exercises

Exercise 3.1: Rationing

Suppose a consumer has the utility function

$$U(x_1, x_2, x_3) = \alpha_1 \ln(x_1) + \alpha_2 \ln(x_2) + \alpha_3 \ln(x_3), \qquad (3.21)$$

where the α_j are positive constants summing to one. The budget constraint is

$$p_1\,x_1 + p_2\,x_2 + p_3\,x_3 \leq I.$$

In addition, the consumer faces a rationing constraint: he is not allowed to buy more than k units of good 1.

Solve the optimization problem. Under what condition on the various parameters is the rationing constraint binding?

Show that when the rationing constraint binds, the income that the consumer would have liked to spend on good 1 but cannot do so is now split between goods 2 and 3 in the proportions $\alpha_2 : \alpha_3$. Would you expect rationing of bread purchases to affect demands for butter and rice in this way? What is the property of the utility function (3.21) that produces the result, and how would you expect the bread–butter–rice case to differ?

Exercise 3.2: Distribution Between Envious Consumers

There is a fixed total Y of goods at the disposal of society. There are two consumers who envy each other. If consumer 1 gets Y_1 and consumer 2 gets Y_2, their utilities are

$$U_1 = Y_1 - k\,Y_2^2, \qquad U_2 = Y_2 - k\,Y_1^2,$$

where k is a positive constant. The allocation must satisfy $Y_1 + Y_2 \le Y$, and maximize $U_1 + U_2$.

Show that if $Y > 1/k$, the resource constraint will be slack at the optimum. Interpret the result.

Exercise 3.3: Investment Allocation

A capital sum C is available for allocation among n investment projects. If the non-negative amount x_j is allocated to project j for $j = 1, 2, \ldots n$, the expected return from this portfolio of projects is

$$\sum_{j=1}^{n} [\,\alpha_j\, x_j - \frac{1}{2}\,\beta_j\; x_j^2\,].$$

The allocation is to be chosen to maximize this.

Find the first-order necessary conditions from the Kuhn–Tucker Theorem. Define

$$H = \sum_{j=1}^{n} (\alpha_j/\beta_j), \qquad K = \sum_{j=1}^{n} (1/\beta_j).$$

Show that

(i) If $C > H$, then a part of the total sum available is left unused.

(ii) If $\alpha_j > (H - C)/K$ for all j, then every project will receive some funding.

(iii) If any project receives zero funding, then it must have a lower α than any project that gets some funding.

Further Reading

A more rigorous treatment of the more general Lagrange and Kuhn–Tucker theorems is in Intriligator (op.cit.), ch. 4.

Quasi-linear preferences and their applications are discussed in more detail by

HAL VARIAN, *Microeconomic Analysis*, New York: Norton, 2nd edn., 1984, pp. 278–83.

A good modern treatment of rationing is in

PETER NEARY and KEVIN ROBERTS, 'The theory of household behavior under rationing', *European Economic Review*, 13 (1980), pp. 25–42. They use methods that will not be developed in this book until Chapter 8, but it will be worth while to return to this reference then.