1 Introduction

This chapter covers two major topics. The first is nonlinear programming, or Kuhn-Tucker conditions. Here we extend the techniques constrained optimization covered in chapter 12 by introducing additional constraints which may, or may not, be binding. The examples and models are chosen to be familiar to the student but with an added twist or dimension to the problem.

The second part of the chapter introduces the envelope theorem and the concept of duality. This section will present the student with an alternative approach to the comparative statics exercises carried out in chapters 11 and 12. Again we engage problems that are familiar to the student and draw comparisons between the traditional methodology versus applying the envelope theorem.

2 Kuhn-Tucker Conditions

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 multiplier will be zero. In nonlinear programming, there exists a similar type of first-order condition, known as the Kuhn-Tucker conditions.\(^1\) As we shall see, however, while the classical first-order condition is always necessary, the Kuhn-Tucker conditions cannot be accorded the 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.

2.1 Effect of Nonnegativity Restrictions

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

\[
\begin{align*}
\text{Maximize} & \quad \pi = f(x_1) \\
\text{Subject to} & \quad x_1 \geq 0
\end{align*}
\]

(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 \(\pi\) occurs in the interior of the shaded feasible region in Fig. 1, such as a point \(A\) in diagram \(a\), then we have an interior solution.

The first-order condition in this case is \( \frac{d\pi}{dx_1} = f'(x_1) = 0 \), same as in the classical problem. Second, as illustrated by point \( B \) in diagram \( b \), 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, a third possibility, a local maximum may in the present context take the position of point \( C \) or point \( D \) in diagram \( c \), because to qualify as a local maximum in the problem (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 (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 diagram \( a \).

The upshot of the above discussion is that, in order for a value of \( x_1 \) to give a local maximum of \( \pi \) in the problem (1), it must satisfy one of the following three conditions

\[
f'(x_1) = 0 \quad \text{and} \quad x_1 > 0 \quad \text{[point A]} \tag{2}
\]

\[
f'(x_1) = 0 \quad \text{and} \quad x_1 = 0 \quad \text{[point B]} \tag{3}
\]

\[
f'(x_1) < 0 \quad \text{and} \quad x_1 = 0 \quad \text{[points C and D]} \tag{4}
\]

Actually, these three conditions can be consolidated into a single
statement
\[ f'(x_1) \leq 0 \quad x_1 \geq 0 \quad \text{and} \quad x_1 f'(x_1) = 0 \] (5)

The first inequality in (5) is a summary of the information regarding \( f'(x_1) \) enumerated in (2) through (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 (5), we have an equation which expresses an important feature common to (2), (3), as well as (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. Taken together, the three parts of (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:
Maximize
\[ \pi = f(x_1, x_2, \ldots, x_n) \] (6)

Subject to
\[ x_j \geq 0 \quad (j = 1, 2, \ldots, n) \]

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 (5) to each choice variable, \( x_j \), taken by itself. Graphically, this amounts to viewing the horizontal axis in Fig. 1 as representing each \( x_j \) in turn. The required modification of the first-order condition then readily suggests itself:

\[ f_j \leq 0 \quad x_j \geq 0 \quad \text{and} \quad x_j f_j = 0 \quad (j = 1, 2, \ldots, n) \] (7)

where \( f_j \) is the partial derivative \( \partial \pi / \partial x_j \).
2.2 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 variables \(n = 3\) and two constraints \(m = 2\):

Maximize
\[
\pi = f(x_1, x_2, x_3)
\]  
(8)

Subject to
\[
g^1(x_1, x_2, x_3) \leq r_1
\]
\[
g^2(x_1, x_2, x_3) \leq r_2
\]
and
\[
x_1, x_2, x_3 \geq 0
\]
which, with the help of two dummy variables \(s_1\) and \(s_2\), can be transformed into the equivalent form

Maximize
\[
\pi = f(x_1, x_2, x_3)
\]  
(9)

Subject to
\[
g^1(x_1, x_2, x_3) + s_1 = r_1
\]
\[
g^2(x_1, x_2, x_3) + s_2 = r_2
\]
and
\[
x_1, x_2, x_3, s_1, s_2 \geq 0
\]
If the nonnegativity restrictions are absent, we may, in line with the classical approach, form the Langrangian function:
\[
Z^* = f(x_1, x_2, x_3) + \lambda_1\left[r_1 - g^1(x_1, x_2, x_3) - s_1\right] + \lambda_2\left[r_2 - g^2(x_1, x_2, x_3) - s_2\right]
\]  
(10)

and write the first-order condition as
\[
\frac{\partial Z^*}{\partial x_1} = \frac{\partial Z^*}{\partial x_2} = \frac{\partial Z^*}{\partial x_3} = \frac{\partial Z^*}{\partial s_1} = \frac{\partial Z^*}{\partial s_2} = \frac{\partial Z^*}{\partial \lambda_1} = \frac{\partial Z^*}{\partial \lambda_2} = 0
\]  
(11)
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 (7). Consequently, we obtain the following set of conditions instead:

$$\begin{align*}
\frac{\partial Z^*}{\partial x_j} & \leq 0 \quad x_j \geq 0 \quad \text{and} \quad x_j \frac{\partial Z^*}{\partial x_j} = 0 \\
\frac{\partial Z^*}{\partial s_i} & \leq 0 \quad s_i \geq 0 \quad \text{and} \quad s_i \frac{\partial Z_i^*}{\partial s_i} = 0 \\
\frac{\partial Z^*}{\partial \lambda_i} & = 0 \\
\end{align*}$$

(12)

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

Each line (12) 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 tells us that we must have $-\lambda_i \leq 0$, $s_i \geq 0$ and $-s_i \lambda_i = 0$, or equivalently,

$$s_i \geq 0 \quad \lambda_i \geq 0 \quad \text{and} \quad s_i \lambda_i = 0$$

(13)

But the third line- a restatement of the constraints in (9)-means that $s_i = r_i - g^i(x_1, x_2, x_3)$. By substituting the latter into (13), therefore, we can combine the second and third lines of (12) into:

$$r_i - g^i(x_1, x_2, x_3) \geq 0 \quad \lambda_i \geq 0 \quad \text{and} \quad \lambda_i \left[ r_i - g^i(x_1, x_2, x_3) \right] = 0$$

This enables us to express the first-order condition (12) in an equivalent form without the dummy variables. Using the symbol $g^j_i$ to denote $\partial g^i/\partial x_j$, we now write

$$\begin{align*}
\frac{\partial Z^*}{\partial x_j} & = f_j - (\lambda_1 g^1_j + \lambda_2 g^2_j) \leq 0 \quad x_j \geq 0 \quad \text{and} \quad x_j \frac{\partial Z^*}{\partial x_j} = 0 \\
\frac{\partial Z^*}{\partial x_j} & \leq 0 \quad x_j \geq 0 \quad \text{and} \quad x_j \frac{\partial Z^*}{\partial x_j} = 0 \\
\end{align*}$$

(14)

$$r_i - g^i(x_1, x_2, x_3) \geq 0 \quad \lambda_i \geq 0 \quad \text{and} \quad \lambda_i \left[ r_i - g^i(x_1, x_2, x_3) \right] = 0$$
These, then, are the Kuhn-Tucker conditions for the problem (8), or, more accurately, one version of the Kuhn-Tucker conditions, expressed in terms of the Lagrangian function $Z^*$ in (10).

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 (21.10), let us ignore the nonnegativity restrictions as well as the inequality signs in the constraints and write the purely classical type of Lagrangian function $Z$:

$$Z = f(x_1, x_2, x_3) + \lambda_1 \left[ r_1 - g^1(x_1, x_2, x_3) \right] + \lambda_2 \left[ r_2 - g^2(x_1, x_2, x_3) \right]$$

Then let us (1) set the partial derivatives $\partial Z / \partial x_j \leq 0$, but $\partial Z / \partial \lambda_i \geq 0$, (2) impose nonnegativity restrictions on $x_j$ and $\lambda_i$, 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,

$$\frac{\partial Z}{\partial x_j} = f_j - (\lambda_1 g^1_j + \lambda_2 g^2_j) \leq 0 \quad x_j \geq 0 \quad \text{and} \quad x_j \frac{\partial Z^*}{\partial x_j} = 0$$
$$\frac{\partial Z}{\partial \lambda_i} = r_i - g^i(x_1, x_2, x_3) \geq 0 \quad \lambda_i \geq 0 \quad \text{and} \quad \lambda_i \frac{\partial Z}{\partial \lambda_i} = 0$$

are identical with (14), 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 (14)-as the partial derivative $\partial Z / \partial \lambda_i$. In the subsequent discussion, therefore, we shall only use the (16) version of the Kuhn-Tucker conditions, based on the Lagrangian function $Z$. 

7
2.2.1 Example: Utility Maximization with a simple rationing constraint

Consider a familiar problem of utility maximization with a budget constraint:

Maximize \( U = U(x, y) \)

subject to \( B = P_x x + P_y y \)

But where a ration on \( x \) has been imposed equal to \( \bar{x} \). The Lagrange becomes

\[
\max_{x, y} U(x, y) + \lambda_1 (B - P_x x - P_y y) + \lambda_2 (\bar{x} - x)
\]

The Kuhn-Tucker conditions are

\[
\begin{align*}
L_x &= U_x - P_x \lambda_1 - \lambda_2 x \leq 0 & x \geq 0 & \text{ and } & L_x \cdot x &= 0 \\
L_y &= U_y - P_y \lambda_1 \leq 0 & y \geq 0 & \text{ and } & L_y \cdot y &= 0 \\
L_{\lambda_1} &= B - P_x x - P_y y \geq 0 & \lambda_1 \geq 0 & \text{ and } & L_{\lambda_1} \cdot \lambda_1 &= 0 \\
L_{\lambda_2} &= \bar{x} - x \geq 0 & \lambda_2 \geq 0 & \text{ and } & L_{\lambda_2} \cdot \lambda_2 &= 0
\end{align*}
\]

Now let us interpret the Kuhn-Tucker conditions for this particular problem. Looking at the Lagrange

\[
U(x, y) + \lambda_1 (B - P_x x - P_y y) + \lambda_2 (\bar{x} - x)
\]

We require that

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

therefore either

\[
\lambda_1 = 0
\]

or

\[
B - P_x x - P_y y = 0
\]
If we interpret $\lambda_1$ as the marginal utility of the budget (Income), then if the budget constraint is not met the marginal utility of additional $B$ is zero ($\lambda_1 = 0$).

(2) Similarly for the ration constraint, either

$$\bar{x} - x = 0$$

or

$$\lambda_2 = 0$$

$\lambda_2$ can be interpreted as the marginal utility of relaxing the constraint.

Numerical example

Maximize $U = xy$

subject to:

$$100 \geq x + y$$

and

$$x \leq 40$$

The Lagrange is

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

and the Kuhn-Tucker conditions become

\[
\begin{align*}
L_x &= y - \lambda_1 - \lambda_2 \leq 0 & x \geq 0 & \text{ and } & L_x \cdot x &= 0 \\
L_y &= x - \lambda_1 \leq 0 & y \geq 0 & \text{ and } & L_y \cdot y &= 0 \\
L_{\lambda_1} &= 100 - x - y \geq 0 & \lambda_1 \geq 0 & \text{ and } & L_{\lambda_1} \cdot \lambda_1 &= 0 \\
L_{\lambda_2} &= 40 - x \geq 0 & \lambda_2 \geq 0 & \text{ and } & L_{\lambda_2} \cdot \lambda_2 &= 0
\end{align*}
\]

Which gives us four equations and four unknowns: $x, y, \lambda_1$ and $\lambda_2$. 

9
To solve, we typically approach the problem in a stepwise manner. First, ask if any \( \lambda_i \) could be zero. Try \( \lambda_2 = 0 \) (\( \lambda_1 = 0 \) does not make sense, given the form of the utility function), then

\[
x - \lambda_1 = y - \lambda_1 \quad \text{or} \quad x = y
\]

from the constraint \( 100 - x - y \) we get \( x^* = y^* = 50 \) which violates our constraint \( x \leq 40 \). Therefore \( x^* = 40 \) and \( y^* = 60 \), also \( \lambda_1^* = 40 \) and \( \lambda_2^* = 20 \).

### 2.3 Interpretation of the Kuhn-Tucker Conditions

Parts of the Kuhn-Tucker conditions (16) are merely a restatement of certain aspects of the given problem. Thus the conditions \( x_j \geq 0 \) merely repeat the nonnegativity restrictions, and the conditions \( \partial Z/\partial \lambda_i \geq 0 \) merely reiterate the constraints. To include these in (16), however, has the important advantage of revealing more clearly the remarkable symmetry between the two types of variables, \( x_j \) (choice variable) and \( \lambda_i \) (Lagrange multipliers). To each variable in each category, there corresponds a marginal condition-\( \partial Z/\partial x_j \leq 0 \) or \( \partial Z/\partial \lambda_i \geq 0 \)-to be satisfied by the optimal solution. Each of the variables must be non-negative 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 an equality, as in the classical context, *or* the choice variable in question must take a zero value, *or* both. Analogously, for each \( \lambda_i \), we must find in the optimal solution that *either* the marginal condition holds as an equality-meaning that the \( i \)th 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 (16). Assume the problem to be the familiar production problem. Then we have
\( f_j \equiv \text{the marginal gross profit of the } j \text{ th product} \)
\( \lambda_i \equiv \text{the shadow price of the } i \text{ th resource} \)
\( g^i_j \equiv \text{the amount of the } i \text{ th resource used up in producing the marginal unit of the } j \text{ th product} \)
\( \lambda_i g^i_j \equiv \text{the marginal imputed cost of the } i \text{ th resource incurred in producing a unit of the } j \text{ th product} \)
\( \sum_i \lambda_i g^i_j \equiv \text{the aggregate marginal imputed cost of the } j \text{ th product} \)

Thus the marginal condition
\[ \frac{\partial Z}{\partial x_j} = f_j - \sum_i \lambda_i g^i_j \leq 0 \]

requires that the marginal gross profit of the \( j \) th product be no greater than its aggregate marginal imputed cost; i.e., no underimputation is permitted. The complementary-slackness condition then means that, if the optimal solution calls for the active production of the \( j \) th 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 \((x_j^* = 0)\).\(^2\) 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.

\(^2\)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 \( a = 0 \).
As for the remaining conditions, which relate to the variables \( \lambda_i \), 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 (\( \partial Z / \partial \lambda_i^* = 0 \)). These, of course, are nothing but the implications of Duality Theorem II of the preceding chapter.

It is also possible, of course, to take the Lagrange-multiplier value \( \lambda_i^* \) 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 minding (\( \partial Z / \partial \lambda_i^* > 0 \)), 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 (\( \partial Z / \partial \lambda_i^* = 0 \)).

### 2.4 The n-Variable m-Constraint Case

The above 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

$$Z = f(x_1, x_2, \ldots, x_n) + \sum_{i=1}^{m} \lambda_i \left[ r_i - g^i(x_1, x_2, \ldots, x_n) \right] \quad (17)$$

And the Kuhn-Tucker conditions will be simply be

$$\frac{\partial Z}{\partial x_j} \leq 0 \quad x_j \geq 0 \quad \text{and} \quad x_j \frac{\partial Z}{\partial x_j} = 0 \quad [\text{maximization}]$$

$$\frac{\partial Z}{\partial \lambda_i} \geq 0 \quad \lambda_i \geq 0 \quad \text{and} \quad \lambda_i \frac{\partial Z}{\partial \lambda_i} = 0 \quad \left( \begin{array}{c} i = 1, 2, \ldots, m \\ j = 1, 2, \ldots, n \end{array} \right) \quad (18)$$

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 (16). Note that, aside from the change in the dimension of the problem, the Kuhn-Tucker conditions remain entirely the same as in the three-variable, two constraint case discussed before. The interpretation of these conditions should naturally also remain the same.

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 (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 (17)—directly apply the minimization version of the Kuhn-Tucker conditions as follows:

$$\frac{\partial Z}{\partial x_j} \geq 0 \quad x_j \geq 0 \quad \text{and} \quad x_j \frac{\partial Z}{\partial x_j} = 0 \quad [\text{minimization}]$$

$$\frac{\partial Z}{\partial \lambda_i} \leq 0 \quad \lambda_i \geq 0 \quad \text{and} \quad \lambda_i \frac{\partial Z}{\partial \lambda_i} = 0 \quad \left( \begin{array}{c} i = 1, 2, \ldots, m \\ j = 1, 2, \ldots, n \end{array} \right) \quad (19)$$
This you should compare with (18).

Reading (18) and (19) horizontally (rowwise), 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 $\lambda_i$ (second row). Reading them vertically (columnwise) on the other hand, we note that, for each $x_j$ and $\lambda_i$, 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 the next section, the Kuhn-Tucker maximum conditions (18) and minimum conditions (19) 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.

### 2.4.1 An Example

Let us check whether the Kuhn-Tucker conditions are satisfied by the solution to the following:

Minimize  

$$ C = (x_1 - 4)^2 + (x_2 - 4)^2 $$

Subject to  

$$ 2x_1 + 3x_2 \geq 6 $$

$$ -3x_1 - 2x_2 \geq -12 $$

and  

$$ x_1, x_2 \geq 0 $$

14
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 (19), which include the four marginal conditions

\[
\begin{align*}
\frac{\partial Z}{\partial x_1} &= 2(x_1 - 4) - 2\lambda_1 + 3\lambda_2 \geq 0 \\
\frac{\partial Z}{\partial x_2} &= 2(x_2 - 4) - 3\lambda_1 + 2\lambda_2 \geq 0 \\
\frac{\partial Z}{\partial \lambda_1} &= 6 - 2x_1 - 3x_2 \leq 0 \\
\frac{\partial Z}{\partial \lambda_2} &= -12 + 3x_1 + 2x_2 \leq 0
\end{align*}
\]

plus the nonnegativity and complementary-slackness conditions. The question is: Can we find nonnegativity values \( \lambda_1^* \) and \( \lambda_2^* \) which, together with the optimal values \( x_1^* = \frac{2}{13} = \frac{28}{13} \) and \( x_2^* = \frac{10}{13} = \frac{36}{13} \), will satisfy all those conditions?

Given that \( x_1^* \) and \( x_2^* \) are both nonzero, complementary-slackness dictates that \( \partial Z/\partial x_1 = 0 \) and \( \partial Z/\partial x_2 = 0 \), \( \partial Z/\partial \lambda_1^* < 0 \) and \( \partial Z/\partial \lambda_2^* = 0 \), which satisfy the marginal inequalities as well as the complementary-slackness conditions, all the Kuhn-Tucker minimum conditions are satisfied.

### 2.4.2 Problems:

1. Draw a set of diagrams similar to Fig. 21.3 for the minimization case, and deduce a set of necessary conditions for a local minimum corresponding to (1.5) through (1.7). Then condense these conditions into a single statement similar to (1.8).

   (a) Show that, in (1.21), instead of writing

   \[ y_i \frac{\partial Z}{\partial \lambda_i} = 0 \quad (i = 1, \ldots, m) \]
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 set of conditions

\[
x_j \frac{\partial Z}{\partial x_j} = 0 \quad (j = 1, \ldots, n)
\]

2. Based on the reasoning used in the preceding problem, which set (or sets) of conditions in (1.22) can be condensed into a single equation?

3. Given the minimization problem (1.3), and using the Lagrangian function (1.20), take the derivatives of \( \partial Z / \partial x_j \) and \( \partial Z / \partial \lambda_i \) and write out the expanded version of the Kuhn-Tucker minimum conditions (1.22).

4. Convert the minimization problem (1.3) into a maximization problem, formulate the Lagrangian function, take the derivatives with respect to \( x_j \) and \( y_i \), and apply the Kuhn-Tucker maximum conditions (1.21). Are the results consistent with those obtained in the preceding problem?

5. Check the applicability of the Kuhn-Tucker conditions to Example 2 of Sec. 21.1 as follows:

(a) Write the Lagrangian function and the Kuhn-Tucker conditions

(b) From the solution given in Fig. 21.1b, find the optimal values of \( \partial Z / \partial \lambda_i (i = 1, 2, 3) \). What can we conclude about \( \lambda_i^* \)?
(c) Now find the optimal values of $\partial Z/\partial x_1$ and $\partial Z/\partial x_2$.
(d) Are all Kuhn-Tucker conditions satisfied?

3 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 program, 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.

3.1 Irregularities at Boundary Points

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

**Example 1**

Maximize

$$\pi = x_1$$

Subject to

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

and

$$x_1, x_2 \geq 0$$

As shown in Fig. 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

$$Z = x_1 + \lambda_1 \left[ -x_2 + (1 - x_1)^3 \right]$$
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 \( \frac{\partial Z}{\partial x_1^*} = 1 \), thus violating the above marginal condition.

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
the 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. The following two examples will confirm this.

**Example 2**

To the problem of the preceding example, let us add a new con-
straint

\[ 2x_1 + x_2 \leq 2 \]

whose border, \( x_2 = 2 - 2x_1 \), plots as a straight line with slope -2
which passes through the optimal point in Fig. 2. Clearly, the feasible
region remains the same as before, as 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

\[
\begin{align*}
\frac{\partial Z}{\partial x_1} &= 1 - 3\lambda_1 (1 - x_1)^2 - 2\lambda_2 \leq 0 \\
\frac{\partial Z}{\partial x_2} &= -\lambda_1 - \lambda_2 \leq 0 \\
\frac{\partial Z}{\partial \lambda_1} &= -x_2 + (1 - x_1)^3 \geq 0 \\
\frac{\partial Z}{\partial \lambda_2} &= 2 - 2x_1 - x_2 \geq 0
\end{align*}
\]

it turns out that the values \( x_1^* = 1, \ x_2^* = 0, \ \lambda_1^* = 1, \ \lambda_2^* = \frac{1}{2} \) do
satisfy the above four inequalities, as well as the nonnegativity and
complementary-slackness conditions. As a mater of fact, \( \lambda_1^* \) can be
assigned any nonnegativity 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, how-
ever, this example shows that the Kuhn-Tucker conditions can remain
valid despite the cusp.
**Example 3**
The feasible region of the problem
Maximize
\[ \pi = x_2 - x_1^2 \]
Subject to
\[ -(10 - x_1^2 - x_2)^3 \leq 0 \]
\[ -x_1 \leq -2 \]
and
\[ x_1, x_2 \geq 0 \]
as shown in Fig. 21.5, contains no cusp anywhere. Yet, at the optimal solution, (2,6), the Kuhn-Tucker conditions nonetheless fail to hold. For, with the Lagrangian function
\[ Z = x_2 - x_1^2 + \lambda_1(10 - x_1^2 - x_2)^3 + \lambda_2(-2 + x_1) \]
the second marginal condition would require that
\[ \frac{\partial Z}{\partial x_2} = 1 - 3\lambda_1(10 - x_1^2 - x_2)^2 \leq 0 \]
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 \( \lambda_1 \). 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. 21.5. The fundamental reason why cusps are neither nor sufficient for the failure of the Kuhn-Tucker conditions is that the irregularities referred to above relate, not to the shape of the feasible region per se, but to the forms of the constraint functions themselves.
3.2 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 \( j \) th 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,

\[
dx_j \geq 0 \quad \text{if} \quad x_j^* = 0
\]  

(20)

Second, if the \( i \) th 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) or will not decrease (for a minimization problem), that is,

\[
dg^i(x^*) = g_1^i dx_1 + g_2^i dx_2 + \cdots + g_n^i dx_n \begin{cases} \leq 0 \text{ (maximization)} \\ \geq 0 \text{ (minimization)} \end{cases}
\]

if \( g^i(x^*) = r_i \)

where all the partial derivatives of \( g^i_j \) are to be evaluated at \( x^* \). If a vector \( dx \) satisfies (20) and (21), 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. 2, which fails the Kuhn-Tucker conditions, also fails the constraint qualification. At that point, \( x^*_2 = 0 \), thus the test vector must satisfy:

\[
dx_2 \geq 0 \quad \text{[by (20)]}
\]

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

\[
g_1^1 dx_1 + g_2^1 dx_2 = 3(1 - x_1)^2 dx_1 + dx_2 = dx_2 \leq 0
\]

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. 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 exists for it. Hence the optimal solution point \((1,0)\) violates the constraint qualification.

**Example 5**

Referring to Example 2 above, let us illustrate that, after an additional constraint \(2x_1 + x_2 \leq 2\) is added to Fig. 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

\[
g_1^2 dx_1 + g_2^2 dx_2 = 2dx_1 + dx_2 = 2dx_1 \leq 0
\]

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. 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.
3.3 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 program with linear constraints, or, as a special case, a linear program per se.

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. 21.6, 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 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 (21). But since \(x_1 = x_2 = 0\), we must choose test vectors with \(dx_1 \geq 0\) and \(dx_2 \geq 0\), by (20). Hence all test vectors from the origin must point in the due-east, due-north, or northeast directions, as depicted in Fig. 3.3. 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 (12). 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 (20) and (21) 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

\[
g_1^2dx_1 + g_2^2dx_2 = a_{21}dx_1 + a_{22}dx_2 \leq 0 \tag{22}\]

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 (22) and modify the restriction on the test vector to the form

\[
a_{21}(x_1 + dx_1) + a_{22}(x_2 + dx_2) \leq r_2 \tag{23}\]
Interpreting \((x_j + dx_j)\) to the new value of \(x_j\) attained at the arrowhead of a test vector, we may construe (23) 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 (20) can be ignored. However, for point \(N\), (21) dictates that

\[
g_1^1dx_1 + g_2^1dx_2 = a_{11}dx_1 + a_{12}dx_2 \leq 0
\]

Since point \(N\) satisfies \(a_{11}dx_1 + a_{12}dx_2 = r_1\) (first constraint border), we may add this equality to (21.23) and write

\[
a_{11}(x_1 + dx_1) + a_{12}(x_2 + dx_2) \leq r_1
\]

This would require the test vectors to have arrowheads located on or below the first constraint border in Fig. 21.6. 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 (20), but this time (21) requires all test vectors to satisfy both (22) and (24). Since we may modify the latter conditions to the forms in (23) and (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.

### 3.3.1 Problems:

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 \(C = 2x_1 + x_2\)
   Subject to \(x_1^2 - 4x_1 + x_2 \geq 0\)
   \(-2x_1 - 3x_2 \geq -12\)
   and \(x_1, x_2 \geq 0\)
   Solve graphically for the global minimum, and check whether the optimal solution satisfies (a) the constraint qualification and (b) the Kuhn-Tucker conditions.

5. Minimize \(C = x_1\)
   Subject to \(-x_2 - (1 - x_1)^3 \geq 0\)
   and \(x_1, x_2 \geq 0\)

27
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 (1.20) to the form

$$Z_0 = y_0 f(x_1, x_2, \ldots, x_n) + \sum_{i=1}^{m} \lambda_1 \left[ r_i - g_i(x_1, x_2, \ldots, x_n) \right]$$

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 $\lambda_0$.)

4 Economic Applications

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

Maximize

$$U = U(x, y)$$

Subject to

$$B \geq P_x x + P_y y$$

and

$$C \geq c_x x + c_y y$$

in addition, the non-negativity constraint $x \geq 0$ and $y \geq 0$.

The Lagrangian for the problem is

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

where $\lambda, \lambda_2$ are the Lagrange multiplier on the budget and coupon constraints respectively. The Kuhn-Tucker conditions are

$$Z_x = U_x - \lambda_1 P_x - \lambda_2 c_x \leq 0 \quad x \geq 0 \quad x \cdot Z_x = 0$$

$$Z_y = U_y - \lambda_1 P_y - \lambda_2 c_y \leq 0 \quad y \geq 0 \quad y \cdot Z_y = 0$$

$$Z_{\lambda_1} = B - P_x x - P_y y \geq 0 \quad \lambda_1 \geq 0 \quad \lambda_1 \cdot Z_{\lambda_1} = 0$$

$$Z_{\lambda_2} = C - c_x x - c_y y \geq 0 \quad \lambda_2 \geq 0 \quad \lambda_2 \cdot Z_{\lambda_2} = 0$$

**Numerical Example**

Let’s suppose the utility function is of the form $U = x \cdot y^2$. Further, let $B = 100, P_x = P_y = 1$ while $C = 120$ and $c_x = 2, c_y = 1$.

The Lagrangian becomes

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

The Kuhn-Tucker conditions are now

$$Z_x = y^2 - \lambda_1 - 2\lambda_2 \leq 0 \quad x \geq 0 \quad x \cdot Z_x = 0$$

$$Z_y = 2xy - \lambda_1 - \lambda_2 \leq 0 \quad y \geq 0 \quad y \cdot Z_y = 0$$

$$Z_{\lambda_1} = 100 - x - y \geq 0 \quad \lambda_1 \geq 0 \quad \lambda_1 \cdot Z_{\lambda_1} = 0$$

$$Z_{\lambda_2} = 120 - 2x - y \geq 0 \quad \lambda_2 \geq 0 \quad \lambda_2 \cdot Z_{\lambda_2} = 0$$

29
Solving the problem:

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

**Step one:** Assume \( \lambda_2 = 0, \lambda_1 > 0 \)

By ignoring the coupon constraint, the first order conditions become

\[
\begin{align*}
Z_x &= y^2 - \lambda_1 = 0 \\
Z_y &= 2xy - \lambda_1 = 0 \\
Z_{\lambda_1} &= 100 - x - y = 0
\end{align*}
\]

Solving for \( x \) and \( y \) yields

\[
x^* = 33.33 \quad y^* = 66.67
\]

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

\[
2(33.33) + 66.67 = 133.67 > 120
\]

The solution violates the coupon constraints.

**Step two:** Assume \( \lambda_1 = 0, \lambda_2 > 0 \)

Now the first order conditions become

\[
\begin{align*}
Z_x &= y^2 - 2\lambda_2 = 0 \\
Z_y &= 2xy - \lambda_2 = 0 \\
Z_{\lambda_2} &= 120 - 2x - y = 0
\end{align*}
\]

Solving this system of equations yields

\[
x^* = 20 \quad y^* = 80
\]
When we check our solution against the budget constraint, we find that the budget constraint is just met. In this case, we have the unusual result that the budget constraint is met but is not binding due to the particular location of the coupon constraint. The student is encouraged to carefully graph the solution, paying careful attention to the indifference curve, to understand how this result arose.

4.2 Peak Load Pricing

Peak and off-peak pricing and planning problems are common place 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 has been purchased to service the firm’s primary market, the capital is freely available (up to capacity) to be used in the secondary market. Typical examples include: schools and universities who build to meet day-time needs (peak), but may offer night-school classes (off-peak); theatres who offer shows in the evening (peak) and matinees (off-peak); or trucking companies who have dedicated routes but may choose to enter ”back-haul” markets. Since the capacity price 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 given that it is 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 at the lower (profit maximizing) price that 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 Kuhn-Tucker.

Consider a profit maximizing Company who faces two demand
curves

\[ P_1 = D^1(Q_1) \quad \text{in the day time (peak period)} \]
\[ P_2 = D^2(Q_2) \quad \text{in the night time (off-peak period)} \]

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

\[
\text{Maximize}_{Q_1, Q_2, K} \quad P_1 Q_1 + P_2 Q_2 - b(Q_1 - Q_2) - cK
\]

Subject to

\[ K \geq Q_1 \]
\[ K \geq Q_2 \]

Where

\[ P_1 = D^1(Q_1) \]
\[ P_2 = D^2(Q_2) \]

The Lagrangian for this problem is:

\[
Z = D^1(Q_1)Q_1 + D^2(Q_2)Q_2 - b(Q_1 + Q_2) - cK + \lambda_1(K - Q_1) + \lambda_2(K - Q_2)
\]

The Kuhn-Tucker conditions are

\[ Z_1 = D^1 + Q_1 \frac{\partial D^1}{\partial Q_1} - b - \lambda_1 \leq 0 \quad Q_1 \geq 0 \quad (MR_1 - b - \lambda_1 \leq 0) \]
\[ Z_2 = D^2 + Q_2 \frac{\partial D^2}{\partial Q_2} - b - \lambda_2 \leq 0 \quad Q_2 \geq 0 \quad (MR_2 - b - \lambda_2 \leq 0) \]
\[ Z_K = -c + \lambda_1 + \lambda_2 \leq 0 \quad K \geq 0 \quad (c \geq \lambda_1 + \lambda_2) \]
\[ Z_{\lambda_1} = K - Q_1 \geq 0 \quad \lambda_1 \geq 0 \]
\[ Z_{\lambda_2} = K - Q_2 \geq 0 \quad \lambda_2 \geq 0 \]
Assuming that \( Q_1, Q_2, K > 0 \) the first-order conditions become

\[
MR_1 = b + \lambda_1 = b + c - \lambda_2 \quad (\lambda_1 = c - \lambda_2)
MR_2 = b + \lambda_2
\]

Finding a solution:
Step One: Since \( D^2(Q_2) \) is smaller than \( D^1(Q_1) \) try \( \lambda_2 = 0 \)
Therefore from the Kuhn-Tucker conditions

\[
MR_1 = b + c - \lambda_2 = b + c
MR_2 = b + \lambda_2 = b
\]

which implies that \( K = Q_1 \). Then we check to see if \( Q_2^* \leq K \). If true, then we have a valid solution. Otherwise the second constraint is violated and the assumption that \( \lambda_2 = 0 \) was false. Therefore we proceed to the next step.

Step Two: if \( Q_2^* > K \) then \( Q_1^* = Q_2^* = K \) and

\[
MR_1 = b + \lambda_1
MR_2 = b + \lambda_2
\]

Since \( c = \lambda_1 + \lambda_2 \) then \( \lambda_1 \) and \( \lambda_2 \) represent the share of \( c \) each group pays. Both cases are illustrated in figure 4.2

**Numerical Example**
Suppose the demand during peak hours is
\[ P_1 = 22 - 10^{-5}Q_1 \]

and during off-peak hours 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 (labour and fuel) to produce 1 unit per half day (both day and evening).

If we assume that the capacity constraint is binding \((\lambda_2 = 0)\), then the Kuhn-Tucker conditions (above) become

\[
\begin{align*}
\lambda_1 &= c = 8 \\
\frac{MR}{22 - 2 \times 10^{-5}Q_1} &= b + c = 14 \\
\frac{MC}{18 - 2 \times 10^{-5}Q_2} &= b = 6
\end{align*}
\]

Solving this system gives us

\[
\begin{align*}
Q_1 &= 40000 \\
Q_2 &= 60000
\end{align*}
\]

which violates the assumption that the second constraint is non-binding \((Q_2 > Q_1 = K)\).

Therefore, assuming that both constraints are binding, then \(Q_1 = Q_2 = Q\) and the Kuhn-Tucker conditions become

\[
\begin{align*}
\lambda_1 + \lambda_2 &= 8 \\
22 - 2 \times 10^{-5}Q &= 6 + \lambda_1 \\
18 - 2 \times 10^{-5}Q &= 6 + \lambda_2
\end{align*}
\]
which yields the following solutions

\[ Q = K = 50000 \]
\[ \lambda_1 = 6 \quad \lambda_2 = 2 \]
\[ P_1 = 17 \quad P_2 = 13 \]

Since the capacity constraint is binding in both markets, market one pays \( \lambda_1 = 6 \) of the capacity cost and market two pays \( \lambda_2 = 2 \).

### 4.2.1 Problems

1. Suppose in the above example a unit of capacity cost 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. Skippy lives on an island where she produces two goods, \( x \) and \( y \), according the the production possibility frontier \( 200 \geq x^2 + y^2 \), and she consumes all the goods herself. Her utility function is

   \[ u = x \cdot y^3 \]

   Skippy also faces and environmental constraint on her total output of both goods. The environmental constraint is given by \( x + y \leq 20 \)

   (a) Write down the Kuhn Tucker first order conditions.

   (b) Find Skippy’s optimal \( x \) and \( y \). Identify which constraints are binding.

35
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 is given by
\( p_2 = 380 - q_2 \). The variable cost to 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 down 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 \( \lambda_1 \) and \( \lambda_2 \))? 

(d) Now suppose capacity cost is 30 per unit (paid only once).
Find quantities, capacity and how much of the capacity is
paid for by each market (i.e. \( \lambda_1 \) and \( \lambda_2 \))? 

5 Maximum Value Functions and the En-
velope Theorem\(^3\)

A maximum (or minimum) 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 (through the parameters’ influence on the optimal

\(^3\)This section of the chapter presents an overview of the envelope theorem for the purpose of
introducing the concept to the student. For a richer treatment of this topic is found in chapter 7 of
The Structure of Economics: a Mathematical Analysis (3rd Ed.) by Eugene Silberberg and Wing
Suen (McGraw-Hill, 2001) from which parts of this section are based.
values of the choice variables). Thus the maximum value function is also referred to as the indirect objective function.

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 gives 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 considering in the following sections.

To illustrate, consider the following maximization problem with two choice variables $x$ and $y$, and one parameter, $\alpha$:

Maximize

$$U = f(x, y, \alpha)$$  \hspace{1cm} (26)

The first order necessary condition are

$$f_x(x, y, \alpha) = f_y(x, y, \alpha) = 0$$  \hspace{1cm} (27)

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

$$x = x^*(\alpha) \quad y = x^*(\alpha)$$  \hspace{1cm} (28)

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

$$V(\alpha) = f(x^*(\alpha), y^*(\alpha), \alpha)$$  \hspace{1cm} (29)

where this function is the value of $f$ when the values of $x$ and $y$ are those that maximize $f(x, y, \alpha)$. Therefore, $V(\alpha)$ is the maximum
value function (or indirect objective function). If we differentiate $V$ with respect to $\alpha$

$$\frac{\partial V}{\partial \alpha} = f_x \frac{\partial x^*}{\partial \alpha} + f_y \frac{\partial y^*}{\partial \alpha} + f_\alpha$$  

(30)

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

$$\frac{\partial V}{\partial \alpha} = f_\alpha$$  

(31)

This result says that, at the optimum, as $\alpha$ varies, with $x^*$ and $y^*$ allowed to adjust optimally gives the same result as if $x^*$ and $y^*$ were held constant! Note that $\alpha$ enters maximum value function (equation 29) in three places: one direct and two indirect (through $x^*$ and $y^*$). Equations 30 and 31 show that, at the optimum, only the direct effect of $\alpha$ on the objective function matters. This is the essence of the envelope theorem. The envelope theorem says only the direct effects of a change in an exogenous variable need be considered, even though the exogenous variable may enter the maximum value function indirectly as part of the solution to the endogenous choice variables.

5.1 The Profit Function

Let’s apply the above approach to an economic application, namely the profit function of a competitive firm. Consider the case where a firm uses two inputs: capital, $K$, and labour, $L$. The profit function is

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

(32)

where $p$ is the output price and $w$ and $r$ are the wage rate and rental rate respectively.
The first order conditions are
\[
\begin{align*}
\pi_L &= f_L(K, L) - w = 0 \\
\pi_K &= f_K(K, L) - r = 0
\end{align*}
\]
which respectively define the factor demand equations
\[
\begin{align*}
L &= L^*(w, r, p) \\
K &= K^*(w, r, p)
\end{align*}
\]
substituting the solutions $K^*$ and $L^*$ into the objective function gives us
\[
\pi^*(w, r, p) = pf(K^*, L^*) - wL^* - rK^*
\]
\(\pi^*(w, r, p)\) is the profit function (or 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 (equation 32) with respect to $w$, holding all other variables constant and we get
\[
\frac{\partial \pi}{\partial w} = -L
\]
However, this result does not take into account the profit maximizing firms ability to make a substitution of capital for labour and adjust the level of output in accordance with profit maximizing behavior.

Since $\pi^*(w, r, p)$ is the maximum value of profits for any values of $w$, $r$, and $p$, changes in $\pi^*$ from a change in $w$ takes all captial for labour subsitutions into account. To evaluate a change in the maximum profit function from a change in $w$, we differentiate $\pi^*(w, r, p)$ with respect to $w$ yielding
\[
\frac{\partial \pi^*}{\partial w} = [pf_L - w] \frac{\partial L^*}{\partial w} + [pf_K - r] \frac{\partial K^*}{\partial w} - L^*
\]
From the first order conditions, the two bracketed terms are equal to zero. Therefore, the resulting equation becomes

$$\frac{\partial \pi^*}{\partial w} = -L^*(w, r, p)$$  \hspace{1cm} (38)

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

$$\frac{\partial \pi^*(w, r, p)}{\partial r} = -K^*(r, w, p)$$  \hspace{1cm} (39)

and

$$\frac{\partial \pi^*(w, r, p)}{\partial p} = f(K^*, L^*) = q^*$$  \hspace{1cm} (40)

The simple comparative static results derived from the profit function is known as ”Hotelling’s Lemma”. Hotelling’s Lemma is simply an application of the envelope theorem.

### 5.2 Reciprocity Conditions

Consider again our two variable maximization problem

$$\text{Maximize } U = f(x, y, \alpha)$$

where $x$ and $y$ are the choice variable, and $\alpha$ is a parameter. The first order equations are $f_x = f_y = 0$. which imply the functions $x = x^*(\alpha)$ and $y = y^*(\alpha)$. 

40
We are interested in the comparative statics regarding the directions of change in $x^*(\alpha)$ and $y^*(\alpha)$ as $\alpha$ changes and the implications to the value function. The maximum value function is

$$V(\alpha) = f(x^*(\alpha), y^*(\alpha), \alpha)$$

By definition, $V(\alpha)$ gives the maximum value of $f$ for a given $\alpha$.

Now consider the new function, sometimes called the *primal-dual objective function*, the difference between the actual and maximum value for a given $\alpha$,

$$Z = Z(x, y, \alpha) = f(x, y, \alpha) - V(\alpha)$$

This new function, $Z$, has a maximum of zero when $x = x^*$, $y = y^*$ and for any $x \neq x^*$, $y \neq y^*$ we find that $f \leq V$. In this framework $Z(x, y, \alpha)$ can be considered a function of three independent variables, $x, y,$ and $\alpha$. The maximum of $Z(x, y, \alpha) = f(x, y, \alpha) - V(\alpha)$ can be described by the first and second order conditions.

The first order conditions are:

$$Z_x(x, y, \alpha) = f_x = 0$$
$$Z_y(x, y, \alpha) = f_y = 0$$

and

$$Z_\alpha(x, y, \alpha) = f_\alpha - V_\alpha = 0$$

We can see that the first-order conditions of our new function $Z$ are nothing more that are the original maximum conditions for $f(x, y, \alpha)$ (equations 43), and the envelope theorem (equation 44). These first order conditions hold whenever $x = x^*(\alpha)$ and $y = y^*(\alpha)$ The sufficient second order conditions are

$$H = \begin{vmatrix} f_{xx} & f_{xy} & f_{x\alpha} \\ f_{yx} & f_{yy} & f_{y\alpha} \\ f_{\alpha x} & f_{\alpha y} & f_{\alpha\alpha} - V_{\alpha\alpha} \end{vmatrix}$$ (45)
where the Hessian Matrix is negative definite, or

\[ f_{xx} < 0, \quad f_{xx}f_{yy} - f_{xy}^2 > 0, \quad H < 0 \]  \tag{46}

In addition, if the second-order conditions are met, then \( f_{\alpha \alpha} - V_{\alpha \alpha} < 0 \) is also implied. This inequality is an important result which plays an essential role in many comparative static exercises. We know already that

\[ V_{\alpha}(\alpha) = f_{\alpha}(x^*(\alpha), y^*(\alpha), \alpha) \]  \tag{47}

Differentiating both sides with respect to \( \alpha \) yields

\[ V_{\alpha \alpha} = f_{\alpha x} \frac{\partial x^*}{\partial \alpha} - f_{\alpha y} \frac{\partial y^*}{\partial \alpha} + f_{\alpha \alpha} \]  \tag{48}

From the sufficient second order conditions and using Young’s theorem

\[ V_{\alpha \alpha} - f_{\alpha \alpha} = f_{\alpha x} \frac{\partial x^*}{\partial \alpha} + f_{\alpha y} \frac{\partial y^*}{\partial \alpha} > 0 \]  \tag{49}

Suppose that \( \alpha \) enters only in the \( x \) first order condition such that \( f_{\alpha y} = 0 \) Then equation 49 reduces to

\[ f_{\alpha x} \frac{\partial x^*}{\partial \alpha} > 0 \]  \tag{50}

which implies that \( f_{\alpha x} \) and \( \frac{\partial x^*}{\partial \alpha} \) will have the same sign. For example, in the profit maximization model:

\[ \pi = pf(K, L) - wL - rK \]  \tag{51}

Where the first order conditions are

\[ \pi_L = pf_L - w = 0 \]
\[ \pi_K = pf_K - r = 0 \]  \tag{52}

\[^4\text{This analysis is easily generalized to the n-variable case.}\]
The exogenous variable $w$ enters only the first order equation $pfL - w = 0$; it enters with a negative sign

$$\frac{\partial \pi_L}{\partial w} = -1$$  \hspace{1cm} (53)

Therefore 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 show the reciprocity condition $\frac{\partial L^*}{\partial r} = \frac{\partial K^*}{\partial w}$. From the indirect profit function $\pi^*(w, r, p)$ Hotelling’s Lemma gave us

$$\pi_w^* = \frac{\partial \pi^*}{\partial w} = -L^*(w, r, p)$$
$$\pi_r^* = \frac{\partial \pi^*}{\partial r} = -K^*(w, r, p)$$  \hspace{1cm} (54)

differentiating again and applying Young’s theorem

$$\pi_{wr}^* = -\frac{\partial L^*}{\partial r} = -\frac{\partial K^*}{\partial w} = \pi_{rw}^*$$  \hspace{1cm} (55)

or

$$\frac{\partial L^*}{\partial r} = \frac{\partial K^*}{\partial w}$$  \hspace{1cm} (56)

### 5.3 The Envelope Theorem and Constrained Optimization

Now let us turn our attention to the case of constrained optimization. Again we will have an objective function ($U$), two choice variables, ($x$ and $y$) and one parameter ($\alpha$) except now we introduce the following constraint:

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

The derivation of the envelope theorem for the models with one constraint is as follows:
The problem then becomes
Maximize

\[ U = f(x, y; \alpha) \]  \hspace{1cm} (57)

subject to

\[ g(x, y; \alpha) = 0 \]  \hspace{1cm} (58)

The Lagrangian for this problem is

\[ Z = f(x, y; \alpha) + \lambda g(x, y; \alpha) \]  \hspace{1cm} (59)

The first order conditions are

\[ Z_x = f_x + \lambda g_x = 0 \]
\[ Z_y = f_y + \lambda g_y = 0 \]
\[ Z_\lambda = g(x, y; \alpha) = 0 \]  \hspace{1cm} (60)

Solving this system of equations gives us

\[ x = x^*(\alpha) \quad y = y^*(\alpha) \quad \lambda = \lambda^*(\alpha) \]  \hspace{1cm} (61)

Substituting the solutions into the objective function, we get

\[ U^* = f(x^*(\alpha), y^*(\alpha), \alpha) = V(\alpha) \]  \hspace{1cm} (62)

where \( V(\alpha) \) is the indirect objective function, or maximum value function. This is the maximum value of \( y \) for any \( \alpha \) and \( x_i \)'s that satisfy the constraint.

How does \( V(\alpha) \) change as \( \alpha \) changes? First, we differentiate \( V \) with respect to \( \alpha \)

\[ \frac{\partial V}{\partial \alpha} = f_x \frac{\partial x^*}{\partial \alpha} + f_y \frac{\partial y^*}{\partial \alpha} + f_\alpha \]  \hspace{1cm} (63)

In this case, equation 63 will not simplify to \( \frac{\partial V}{\partial \alpha} = f_\alpha \) since \( f_x \neq 0 \) and \( f_y \neq 0 \). However, if we substitute the solutions to \( x \) and \( y \) into the constraint (producing an identity)

\[ g(x^*(\alpha), y^*(\alpha), \alpha) \equiv 0 \]  \hspace{1cm} (64)
and differentiating with respect to $\alpha$ yields

$$g_x \frac{\partial x^*}{\partial \alpha} + g_x \frac{\partial x^*}{\partial \alpha} + g_\alpha \equiv 0 \quad (65)$$

If we multiply equation 65 by $\lambda$ and combine the result with equation 63 and rearranging terms, we get

$$\frac{\partial V}{\partial \alpha} = (f_x + \lambda g_x) \frac{\partial x^*}{\partial \alpha} + (f_y + \lambda g_y) \frac{\partial y^*}{\partial \alpha} + f_\alpha + \lambda g_\alpha = Z_\alpha \quad (66)$$

Where $Z_\alpha$ is the partial deviative of the Lagrangian function with respect to $\alpha$, holding all other variable constant. In this case, the Langrangian functions serves as the objective function in deriving the indirect objective function.

While the results in equation 66 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 whether they enter 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 unconstrained case. However, if the parameter enters the constraint, the relation

$$V_{\alpha \alpha} \geq f_{\alpha \alpha}$$

will no longer hold.

### 5.4 Interpretation of the Lagrange Multiplier

In the consumer choice problem in chapter 12 we derived the result that the Lagrange multiplier, $\lambda$, represented the change in the value of the Lagrange function when the consumer’s budget changed. We loosely interpreted $\lambda$ 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

Maximize

\[ U = f(x, y) \]  \hspace{1cm} (67)

Subject to

\[ c - g(x, y) = 0 \]  \hspace{1cm} (68)

where \( c \) is a constant. The Lagrangian for this problem is

\[ Z = f(x, y) + \lambda (c - g(x, y)) \]  \hspace{1cm} (69)

The first order equations are

\[
\begin{align*}
Z_x &= f_x(x, y) - \lambda g_x(x, y) = 0 \\
Z_y &= f_y(x, y) - \lambda g_y(x, y) = 0 \\
Z_\lambda &= c - g(x, y) = 0
\end{align*}
\]  \hspace{1cm} (70)

From the first two equations in (70), we get

\[ \lambda = \frac{f_x}{g_x} = \frac{f_y}{g_y} \]  \hspace{1cm} (71)

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

Equations (70) implicitly define the solutions

\[ x = x^*(c) \quad y = y(c) \quad \lambda = \lambda^*(c) \]  \hspace{1cm} (72)

substituting (72) back into the Lagrangian yields the maximum value function

\[ V(c) = Z^*(c) = f(x^*(c), y^*(c)) + \lambda^*(c) (c - g(x^*_1(c), y^*(c))) \]  \hspace{1cm} (73)
differentiating with respect to $c$ yields

$$\frac{\partial Z^*}{\partial c} = f_x \frac{\partial x^*}{\partial c} + f_y \frac{\partial y^*}{\partial c} + (c - g(x^*(c), y^*(c))) \frac{\partial \lambda^*}{\partial c} - \lambda^*(c) \frac{\partial x^*}{\partial c} - \lambda^*(c) \frac{\partial y^*}{\partial c} + \lambda^*(c) \frac{\partial c}{\partial c}$$

(74)

by rearranging we get

$$\frac{\partial Z^*}{\partial c} = (f_x - \lambda^* x) \frac{\partial x^*}{\partial c} + (f_y - \lambda^* y) \frac{\partial y^*}{\partial c} + (c - g(x^*(c), y^*(c))) \frac{\partial \lambda^*}{\partial c} + \lambda^*$$

(75)

Note that the three terms in brackets are nothing more than the first order equations and, at the optimal values of $x$, $y$ and $\lambda$, these terms are all equal to zero. Therefore this expression simplifies to

$$\frac{\partial V(c)}{\partial c} = \frac{\partial Z^*}{\partial c} = \lambda^*$$

(76)

Therefore equals the rate of change of the maximum value of the objective function when $c$ changes ($\lambda$ is sometimes 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.

6 Duality and the Envelope Theorem

A consumer’s expenditure function and his indirect utility function are 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.
Let \( U(x, y) \) be a utility function in \( x \) 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.

Setting up the Lagrangian:

\[
Z = U(x, y) + \lambda(B - P_x x - P_y y) \tag{77}
\]

The first order conditions are

\[
\begin{align*}
Z_x &= U_x - \lambda P_x = 0 \\
Z_y &= U_y - \lambda P_y = 0 \\
Z_\lambda &= B - P_x X - P_y Y = 0 \tag{78}
\end{align*}
\]

This system of equations implicitly defines a solution for \( x^M, y^M \) and \( \lambda^M \) as a function of the exogenous variables \( B, P_x, P_y \).

\[
\begin{align*}
x^M &= x^M(P_x, P_y, B) \\
y^M &= y^M(P_x, P_y, B) \\
\lambda^M &= \lambda^M(P_x, P_y, B, \alpha) \tag{79}
\end{align*}
\]

The solutions to \( x^M \) and \( y^M \) are the consumer’s ordinary demand functions, sometimes called the "Marshallian" demand functions.\(^5\)

Substituting the solutions to \( x^* \) and \( y^* \) into the utility function yields

\[
U^* = U^*(x^M(B, P_x, P_y), y^M(B, P_x, P_y)) = V(B, P_x, P_y) \tag{80}
\]

Where \( V \) is the maximum value function, or indirect utility function.

Now consider the alternative, or dual, problem for the consumer; minimize total expenditure on \( x \) and \( y \) while maintaining a given level of utility, \( U^* \). The Langrangian for this problem is

\[
Z = P_x x + P_y y + \lambda(U^* - U(x, y)) \tag{81}
\]

\(^5\)Named after the famous economist Alfred Marshall, known to most economic students as "another dead guy."
The first order conditions are

\[ Z_x = P_x - \lambda U_x = 0 \]
\[ Z_y = P_y - \lambda U_y = 0 \]
\[ Z_\lambda = U^* - U(x, y; \alpha) = 0 \quad (82) \]

This system of equations implicitly define the solutions to \( x^h, y^h \) and \( \lambda^h \)

\[ x^h = x^h(U^*, P_x, P_y) \]
\[ y^h = y^h(U^*, P_x, P_y) \]
\[ \lambda^h = \lambda^h(U^*, P_x, P_y) \quad (83) \]

\( x^h \) and \( y^h \) are the compensated, or “real income” held constant demand functions. They are commonly referred to as ”Hicksian” demand functions, hence the h superscript.\(^6\)

If we compare the first two equations from the first order conditions in both utility maximization problem and expenditure minimization problem \((Z_x, Z_y)\), we see that both sets can be combined (eliminating \( \lambda \)) to give us

\[ \frac{P_x}{P_y} = \frac{U_x}{U_y} (= MRS) \quad (84) \]

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. If the target level of utility in the minimization problem is set equal to the value of the utility obtained in the solution to the maximization problem, namely \( U^* \), we obtain the following

\[ x^M(B, P_x, P_y) = x^h(U^*, P_x, P_y) \]
\[ y^M(B, P_x, P_y) = y^h(U^*, P_x, P_y) \quad (85) \]

or the solution to both the maximization problem and the minimization problem produce identical values for \( x \) and \( y \). However, the

\(^6\)Yet another famous, but dead economist, Sir John Hicks.
solutions are functions of different exogenous variables so any comparative statics exercises will produce different results.

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

$$P_x x^h(P_x, P_y, U^*) + P_y y^h(P_x, P_y, U^*) = E(P_x, P_y, U^*) \quad (86)$$

where $E$ is the minimum value function or expenditure function. The duality relationship in this case is

$$E(P_x, P_y, U^*, \alpha) = B \quad (87)$$

where $B$ is the exogenous budget from the maximization problem.

Finally, it can be shown from the first order conditions of the two problems that

$$\lambda^M = \frac{1}{\lambda^h} \quad (88)$$

6.1 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 the ratio of partial derivatives of the maximum value function. Substituting the optimal values of $x^M, y^M$ and $\lambda^M$ into the Lagrangian gives us

$$V(B, P_x, P_y) = U(x^M, y^M) + \lambda^M (B - P_x x^M - P_y y^M) \quad (89)$$

First differentiate with respect to $P_x$

$$\frac{\partial V}{\partial P_x} = (U_x - \lambda^M P_x) \frac{\partial x^M}{\partial P_x} + (U_y - \lambda^M P_y) \frac{\partial y^M}{\partial P_x} + (B - P_x x^M - P_y y^M) \frac{\partial \lambda^M}{\partial P_x} - \lambda^M x^M \quad (90)$$
\[
\frac{\partial V}{\partial P_x} = (0) \frac{\partial x^M}{\partial P_x} + (0) \frac{\partial y^M}{\partial P_x} + (0) \frac{\partial \lambda^M}{\partial P_x} - \lambda^M x^M = -\lambda^M x^M \tag{91}
\]

Next, differentiate the value function with respect to \( B \)

\[
\frac{\partial V}{\partial B} = (U_x - \lambda^M P_x) \frac{\partial x^M}{\partial B} + (U_y - \lambda^M P_y) \frac{\partial y^M}{\partial B} + B - P_x x^M - P_y y^M \frac{\partial \lambda^M}{\partial B} + \lambda^M \tag{92}
\]

\[
\frac{\partial V}{\partial B} = (0) \frac{\partial x^M}{\partial B} + (0) \frac{\partial y^M}{\partial B} + (0) \frac{\partial \lambda^M}{\partial B} + \lambda^M = \lambda^M \tag{93}
\]

Finally, taking the ratio of the two partial derivatives

\[
\frac{\partial V}{\partial P_x} = -\frac{\lambda^M x^M}{\lambda^M} = x^M \tag{94}
\]

which is Roy’s identity.

### 6.2 Shephard’s Lemma

Earlier in the chapter an application of the envelope theorem was the derivation of Hotelling’s Lemma, which states that the partial derivatives of the maximum value of the profit function yields the firm’s factory demand functions and the supply functions. A similar approach applied to the expenditure function yields Shepard’s Lemma.

Consider the consumer’s minimization problem. The Lagrangian is

\[
Z = P_x x + P_y y + \lambda(U^* - U(x, y)) \tag{95}
\]

From the first order conditions, the solutions are implicitly defined

\[
x^h = x^h(P_x, P_y, U^*) \]
\[
y^h = y^h(P_x, P_y, U^*) \]
\[
\lambda^h = \lambda^h(P_x, P_y, U^*) \tag{96}
\]
Substituting these solutions into the Lagrangian yields the minimum value function
\[ V(P_x, P_y, U^*) = P_x x^h + P_y y^h + \lambda^h (U^* - U(x^h, y^h)) \] (97)

The partial derivatives of the value function with respect to \( P_x \) and \( P_y \) are the consumer’s conditional, or Hicksian, demands:
\[
\begin{align*}
\frac{\partial V}{\partial P_x} &= (P_x - \lambda^h U_x) \frac{\partial x^h}{\partial P_x} + (P_y - \lambda^h U_y) \frac{\partial y^h}{\partial P_x} + (U^* - U(x^h, y^h)) \frac{\partial \lambda^h}{\partial P_x} + x^h \\
\frac{\partial V}{\partial P_y} &= (0) \frac{\partial x^h}{\partial P_y} + (0) \frac{\partial y^h}{\partial P_y} + (0) \frac{\partial \lambda^h}{\partial P_y} + x^h = x^h
\end{align*}
\] (98)

and
\[
\begin{align*}
\frac{\partial V}{\partial P_y} &= (P_x - \lambda^h U_x) \frac{\partial x^h}{\partial P_y} + (P_y - \lambda^h U_y) \frac{\partial y^h}{\partial P_y} + (U^* - U(x^h, y^h)) \frac{\partial \lambda^h}{\partial P_y} + y^h \\
\frac{\partial V}{\partial P_x} &= (0) \frac{\partial x^h}{\partial P_x} + (0) \frac{\partial y^h}{\partial P_x} + (0) \frac{\partial \lambda^h}{\partial P_x} + y^h = y^h
\end{align*}
\] (99)

Differentiating \( V \) with respect to the constraint \( U^* \) yields \( \lambda^h \), the marginal cost of the constraint
\[
\begin{align*}
\frac{\partial V}{\partial U^*} &= (P_x - \lambda^h U_x) \frac{\partial x^h}{\partial U^*} + (P_y - \lambda^h U_y) \frac{\partial y^h}{\partial U^*} + (U^* - U(x^h, y^h)) \frac{\partial \lambda^h}{\partial U^*} + \lambda^h \\
\frac{\partial V}{\partial U^*} &= (0) \frac{\partial x^h}{\partial U^*} + (0) \frac{\partial y^h}{\partial U^*} + (0) \frac{\partial \lambda^h}{\partial U^*} + y^h = \lambda^h
\end{align*}
\]

Together, these three partial derivatives are Shepard’s Lemma.

6.3 Example of duality for the consumer choice problem

6.3.1 Utility Maximization

Consider a consumer with the utility function \( U = xy \), who faces a budget constraint of \( B = P_x x P_y y \), where all variables are defined as before.
The choice problem is
Maximize
\[ U = xy \]  
(100)
Subject to
\[ B = P_x x P_y y \]  
(101)
The Lagrangian for this problem is
\[ Z = xy + \lambda(B - P_x x P_y y) \]  
(102)
The first order conditions are
\[ Z_x = y - \lambda P_x = 0 \]
\[ Z_y = x - \lambda P_y = 0 \]
\[ Z_\lambda = B - P_x x - P_y y = 0 \]  
(103)
Solving the first order conditions yield the following solutions
\[ x^M = \frac{B}{2P_x} \quad y^M = \frac{B}{2P_y} \quad \lambda = \frac{B}{2P_x P_y} \]  
(104)
where \( x^M \) and \( y^M \) are the consumer’s Marshallian demand functions. Checking second order conditions, the bordered Hessian is
\[
|\mathbf{H}| = \begin{vmatrix}
0 & 1 & -P_x \\
1 & 0 & -P_y \\
-P_x & -P_y & 0
\end{vmatrix} = 2P_x P_y > 0 \]  
(105)
Therefore the solution does represent a maximum. Substituting \( x^M \) and \( y^M \) into the utility function yields the indirect utility function
\[ V(P_x, P_y, B) = \left( \frac{B}{2P_x} \right) \left( \frac{B}{2P_y} \right) = \frac{B^2}{4P_x P_y} \]  
(106)
If we denote the maximum utility by \( U_0 \) and re-arrange the indirect utility function to isolate \( B \)
\[ \frac{B^2}{4P_x P_y} = U_0 \]  
(107)
\[ B = (4P_xP_yU_0)^{\frac{1}{2}} = 2P_x^2P_y^2U_x^2 = E(P_x, P_y, U_0) \] (108)

We have the expenditure function

**Roy’s Identity**  Let’s verify Roy’s identity which states

\[ x^M = -\frac{\partial V}{\partial P_x} \] (109)

Taking the partial derivative of \( V \)

\[ \frac{\partial V}{\partial P_x} = -\frac{B^2}{4P_xP_y} \] (110)

and

\[ \frac{\partial V}{\partial B} = -\frac{B}{P_xP_y} \] (111)

Taking the negative of the ratio of these two partials

\[ -\frac{\partial V}{\partial P_x} = -\left( \frac{B^2}{4P_xP_y} \right) = \frac{B}{2P_x} = x^M \] (112)

Thus we find that Roy’s Identity does hold.

### 6.3.2 The dual and Shepard’s Lemma

Now consider the dual problem of cost minimization given a fixed level of utility. Letting \( U_0 \) denote the target level of utility, the problem is

Minimize

\[ P_x x + P_y y \] (113)
Subject to

\[ U_0 = xy \]  \hspace{1cm} (114)

The Lagrangian for the problem is

\[ Z = P_xx + P_yy + \lambda(U_0 - xy) \]  \hspace{1cm} (115)

The first order conditions are

\[ Z_x = P_x - \lambda y = 0 \]
\[ Z_y = P_y - \lambda x = 0 \]
\[ Z_\lambda = U_0 - xy = 0 \]  \hspace{1cm} (116)

Solving the system of equations for \( x, y \) and \( \lambda \)

\[ x^h = \left( \frac{P_y U_0}{P_x} \right)^{\frac{1}{2}} \]
\[ y^h = \left( \frac{P_x U_0}{P_y} \right)^{\frac{1}{2}} \]
\[ \lambda^h = \left( \frac{P_x P_y}{U_0} \right)^{\frac{1}{2}} \]  \hspace{1cm} (117)

where \( x^h \) and \( y^h \) are the consumer’s compensated (Hicksian) demand functions. Checking the second order conditions for a minimum

\[ |H| = \begin{vmatrix} 0 & -\lambda & -y \\ -\lambda & 0 & -x \\ -y & -x & 0 \end{vmatrix} = -2xy\lambda < 0 \]  \hspace{1cm} (118)

Thus the sufficient conditions for a minimum are satisfied. Substituting \( x^h \) and \( y^h \) into the original objective function gives us the minimum value function, or expenditure function

\[ P_xx^h + P_yy^h = P_x \left( \frac{P_y U_0}{P_x} \right)^{\frac{1}{2}} + P_y \left( \frac{P_x U_0}{P_y} \right)^{\frac{1}{2}} \]
\[ = (P_x P_y U_0)^{\frac{1}{2}} + (P_x P_y U_0)^{\frac{1}{2}} \]
\[ = 2P_x^{\frac{1}{2}} P_y^{\frac{1}{2}} U_0^{\frac{1}{2}} \]  \hspace{1cm} (119)

55
Note that the expenditure function derived here is identical to the expenditure function obtained by re-arranging the indirect utility function from the maximization problem.

**Shepard’s Lemma** We can now test Shepard’s Lemma by differentiating the expenditure function directly.

First, we derive the conditional demand functions

\[
\frac{\partial E(P_x, P_y, U_0)}{\partial P_x} = \frac{\partial}{\partial P_x} \left( 2P_x^{\frac{1}{2}} P_y^{\frac{1}{2}} U_0^{\frac{1}{2}} \right) = \frac{P_y^{\frac{1}{2}} U_0^{\frac{1}{2}}}{P_x^{\frac{1}{2}}} = x^h \tag{120}
\]

and

\[
\frac{\partial E(P_x, P_y, U_0)}{\partial P_y} = \frac{\partial}{\partial P_y} \left( 2P_x^{\frac{1}{2}} P_y^{\frac{1}{2}} U_0^{\frac{1}{2}} \right) = \frac{P_x^{\frac{1}{2}} U_0^{\frac{1}{2}}}{P_y^{\frac{1}{2}}} = y^h \tag{121}
\]

Next, we can find the marginal cost of utility (the Lagrange multiplier)

\[
\frac{\partial E(P_x, P_y, U_0)}{\partial U_0} = \frac{\partial}{\partial U_0} \left( 2P_x^{\frac{1}{2}} P_y^{\frac{1}{2}} U_0^{\frac{1}{2}} \right) = \frac{P_x^{\frac{1}{2}} P_y^{\frac{1}{2}}}{U_0^{\frac{1}{2}}} = \lambda^h \tag{122}
\]

Thus, Shepard’s Lemma holds in this example.

7 Income and Substitution Effects: The Slutsky Equation

7.1 The Traditional Approach

Consider a representative consumer who chooses only two goods: x and y. The price of both goods are determined in the market and
are therefore exogenous. As well, the consumer’s budget is also exogenously determined. The consumer choice problem then is

Maximize

\[ U(x, y) \]  \hspace{1cm} (123)\\

Subject to

\[ B = P_x X + P_y Y \]  \hspace{1cm} (124)\\

The Langrangian function for this optimization problem is

\[ Z = U(x, y) + \lambda(B - P_x x + P_y y) \]  \hspace{1cm} (125)\\

The first order conditions yield the following set of simultaneous equations:

\[ Z_{\lambda} = B - P_x x - P_y y = 0 \]
\[ Z_x = U_x - \lambda P_x = 0 \]
\[ Z_y = U_y - \lambda P_y = 0 \]  \hspace{1cm} (126)\\

Solving this system will allow us to express the optimal values of the endogenous variables as implicit functions of the exogenous variables:

\[ \lambda^* = \lambda^*(P_x, P_y, B) \]
\[ x^* = x^*(P_x, P_y, B) \]
\[ y^* = y^*(P_x, P_y, B) \]

If the bordered Hessian in the present problem is positive

\[ |H| = \begin{vmatrix} 0 & -P_x & -P_y \\ -P_x & U_{xx} & U_{xy} \\ -P_y & U_{yx} & U_{yy} \end{vmatrix} = 2P_x P_y U_{xy} - P_y^2 U_{xx} - P_x^2 U_{yy} > 0 \]  \hspace{1cm} (127)\\

then the value of \( U \) will be a maximum.
By substituting the optimal values $x^*$, $y^*$ and $\lambda^*$ into the first order equations, we convert these equations into equilibrium identities:

\[
B - P_y x^* - P_y y^* \equiv 0 \\
U_x (x^*, y^*) - \lambda^* P_x \equiv 0 \\
U_y (x^*, y^*) - \lambda^* P_y \equiv 0
\]

By taking the total differential of each identity in turn, and noting that $U_{xy} = U_{yx}$ (Young’s Theorem), we then arrive at the linear system

\[
-P_x dx^* - P_y dy^* = x^* dP_x + y^* dP_y - dB \\
-P_x d\lambda^* + U_{xx} dx^* + U_{xy} dy^* = \lambda^* dP_x \\
-P_y d\lambda^* + U_{yx} dx^* + U_{yy} dy^* = \lambda^* dP_y
\]

Writing these equations in matrix form

\[
\begin{bmatrix}
0 & -P_x & -P_y \\
-P_x & U_{xx} & U_{xy} \\
-P_y & U_{yx} & U_{yy}
\end{bmatrix}
\begin{pmatrix}
d\lambda^* \\
dx^* \\
dy^*
\end{pmatrix}
= 
\begin{pmatrix}
x^* dP_x + y^* dP_y - dB \\
\lambda^* dP_x \\
\lambda^* dP_y
\end{pmatrix}
\]

To study the effect of a change in the budget, let the other exogenous differentials equal zero ($dP_x = dP_y = 0, dB \neq 0$). Then dividing through by $dB$, and applying the implicit function theorem, we have

\[
\begin{bmatrix}
0 & -P_x & -P_y \\
-P_x & U_{xx} & U_{xy} \\
-P_y & U_{yx} & U_{yy}
\end{bmatrix}
\begin{pmatrix}
d\lambda^*/dB \\
dx^*/dB \\
dy^*/dB
\end{pmatrix}
= 
\begin{pmatrix}
-1 \\
0 \\
0
\end{pmatrix}
\]

The coefficient matrix of this system is the Jacobian matrix, which has the same value as the bordered Hessian $|H|$ which is positive if the second order conditions are met. By using Cramer’s rule we can solve for the following comparative static

\[
\frac{\partial x^*}{\partial B} = \frac{1}{|H|} \begin{vmatrix}
0 & -1 & -P_y \\
-P_x & U_{xy} \\
-P_y & U_{yy}
\end{vmatrix} = \frac{1}{|H|} \begin{vmatrix}
-P_x & U_{xy} \\
-P_y & U_{yy}
\end{vmatrix} = \frac{|\bar{H}_{12}|}{|H|} \leq 0
\]

58
As before, in the absence of additional information about the relative magnitudes of $P_x$, $P_y$ and the cross partials, $U_{ij}$, we are unable to ascertain the sign of this comparative-static derivative. This means that the optimal $x^*$ may increase in the budget, $B$, depending on whether it is a normal or inferior good (ambiguous income effect).

Next, we may analyze the effect of a change in $P_x$. Letting $dP_y = dB = 0$ but keeping $dP_x \neq 0$ and dividing Equation 128 by $dP_x$ we obtain

$$\begin{bmatrix} 0 & -P_x & -P_y \\ -P_x & U_{xx} & U_{xy} \\ -P_y & U_{yx} & U_{yy} \end{bmatrix} \begin{bmatrix} \partial x^*/\partial P_x \\ \partial x^*/\partial P_x \\ \partial y^*/\partial P_x \end{bmatrix} = \begin{bmatrix} x^* \\ \lambda^* \\ 0 \end{bmatrix}$$

(131)

From this, the following comparative static emerges:

$$\frac{\partial x^*}{\partial P_x} = \frac{1}{|H|} \begin{vmatrix} 0 & x^* & -P_y \\ -P_x & \lambda^* & U_{xy} \\ -P_y & 0 & U_{yy} \end{vmatrix} = \frac{-x^*}{|H|} \begin{vmatrix} -P_x & U_{xy} \\ -P_y & U_{yy} \end{vmatrix} + \lambda^* \frac{H_{22}}{|H|} -P_y \begin{vmatrix} 0 & -P_y \\ -P_y & U_{yy} \end{vmatrix}$$

(132)

Note that there are two components in $\left( \frac{\partial x^*}{\partial P_x} \right)$. By comparing the first term to our previous comparative static $\left( \frac{\partial x^*}{\partial B} \right)$, we see that

$$(-x^*) \left| \frac{H_{12}}{|H|} \right| = (-x^*) \left( \frac{\partial x^*}{\partial B} \right) \leq 0$$

(133)

which can be interpreted as the income effect of a price change. The second term is the income compensated version of $\partial x^*/\partial P_x$, or the substitution effect of a price change, which is unambiguously negative:

$$\left( \frac{\partial x^*}{\partial P_x} \right)_{\text{compensated}} = \lambda^* \begin{vmatrix} 0 & -P_y \\ -P_y & U_{yy} \end{vmatrix} = \lambda^* \frac{H_{22}}{|H|} = \lambda^* \left( \frac{P^2_y}{|H|} \right) < 0$$

(134)
Hence, we can express Equation 132 in the form

\[
\frac{\partial x^*}{\partial P^*} = -\left(\frac{\partial x^*}{\partial B}\right)x^* + \left(\frac{\partial x^*}{\partial P_x}\right)_{\text{compensated}}
\]  

(135)

This result, which decomposes the comparative static derivative \((\partial x^*/\partial P_x)\) into two components, an income effect and a substitution effect, is the two-good version of the "Slutsky Equation."

### 7.2 Duality and the Alternative Slutsky

From the envelope theorem, we can derive the Slutsky decomposition in a more succinct manner. Consider first that from the utility maximum problem we derived solutions for \(x\) and \(y\)

\[
x^M = x^M(P_x, P_y, B) \\
y^M = y^M(P_x, P_y, B)
\]

(136)

which were the marshallian demand functions. Substituting these solutions into the utility function yielded the indirect utility function (or maximum value function)

\[
U^* = U(x^M(P_x, P_y, B), y^M(P_x, P_y, B)) = U^*(P_x, P_y, B)
\]

(137)

which could be rewritten to isolate \(B\) and giving us the expenditure function

\[
B^* = B(P_x, P_y, U^*)
\]

(138)

Second, from the budget minimization problem we derived the Hicksian, or compensated, demand function

\[
x^* = x^h(P_x, P_y, U^*)
\]

(139)
which, by Shephard’s lemma, is equivalent to the partial derivative of the expenditure function with respect to $P_x$:

$$\frac{\partial B(P_x, P_y, U^*)}{\partial P_x} = x^c(P_x, P_y, U^*)$$

(140)

Thus we know that if the maximum value of utility obtained from

$$\text{Max } U(x, y) + \lambda(B - P_x x - P_y y)$$

is the same value as the exogenous level of utility found in the constrained minimization problem

$$\text{Min } P_x x + P_y y + \lambda(U_0 - U(x, y))$$

(141)

the values of $x$ and $y$ that satisfy the first order conditions of both problems will be identical, or

$$x^c(P_x, P_y, U_0) = x^m(P_x, P_y, B)$$

(142)

at the optimum. If we substitute the expenditure function into $x^M$ in place of the budget, $B$, we get

$$x^c(P_x, P_y, U_0) = x^M(P_x, P_y, B^*(P_x, P_y, U_0))$$

(143)

Differentiate both sides of equation 143 with respect to $P_x$

$$\frac{\partial x^c(P_x, P_y, U_0)}{\partial P_x} = \frac{\partial x^M(P_x, P_y, B^*(P_x, P_y, U_0))}{\partial P_x}$$

$$+ \frac{\partial x^M(P_x, P_y, B^*(P_x, P_y, U_0))}{\partial B} \frac{\partial B(P_x, P_y, U_0)}{\partial P_x}$$

(144)

But we know from Shephard’s lemma that

$$\frac{\partial B(P_x, P_y, U_0)}{\partial P_x} = x_c$$

(145)
substituting equation 145 in to equation 144 we get

$$\frac{\partial x^c}{\partial P_x} = \frac{\partial x^M}{\partial P_x} + x^c \frac{\partial x}{\partial B}$$

(146)

Subtract \((x^c \frac{\partial x}{\partial B})\) from both sides gives us

$$\frac{\partial x^M}{\partial P_x} = -x^c \frac{\partial x}{\partial B} + \frac{\partial x^c}{\partial P_x}$$

(147)

If we compare equation (147) to equation (135) we see that we have arrived at the identical result. The method of deriving the slutsky decomposition through the application of duality and the envelope theorem is sometimes referred to as the ”instant slutsky”.

7.2.1 Problems:

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 maximization problem 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 \(p_y > B/2\)?

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

\[U^* = U(p_x, p_y, B)\]
and derive an expression for the expenditure function

\[ B^* = B(p_x, p_y, U^*) \]

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

\[
\text{Minimize } p_x x + p_y y \text{ subject to } U^* = x(y + 1)
\]

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, \( \partial B/\partial p_x \) and \( \partial B/\partial p_y \) respectively.