# Chapter 5. Maximum Value Functions

In Chapter 4, we learned that the Lagrange multipliers could be interpreted as the rate of change of the maximum attainable value of the objective function with repsect to the right-hand sides of the constraint equations. Apart from the parameter on the right-hand sides of the constraint equations $c$, there are other parameters in the objective funtion $F(x)$ as well as the constraint functions $G(x)$. In this chapter, we will learn how these parameters would affect the maximum attainable value in general.

## 5.A. Parameters in the Objective Function

We will begin with equality constraints. And consider first the case where the parameters affect the maximand alone. We have already seen such an example in Exercise 2.3 *Production and Cost Minimization.*

> Exercise 2.3: Consider a producer who rents machines $K$ at $r$ per year and hires labor $L$ at wage $w$ per year to produce output $Q$, where $Q = \sqrt{K} + \sqrt{L}$. Suppose he wishes to produce a fixed quantity $Q$ at minimum cost.
>
> Find his factor demand function, that is, the optimal amount of $K$ and $L$.

The maximization problem for this example is:

$$\max_{K,L} -Kr - Lw \tag{MP1}$$
$$\text{s.t. } \sqrt{K} + \sqrt{L} = Q.$$

The parameters $r$ and $w$ only appear in the objective function, i.e., the maximand.

The general representation for such maximization problem is

$$v = \max_{x} F(x, \theta)$$
$$\text{s.t. } G(x) = c,$$

where $\theta$ is a vector of parameters. The Lagrangian is

$$\mathcal{L}(x, \lambda, \theta) = F(x, \theta) + \lambda \left[c - G(x)\right].$$

The first-order necessary conditions are

$$\mathcal{L}_x(x^*, \lambda, \theta) = F_x(x^*, \theta) - \lambda G_x(x^*) = 0, \text{ and } \mathcal{L}_\lambda(x^*, \lambda, \theta) = c - G(x^*) = 0.$$

Now suppose that $\theta$ changes to $\theta + \mathrm{d}\theta$. Correspondingly, the optimum $x^*$ changes to $x^* + \mathrm{d}x^*$, and the maximum value $v$ changes to $v + \mathrm{d}v$.

Next, we evaluate the change of maximum value $\mathrm{d}v$ using the method we learned in Chapter 4.

$$\mathrm{d}v \underbrace{=}_{\text{by definition}} (v + \mathrm{d}v) - v \underbrace{=}_{\text{by definition}} F(x^* + \mathrm{d}x^*, \theta + \mathrm{d}\theta) - F(x^*, \theta) \underbrace{=}_{\text{Taylor approximation}} F_x(x^*, \theta)\mathrm{d}x^* + F_\theta(x^*, \theta)\mathrm{d}\theta$$

$$\underbrace{=}_{\text{First-order condition}} \lambda G_x(x^*)\mathrm{d}x^* + F_\theta(x^*, \theta)\mathrm{d}\theta \underbrace{=}_{\text{Taylor approximation}} \lambda\left[G(x^* + \mathrm{d}x^*) - G(x^*)\right] + F_\theta(x^*, \theta)\mathrm{d}\theta \underbrace{=}_{G(x^*+\mathrm{d}x^*)=G(x^*)=c} F_\theta(x^*, \theta)\mathrm{d}\theta.$$

Therefore, we get

$$\mathrm{d}v = F_\theta(x^*, \theta)\mathrm{d}\theta. \tag{5.1}$$

The result tells us that to find the first-order change in the maximum value of the objective function in response to changes in parameters that do not affect the constraints, we need not worry about the simultaneous change in the optimum choice $x^*$ itself. All we have to do is to calculate the partial effect of the parameter change, and evaluate the expression at the initial optimum choice.

We could check (5.1) in our cost minimization example. Return to Exercise 2.3. Solving the problem (MP1), we obtain

$$K^* = \left(\frac{wQ}{r + w}\right)^2 \text{ and } L^* = \left(\frac{rQ}{r + w}\right)^2.$$

In the problem, $F(K, L, w, r) = -Kr - Lw$ and $G(K, L) = \sqrt{K} + \sqrt{L}$.

Plugging the optimal $K^*$ and $L^*$ into the objective function, we obtain

$$v(w, r, Q) = F(K^*, L^*, w, r) = -\left(\frac{wQ}{r + w}\right)^2 r - \left(\frac{rQ}{r + w}\right)^2 w = -\frac{wrQ^2}{r + w}.$$

We could check the equivalence of $\mathrm{d}v(w, r, Q)/\mathrm{d}\theta$ and $F_\theta(K^*, L^*, r, w)$, for $\theta = w$ and $\theta = r$.

Here, we only check $\mathrm{d}v(w,r,Q)/\mathrm{d}w$ and $F_w(K^*,L^*,w,r)$. The other set of equivalence could be checked similarly.

$$\mathrm{d}v(w,r,Q)/\mathrm{d}w = -\frac{rQ^2(r+w) - wrQ^2}{(r+w)^2} = -\left(\frac{rQ}{r+w}\right)^2 = F_w(K^*,L^*,w,r).$$

## 5.B. The Envelope Theorem

The algebra of the previous section is illustrated geometrically in Figure 5.1.

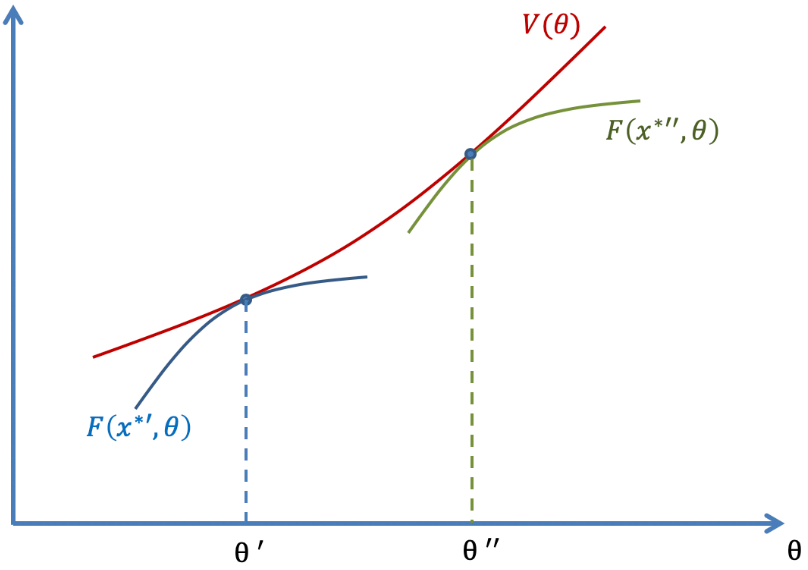


Figure 5.1: The Envelope Theorem

For particular value of $\theta$, say $\theta'$, the optimal choice is $x^{*\prime}$. That is , $x^{*\prime}$ solves the below maximization problem:

$$\max_x F(x,\theta')$$
$$\text{s.t. } G(x) = c.$$

The curve $F(x^{*\prime},\theta)$ represents the value of the objective function with respect to $\theta$, where $x$ is held fixed at $x^{*\prime}$.

The curve $V(\theta)$ represents the optimum value function with respect to $\theta$, where $x$ is allowed to vary optimally as $\theta$ varies. Formally, the function $V(\theta)$ is defined by

$$V(\theta) = \max_x\{F(x,\theta)|G(x) = c\}, \tag{5.2}$$

which is read as "$V(\theta)$ is the maximum over $x$ of $F(x,\theta)$ subject to $G(x) = c$."

Next, write the optimum choice $x^*$ as a function of $\theta$: $x^* = X(\theta)$. Pratically, this could be done by solving the maximization problem for fixed $\theta$. For instance, we have $x^{*\prime} = X(\theta')$. Then, we could rewrite $V(\theta)$ as follows:

$$V(\theta) = F(x^*, \theta) = F(X(\theta), \theta).$$

The two functions $V(\theta)$ and $F(x^{*\prime}, \theta)$ coincide at $\theta'$, because $x^{*\prime}$ happens to be the optimal choice there. Algebraically,

$$F(x^{*\prime}, \theta') = F(X(\theta'), \theta') = V(\theta').$$

For the other values of $\theta$, unless $x^{*\prime}$ remains the optimal choice, the curve $V(\theta)$, which is the optimum value, is higher than that of $F(x^{*\prime}, \theta)$. Algebraically,

$$F(x^{*\prime}, \theta) = F(X(\theta'), \theta) \underbrace{\leq}_{X(\theta) \text{ is the optimal choice given } \theta} F(X(\theta), \theta) = V(\theta) \text{ for } \theta \neq \theta'.$$

Therefore, the two curves should be **mutually tangential** at $\theta'$. This is what (5.1) expresses.

Similarly, we could draw the graph of $F(x^{*\prime\prime}, \theta)$, where $x^{*\prime\prime}$ is the optimal choice at $\theta''$. We know from the previous analysis, the curve $F(x^{*\prime\prime}, \theta)$ would touch the curve of the optimal value function $V(\theta)$ at $\theta''$.

We could draw a whole family of curves of $F(x, \theta)$ for a whole range of fixed values $x$, each $x$ being the optimal for some $\theta$. No member of this family of curves could ever cross above the graph of $V(\theta)$, and each would be tangential to the optimal value function at that value of $\theta$ where its $x$ happens to be the optimal choice. In other words, the optimal value function is the upper envelope of the family of the value functions, in each of which the choice variables are held fixed. That is why the formula (5.1) is often referred to as the *Envelope Theorem*.

We apply the *Envelope Theorem* to the cost minimization problem:

$$\max_x(-\theta x)$$

$$\text{s.t. } G(x) = c$$

where $\theta$ is a vector of input prices.

Figure 5.2 shows the the minimum cost curve and the cost lines for fixed $x$ when the first input price $\theta_1$ varies.

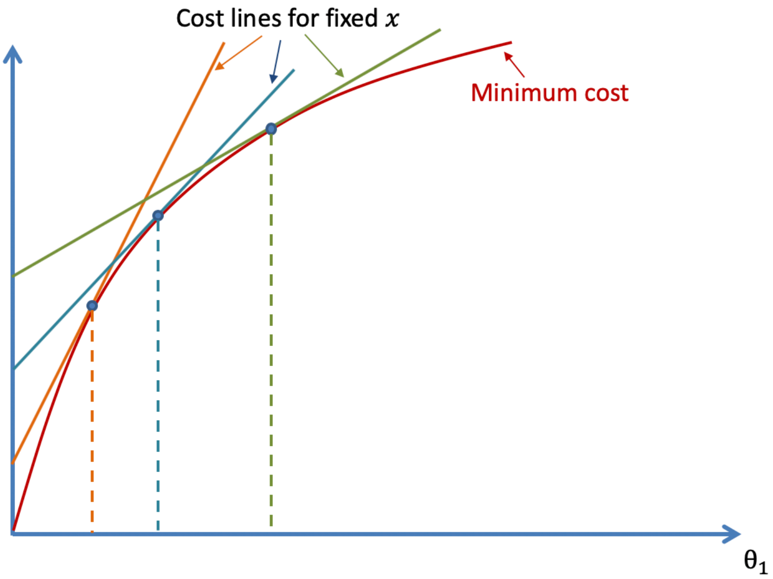


Figure 5.2: The minimum cost function

The cost lines are linear since when $x$ is held fixed, the cost $\theta x$ is a linear function of $\theta_1$. The minimized cost as a function of $\theta$ is the lower envelope (not the upper envelope, since this is a minimization problem) of all these straight lines.

The intuition can be shown graphically. See Figure 5.3 below.

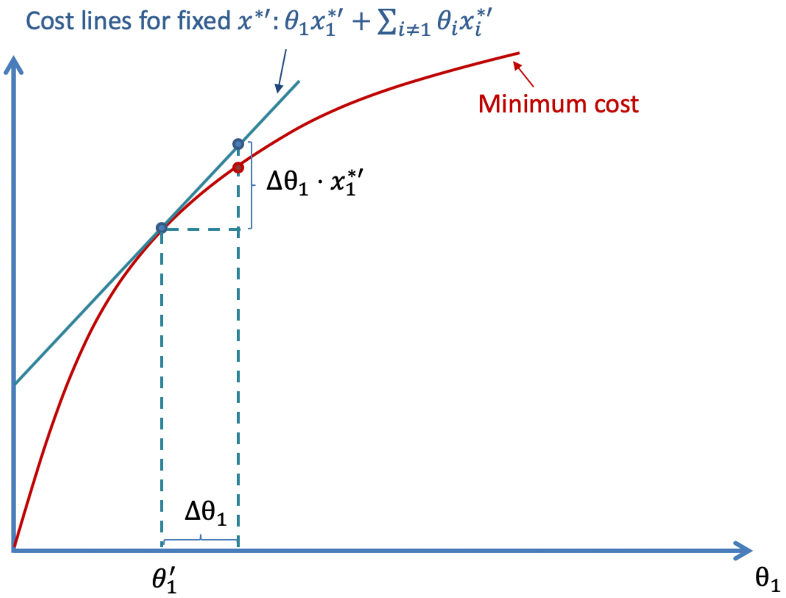


Figure 5.3: Lower Envelope

If $\theta_1'$ increases by $\Delta\theta_1$, assuming that $x^{*\prime}$ is held fixed, the cost increases with $\theta_1$ by the amount $\Delta\theta_1 \cdot x_1^{*\prime}$ (cost lines for fixed $x$ are linear). However, the producer can lower the cost by adjusting $x$ optimally; similarly if $\theta_1'$ decreases by $\Delta\theta_1$, assuming that $x^{*\prime}$ is held fixed, the cost decreases with $\theta_1$ by the amount $\Delta\theta_1 \cdot x_1^{*\prime}$. The producer can lower the cost by adjusting $x$ optimally (minimum cost curve is the lower envelope of the cost lines).

Figures 5.1 and 5.2 also suggest a curvature property. Figure 5.1 shows each $F(x^*, \theta)$ as a concave curve and $V(\theta)$ as a convex curve. Figure 5.2 shows a linear cost function for any fixed input choice but the lower envelope (the minimum cost curve) is concave. In general, the upper envelope must be more convex than any member of the family of which it is the envelope. This property will be studied in detail in Chapter 8.

## 5.C. Parameters Affecting All Functions

Now suppose $G$ as well as $F$ involves $\theta$. The general representation for such maximization problem is

$$v = \max_x F(x, \theta)$$
$$\text{s.t. } G(x, \theta) = c,$$

where $\theta$ is a vector of parameters. The Lagrangian is

$$\mathcal{L}(x, \lambda, \theta) = F(x, \theta) + \lambda \left[c - G(x, \theta)\right].$$

The first-order necessary conditions are

$$\mathcal{L}_x(x^*, \lambda, \theta) = F_x(x^*, \theta) - \lambda G_x(x^*, \theta) = 0, \text{ and } \mathcal{L}_\lambda(x^*, \lambda, \theta) = c - G(x^*, \theta) = 0.$$

The calculation for a change in $\theta$ to $\theta + \mathrm{d}\theta$ proceeds as in Section 5.A, except that the partial derivative of $G$ with respect to $x$ is no longer zero:

$$\begin{cases} G(x^* + \mathrm{d}x^*, \theta + \mathrm{d}\theta) - G(x^*, \theta) \underbrace{=}_{\text{Taylor approximation}} G_x(x^*, \theta)\mathrm{d}x^* + G_\theta(x^*, \theta)\mathrm{d}\theta \\ G(x^* + \mathrm{d}x^*, \theta + \mathrm{d}\theta) = G(x^*, \theta) = c \end{cases}$$

$$\Longrightarrow G_x(x^*, \theta)\mathrm{d}x^* + G_\theta(x^*, \theta)\mathrm{d}\theta = 0 \Longrightarrow G_x(x^*, \theta)\mathrm{d}x^* = -G_\theta(x^*, \theta)\mathrm{d}\theta. \tag{5.3}$$

Using this, and the previous analysis, we have

$$\mathrm{d}v \underbrace{=}_{\text{by definition}} (v+\mathrm{d}v) - v \underbrace{=}_{\text{by definition}} F(x^*+\mathrm{d}x^*, \theta+\mathrm{d}\theta) - F(x^*,\theta) \underbrace{=}_{\text{Taylor approximation}} F_x(x^*,\theta)\mathrm{d}x^* + F_\theta(x^*,\theta)\mathrm{d}\theta$$

$$\underbrace{=}_{\text{First-order condition}} \lambda G_x(x^*,\theta)\mathrm{d}x^* + F_\theta(x^*,\theta)\mathrm{d}\theta \underbrace{=}_{\text{Equation (5.3)}} -\lambda G_\theta(x^*,\theta)\mathrm{d}\theta + F_\theta(x^*,\theta)\mathrm{d}\theta = \mathcal{L}_\theta(x^*,\lambda,\theta)\mathrm{d}\theta.$$

Therefore, we get

$$\mathrm{d}v = \mathcal{L}_\theta(x^*,\lambda,\theta)\mathrm{d}\theta. \tag{5.4}$$

The difference between (5.4) and (5.1) has an intuitive explanation. When $\theta$ affects the constraints, a change $\mathrm{d}\theta$ has the direct effect of increasing the value of $G$ by $G_\theta(x^*,\theta)\mathrm{d}\theta$. This acts exactly like an equal reduction in $c$. The interpretation of the Lagrange multiplier tells us that the equivalent reduction in $c$ reduces $v$ by $\lambda G_\theta(x^*,\theta)\mathrm{d}\theta$. This is just the additional term in (5.4) when compared to (5.1).

In Chapter 4, we have learned a similar comparative static analysis with respect to changes in the parameters $c$. The more general formulation in this chapter can subsume the earlier case. To see this explicitly, define a larger vector of parameters $\widehat{\theta}$, which includes $\theta$ and $c$ as subvectors, and write the constraint as $\widehat{G}(x,\widehat{\theta}) = G(x,\theta) - c = 0$. The maximization problem is now

$$v = \max_x F(x,\theta)$$
$$\text{s.t. } \widehat{G}(x,\widehat{\theta}) = 0,$$

where $\widehat{\theta}$ is a vector of parameters. The Lagrangian is

$$\widehat{\mathcal{L}}(x,\lambda,\widehat{\theta}) = F(x,\theta) - \lambda\widehat{G}(x,\widehat{\theta}).$$

(5.4) becomes

$$\begin{aligned}
\mathrm{d}v &= \widehat{\mathcal{L}}_{\widehat{\theta}}(x^*,\lambda,\widehat{\theta})\mathrm{d}\widehat{\theta} = F_{\widehat{\theta}}(x^*,\theta)\mathrm{d}\widehat{\theta} - \lambda\widehat{G}_{\widehat{\theta}}(x,\widehat{\theta})\mathrm{d}\widehat{\theta} \\
&= F_\theta(x^*,\theta)\mathrm{d}\theta - \lambda\left[G_\theta(x,\theta)\mathrm{d}\theta - I_m\mathrm{d}c\right] \\
&= F_\theta(x^*,\theta)\mathrm{d}\theta - \lambda G_\theta(x,\theta)\mathrm{d}\theta + \lambda\mathrm{d}c \\
&= \mathcal{L}_\theta(x^*,\lambda,\theta)\mathrm{d}\theta + \lambda\,\mathrm{d}c.
\end{aligned}$$

The result $\mathrm{d}v = \mathcal{L}_\theta(x^*, \lambda, \theta)\mathrm{d}\theta + \lambda\,\mathrm{d}c$ includes the previous cases

$$\mathrm{d}v = \mathcal{L}_\theta(x^*, \lambda, \theta)\mathrm{d}\theta, \tag{5.4}$$

$$\text{and } \mathrm{d}v = \lambda\,\mathrm{d}c. \tag{4.3}$$

## 5.D. Some Choice Variables Fixed

In this section, we examine the effect of a change in parameters to the optimum value function when some components of $x$ are kept fixed. Our main focus is to compare such effect with the case where all components of $x$ could be freely adjusted. An economic application is the comparison between the short-run and the long-run outcomes. In the short-run, some choice variables, such as certain types of capital, could not be freely adjusted. However, in in the long-run, all choice variables could be adjusted. Example 5.1 fits such a story.

To tackle this problem, we partition the vector $x$ into two subvectors $y$ and $z$. In the long-run, both $y$ and $z$ are choice variables and could be adjusted freely, while in the short-run, $z$ is held fixed and only $y$ is allowed to vary.

Subsuming $c$ into $\theta$ as explained above, the long-run problem is

$$\max_{y,z} F(y, z, \theta) \qquad \text{s.t. } G(y, z, \theta) = 0. \tag{MP\_LR}$$

And the short-run problem is[1]

$$\max_{y} F(y, z, \theta) \qquad \text{s.t. } G(y, z, \theta) = 0. \tag{MP\_SR}$$

The only difference of these two problems is the choice variable.

Write the long-run optimal choices and the resulting value as functions of $\theta$:

$$y = Y(\theta), \quad z = Z(\theta), \quad v = V(\theta). \tag{5.5}$$

[1] For the short-run problem to be meaningful, the number of constraints must be less than the dimension of $y$.

In the short-run, $z$ should be treated as just another parameter along with $\theta$, and the optimal choice $y$ and the resulting value $v$ are functions of $(z, \theta)$:

$$y = Y(z, \theta), \quad v = V(z, \theta). \tag{5.6}$$

We are now ready to compare the long-run and short-run optimum values. First, the long-run problem (MP_LR) has more choice variables compared to the short-run problem (MP_SR). Therefore, the optimum value in the long-run problem should be at least as large as the optimum value in the short-run problem, since for the long-run problem, we could at least fix $z$ at the short-run level and optimize using $y$ only. That is

$$V(\theta) \geq V(z, \theta) \text{ for all } (z, \theta).$$

And the two values $V(\theta)$ and $V(z, \theta)$ coincide when $z$ is just at the long-run optimal level $Z(\theta)$. Because when $z$ is at the optimal level $Z(\theta)$, being able to adjust it (the long-run case) or not (the short-run case) will not make a difference. Therefore, $V(\theta)$ is the *upper envelope* of the family of value functions $V(z, \theta)$, in each of which $z$ is held fixed.

We could draw a similar graph as Figure 5.1 to show the intuition. See Figure 5.4 below.

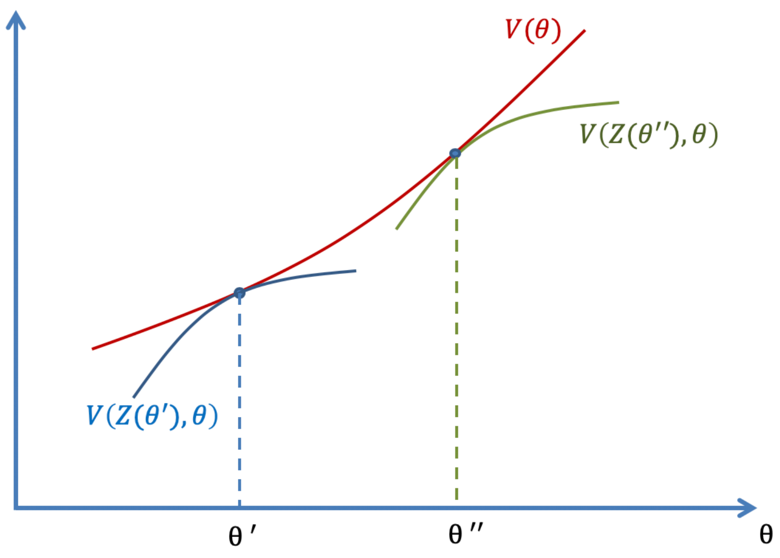


Figure 5.4: Short-run and Long-run

If the functions are differentiable, we would have

$$V'(\theta) = V_\theta(Z(\theta), \theta), \tag{5.7}$$

where the right-hand side is the partial derivative of the short-run optimum value function $V(Z(\theta), \theta)$ taken holding the first argument $z$ fixed, but evaluated at the point $z = Z(\theta)$. Please keep in mind that $V$ functions may not be differentiable even when $F$ and $G$ are. The problem may arise when we have inequality constraints and non-negativity constraints on the choice variables. At some point, there may be a regime change, one constraint from binding to slack or vice versa, and the graph of maximum value function may have a kink. Figure 4.2 provides such an example.

## 5.E. Examples

**Example 5.1: Short-Run and Long-Run Costs.** This example is used to illustrate Envelope Theorem. Consider a producer who rents machines $K$ at $r$ per year and hires labor $L$ at wage $w$ per year to produce output $Q$, where[2]

$$Q = (KL)^{1/\alpha}.$$

Suppose he wishes to produce a fixed quantity $Q$ at minimum cost.

Assume that $K$ is **fixed** in the short run; whereas $L$ could be freely adjusted.

**Question 1:** Calculate the long-run and short-run cost functions.

**Question 2:** Show that Equation (5.7) holds.

**Solution.**

**Question 1:** To make the problem easier to solve, we rewrite the constraint:

$$(KL)^{1/\alpha} = Q \implies KL = Q^\alpha.$$

The **long-run** problem is as follows:

$$C(w, r, Q) = \min_{K,L} rK + wL \iff -C(w, r, Q) = \max_{K,L} -rK - wL$$
$$\text{s.t. } KL = Q^\alpha \text{ and } K \geq 0,\ L \geq 0.$$

[2] Returns to scale are constant if $\alpha = 2$, increasing if $\alpha < 2$, and decreasing if $\alpha > 2$.

This is a maximization problem with non-negative variables. To solve this problem, we invoke Lagrange's Theorem.

i Form Lagrangian:

$$\mathcal{L}(K, L, \lambda) = -rK - wL + \lambda\left[-Q^\alpha + KL\right].$$

ii First-order conditions:

$$\partial\mathcal{L}(K, L, \lambda)/\partial K = -r + \lambda K \leq 0 \text{ and } K \geq 0, \text{ with complementary slackness;} \quad (5.8)$$

$$\partial\mathcal{L}(K, L, \lambda)/\partial L = -w + \lambda L \leq 0 \text{ and } L \geq 0, \text{ with complementary slackness;} \quad (5.9)$$

$$\partial\mathcal{L}(K, L, \lambda)/\partial \lambda = -Q^\alpha + KL = 0. \quad (5.10)$$

It is not hard to see that the constraint (5.10) will be violated if $K = 0$ or $L = 0$. Therefore, we must have $K > 0$ and $L > 0$. By the complementary slackness conditions in (5.8) and (5.9), we have

$$r - \lambda K = 0, \quad (5.11)$$

$$w - \lambda L = 0. \quad (5.12)$$

We could solve $L$, $K$ and $\lambda$ from the three equations (5.10), (5.11) and (5.12):

$$L^* = \left(\frac{rQ^\alpha}{w}\right)^{1/2}, \qquad K^* = \left(\frac{wQ^\alpha}{r}\right)^{1/2}, \qquad \lambda = \left(\frac{Q^\alpha}{wr}\right)^{-1/2}.$$

Then, the long-run cost function $C(w, r, Q)$ is

$$C(w, r, Q) = rK^* + wL^* = 2(wr)^{1/2}Q^{\alpha/2}.$$

In the **short run**, there is no freedom of choice. Since when $K$ is fixed, to produce output $Q$, labor $L = Q^\alpha/K$ must be hired. The short-run cost function $C(w, r, Q, K)$ is

$$C(w, r, Q, K) = rK + wQ^\alpha/K.$$

**Question 2:** Next, we check Equation (5.7).

The left-hand side of (5.7) concerns the long-run cost function:

$$C_Q(w, r, Q) = \alpha(wr)^{1/2}Q^{\alpha/2-1}.$$

The right-hand side of (5.7) concerns the short-run cost function.

First, differentiate the short-run cost function $C(w, r, Q, K)$ with repect to $Q$:

$$C_Q(w, r, Q, K) = w\alpha Q^{\alpha-1}/K.$$

Then evaluate $C_Q(w, r, Q, K)$ at $K^* = ((wQ^\alpha)/r)^{1/2}$:

$$C_Q(w, r, Q, K^*) = \alpha(wr)^{1/2}Q^{\alpha/2-1}.$$

This is exactly the same as $C_Q(w, r, Q)$. Therefore, Equation (5.7) holds.

**Example 5.2: Consumer Demand.**

**Part I: Indirect utility function.** Consider the consumer choice problem:

$$\max_x U(x)$$
$$\text{s.t. } px = I,$$

where $p$ is a row vector of prices, $x$ a column vector of quantities, and $I$ money income. The resulting maximum utility is a function $V(p, I)$, called the *indirect utility function*,[3] and the utility-maximizing quantities $x$ comprise the demand function $D(p, I)$.

Show that

$$D(p, I) = -V_p(p, I)/V_I(p, I). \tag{5.13}$$

**Solution.** The Lagrangian is

$$\mathcal{L}(x, \lambda, p, I) = U(x) + \lambda(I - px).$$

From the general result in this chapter (5.4):

$$V_I(p, I) = \mathcal{L}_I(x^*, \lambda, p, I) = \lambda, \tag{5.14}$$
$$V_p(p, I) = \mathcal{L}_p(x^*, \lambda, p, I) = -\lambda x^*. \tag{5.15}$$

[3] $U(x)$ is called the *direct utility function*.

Dividing (5.15) by (5.14), we have

$$V_p(p,I)/V_I(p,I) = -x^* \implies D(p,I) \underbrace{=}_{\text{by definition}} x^* = -V_p(p,I)/V_I(p,I).$$

**Part II: Expenditure function.** Consider the expenditure minimization problem:

$$\min_x px$$
$$\text{s.t. } U(x) = u,$$

where $p$, $x$ are defined as before, and $u$ is the target utility level. The resulting minimized expenditure is a function $E(p,u)$, called the *expenditure function*. Cost-minimizing commodity choices for a given utility level are called *Hicksian compensated demand function* $C(p,u)$. Show that

$$C(p,u) = E_p(p,u). \tag{5.16}$$

**Solution.** First, we transform the minimization problem into the maximization problem that we are familiar with:

$$\max_x -px$$
$$\text{s.t. } -U(x) = -u.$$

The Lagrangian is

$$\mathcal{L}(x,\mu,p,u) = -px + \mu\left[-u + U(x)\right].$$

From the general result in this chapter (5.4):

$$E_p(p,I) = -\mathcal{L}_p(x^*,\mu,p,u) = x^*. \tag{5.17}$$

That is,

$$C(p,u) \underbrace{=}_{\text{by definition}} x^* = E_p(p,I).$$

*Remark.* The uncompensated demand function (from the indirect utility function) and the compensated demand function (from the expenditure function) are related through the *Slutsky-Hicks equation*:

$$C^j_k(p,u) = D^j_k(p,I) + D^k(p,I)D^j_I(p,I),$$

or equivalently,

$$\left(\frac{\partial x_j}{\partial p_k}\right)_{u \text{ constant}} = \left(\frac{\partial x_j}{\partial p_k}\right)_{I \text{ constant}} + x_k\left(\frac{\partial x_j}{\partial I}\right).$$

To see this, suppose we begin with a utility level $u$, and find the expenditure-minimizing choice $C(p,u)$ and the minimized expenditure $E(p,u)$. Then we assign this $E(p,u)$ as income $I$, and find the utility-maximizing choice $D(p,I)$. Then, we would arrive at the same choice as in the expenditure mininnization problem, that is, $C(p,u) = D(p,I)$ if $I = E(p,u)$.

Take the $j^{th}$ component of $C(p,u) = D(p,I)$ and differentiate it with respect to $p_k$, holding $u$ fixed and recognizing $I = E(p,u)$ as a function of $p_k$, by chain rule,

$$C^j_k(p,u) = D^j_k(p,I) + D^j_I(p,I)E_k(p,u).$$

By (5.17),

$$E_k(p,u) \underbrace{=}_{(5.17)} C^k(p,u) \underbrace{=}_{C(p,u)=D(p,I)} D^k(p,I).$$

Therefore,

$$C^j_k(p,u) = D^j_k(p,I) + D^k(p,I)D^j_I(p,I),$$