

Chapter 7. Concave Programming

7.A. Concave Functions and Their Derivatives

In this chapter, we will combine the idea of convexity with a more conventional calculus approach. The result is that the Lagrange or Kuhn-Tucker conditions, in conjunction with convexity properties of the objective and constraint functions, are sufficient for optimality.

The first step is to express the concavity (convexity) of functions in terms of their derivatives. In chapter 6, we have learned the definition of concave functions:

Definition 6.B.5 (Concave Function). *A function $f : \mathcal{S} \rightarrow \mathbb{R}$, defined on a convex set $\mathcal{S} \subset \mathbb{R}^N$, is concave if*

$$f(\alpha x^a + (1 - \alpha)x^b) \geq \alpha f(x^a) + (1 - \alpha)f(x^b), \tag{6.5}$$

for all $x^a, x^b \in \mathcal{S}$ and for all $\alpha \in [0, 1]$.

We have also shown a similar graph to Figure 7.1 below, and interpreted *concavity* graphically: the graph of the **function** lies on or above the **chord** joining any two points of it.

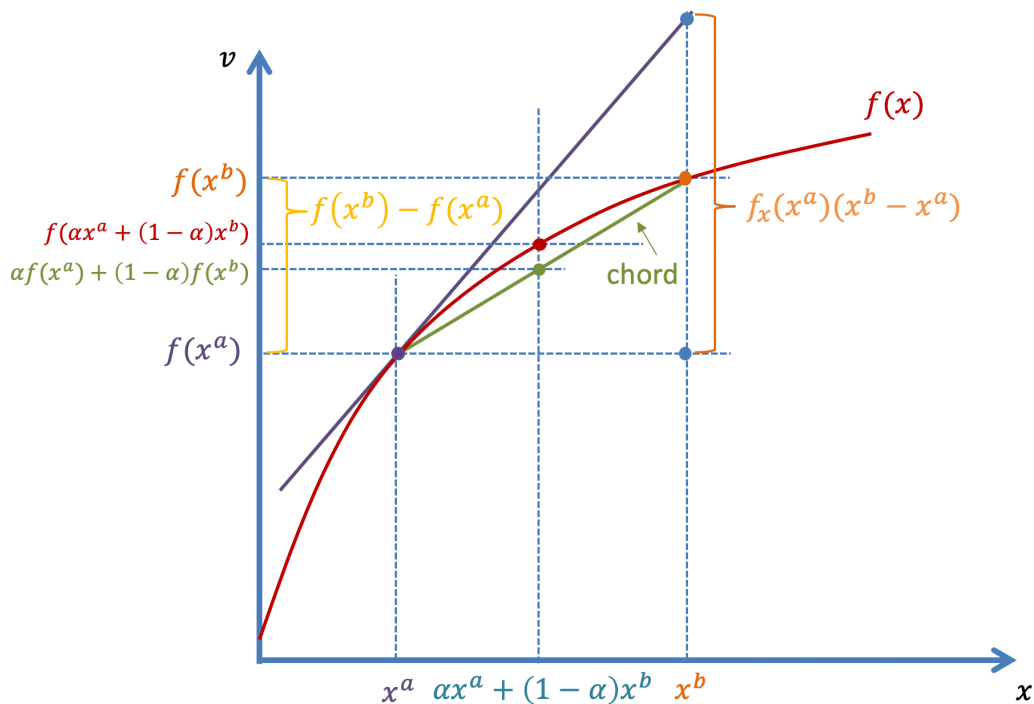


Figure 7.1: Concave Function

To express the concavity of $f(x)$ in terms of its derivative, we now draw the **tangent** to $f(x)$ at x^a . The requirement of concavity says that the graph of the **function** should lie on or below the **tangent**. Or expressed differently,

$$f_x(x^a)(x^b - x^a) \geq f(x^b) - f(x^a),$$

where $f_x(x^a)$ is the slope of the **tangent** to $f(x)$ at x^a .

Such an expression holds for higher dimensions. The result is summarized in Proposition 7.A.1 below.

Proposition 7.A.1 (Concave Function). *A differentiable function $f : \mathcal{S} \rightarrow \mathbb{R}$, defined on a convex set $\mathcal{S} \subset \mathbb{R}^N$, is concave if and only if*

$$f_x(x^a)(x^b - x^a) \geq f(x^b) - f(x^a), \tag{7.1}$$

for all $x^a, x^b \in \mathcal{S}$.

Proof. See Appendix A. □

Similarly, for a differentiable convex function f , we have

$$f_x(x^a)(x^b - x^a) \leq f(x^b) - f(x^a). \tag{7.2}$$

A particularly important class of optimization problems has a concave objective function and convex constraint functions; the term *concave programming* is often used to describe the general problem of this kind, and it is the subject of the next section.

7.B. Concave Programming

Consider the maximization problem

$$\begin{aligned} & \max_x F(x) \\ & \text{s.t. } G(x) \leq c, \end{aligned}$$

where F is differentiable and concave, and each component constraint function G^i is differentiable and convex.

Below we will interpret the problem using the terminology of the production problem. But the mathematics is independent of this interpretation.

$$\begin{aligned} & \max_x \underbrace{F(x)}_{\text{revenue from outputs}} \\ \text{s.t. } & \underbrace{G(x) \leq c}_{\text{resource constraints}} \end{aligned}$$

where x is the vector of outputs, c is a fixed vector of input supplies, and $G(x)$ is the vector of inputs needed to produce x .

Let $X(c)$ denote the optimum choice function, and $V(c)$ the maximum value function.

Claim 1. $V(c)$ is a non-decreasing function.

This is because an x that was feasible for a given value of c remains feasible when any component of c increases, so the maximum value cannot decrease.

Claim 2. $V(c)$ is a concave.

To show concavity of $V(c)$, we need to show that for any two input supply vectors c and c' and any number $\alpha \in [0, 1]$, we have

$$V(\alpha c + (1 - \alpha)c') \geq \alpha V(c) + (1 - \alpha)V(c').$$

That is, it should be possible to achieve revenue at least as high as $\alpha V(c) + (1 - \alpha)V(c')$ when the input supply vector is $\alpha c + (1 - \alpha)c'$.

Let $x^* = X(c)$ and $x^{*'} = X(c')$. Since the optimal choices must be feasible, we have

$$G(x^*) \leq c \quad \text{and} \quad G(x^{*'}) \leq c'. \tag{7.3}$$

We will show that the output vector $\alpha x^* + (1 - \alpha)x^{*'}$ is feasible under the input supply vector $\alpha c + (1 - \alpha)c'$. And that it yields revenue at least as high as $\alpha V(c) + (1 - \alpha)V(c')$.

(i) $\alpha x^* + (1 - \alpha)x^{*'}$ is feasible since for each i , the convexity of G^i implies

$$G^i(\alpha x^* + (1 - \alpha)x^{*'}) \underbrace{\leq}_{\text{convexity}} \alpha G^i(x^*) + (1 - \alpha)G^i(x^{*'}) \underbrace{\leq}_{(7.3)} \alpha c_i + (1 - \alpha)c'_i.$$

(ii) $\alpha x^* + (1 - \alpha)x^{*'}$ yields revenue at least as high as $\alpha V(c) + (1 - \alpha)V(c')$ since the concavity of F implies

$$F(\alpha x^* + (1 - \alpha)x^{*'}) \underset{\text{concavity}}{\geq} \alpha F(x^*) + (1 - \alpha)F(x^{*'}) = \alpha V(c) + (1 - \alpha)V(c'). \quad (7.4)$$

Therefore, we have found a feasible output vector that generates the target revenue. The maximum revenue must be no smaller than the revenue generated from the feasible output vector:

$$V(\alpha c + (1 - \alpha)c') \geq F(\alpha x^* + (1 - \alpha)x^{*'}). \quad (7.5)$$

Combining (7.4) and (7.5), we have

$$V(\alpha c + (1 - \alpha)c') \geq \alpha V(c) + (1 - \alpha)V(c').$$

This is the result we want to prove.

The economics behind this result is that the convexity of G rules out economies of scale or specialization in production, ensuring that a weighted average of outputs can be produced using the same weighted average of inputs. Then, the concavity of F ensures that the resulting revenue is at least as high as the same weighted average of the separate revenues.

Recall that we have established an alternative interpretation of a concave function in Chapter 6:

Claim. f is a concave function if and only if $\mathcal{F} = \{(x, v) | v \leq f(x)\}$ is a convex set.

In our current context, as $V(c)$ is a concave function, the set $\{(c, v) | v \leq V(c)\}$ is a convex set. This is an $(m + 1)$ -dimensional set, the collection of all points (c, v) such that $v \leq V(c)$. That is, revenue of v can be produced using the input vector c .

Figure 7.2 shows this set as the shaded area \mathcal{A} when c is a scalar. Since V is non-decreasing and concave, the set has a frontier that shows a positive but diminishing marginal product of the input in producing revenue. Note that the frontier of V may not be smooth. We will explain this point in the part [Generalized Marginal Products](#).

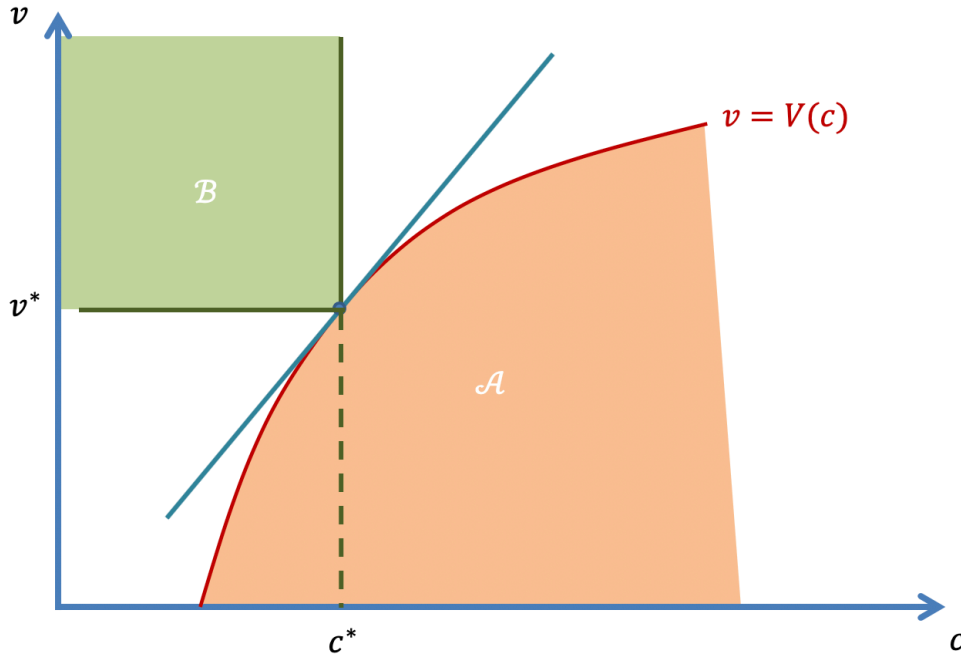


Figure 7.2: Value Function in Concave Programming

Since \mathcal{A} is a convex set, it can be separated from other convex sets. To do this, choose a point (c^*, v^*) in \mathcal{A} such that $v^* = V(c^*)$. (c^*, v^*) must be a boundary point since for any $r > 0$, there exists $\varepsilon \in (0, r)$

- (i) $v^* - (r - \varepsilon) < v^* = V(c^*)$ implies that the point $(c^*, v^* - (r - \varepsilon))$ is in \mathcal{A} ;
- (ii) $v^* + (r - \varepsilon) > v^* = V(c^*)$ implies that the point $(c^*, v^* + (r - \varepsilon))$ is not in \mathcal{A} .

Now, define \mathcal{B} as the set of all points (c, v) such that $c \leq c^*$ and $v \geq v^*$.

Graphically, \mathcal{B} is shown in Figure 7.2 as the shaded green area. It is clear from the graph that \mathcal{B} is a convex set and does not share interior points with \mathcal{A} . Formally,

- (i) **Convexity of \mathcal{B} .** For any two points $(c, v), (c', v') \in \mathcal{B}$, that is, $(c, v), (c', v')$ satisfying $c \leq c^*, v \geq v^*$ and $c' \leq c^*, v' \geq v^*$, and any real number $\alpha \in [0, 1]$, we have $\alpha c + (1 - \alpha)c' \leq \alpha c^* + (1 - \alpha)c^* = c^*$; $\alpha v + (1 - \alpha)v' \geq \alpha v^* + (1 - \alpha)v^* = v^*$, that is, $(\alpha c + (1 - \alpha)c', \alpha v + (1 - \alpha)v') \in \mathcal{B}$.
- (ii) **No Common Interior.** Points in \mathcal{A} satisfy $v \leq V(c)$. For points $(c, v) \in \mathcal{B}$,

$$v \geq v^* = V(c^*) \underset{V(c) \text{ is non-decreasing.}}{\geq} V(c) \implies v \geq V(c).$$

Therefore, \mathcal{A} and \mathcal{B} do not have interior points in common.

We could apply the Separation Theorem. Recognizing that (c^*, v^*) is a common boundary point of \mathcal{A} and \mathcal{B} , we could write the equation of the separating hyperplane as follows:

$$\iota v - \lambda c = b = \iota v^* - \lambda c^*,$$

where ι is a scalar, and λ is a m -dimensional row vector. The signs are so chosen that

$$\iota v - \lambda c \begin{cases} \leq b & \text{for all } (c, v) \in \mathcal{A} \\ \geq b & \text{for all } (c, v) \in \mathcal{B}. \end{cases} \quad (7.6)$$

The parameters and signs are deliberately chosen. As it will become clear later, ι and λ are both non-negative, and are linked to the shadow prices.

Remark. ι and λ must both be non-negative.

- (i) $\iota \geq 0$: Suppose $\iota < 0$. Consider the point $(c^*, v^* + 1) \in \mathcal{B}$. However, $\iota(v^* + 1) - \lambda c^* = b + \iota < b$, contradicting with (7.6).
- (ii) $\lambda_i \geq 0$ for $i = 1, 2, \dots, m$: Suppose $\lambda_i < 0$. Consider the point $(c^* - e^i, v^*)$, where e^i is a vector with its i^{th} component equal to 1 and all other components 0. $(c^* - e^i, v^*) \in \mathcal{B}$. However, $\iota v^* - \lambda(c^* - e^i) = b + \lambda_i < b$, contradicting with (7.6).

Now comes the more subtle question:

Question. Can ι be zero?

Let's see the consequence of $\iota = 0$.

- (i) For the equation of the hyperplane $\iota v - \lambda c = b$ to be meaningful, the combined vector (ι, λ) must be non-zero. Therefore, $\lambda_i \neq 0$ for at least one i . Given that $\lambda_i \geq 0$ for all i , it means $\lambda_i > 0$ for at least one i .
- (ii) The equation of the hyperplane becomes $-\lambda c = b = -\lambda c^*$. For all $(c, v) \in \mathcal{A}$, we have $-\lambda c \leq -\lambda c^*$, or $\lambda(c - c^*) \geq 0$.

In the **scalar constraint case**, in such a situation, we have $\lambda > 0$. Therefore, $\lambda(c - c^*) \geq 0$ implies $c - c^* \geq 0$. Graphically, the separating line is vertical at c^* , and the set \mathcal{A} lies entirely to the right of it. Put it differently, there are no feasible points to the left of c^* : the production is impossible if input supply falls short of this level. In some applications, this can happen because of indivisibilities.

Figure 7.3 shows two ways in which this can happen.

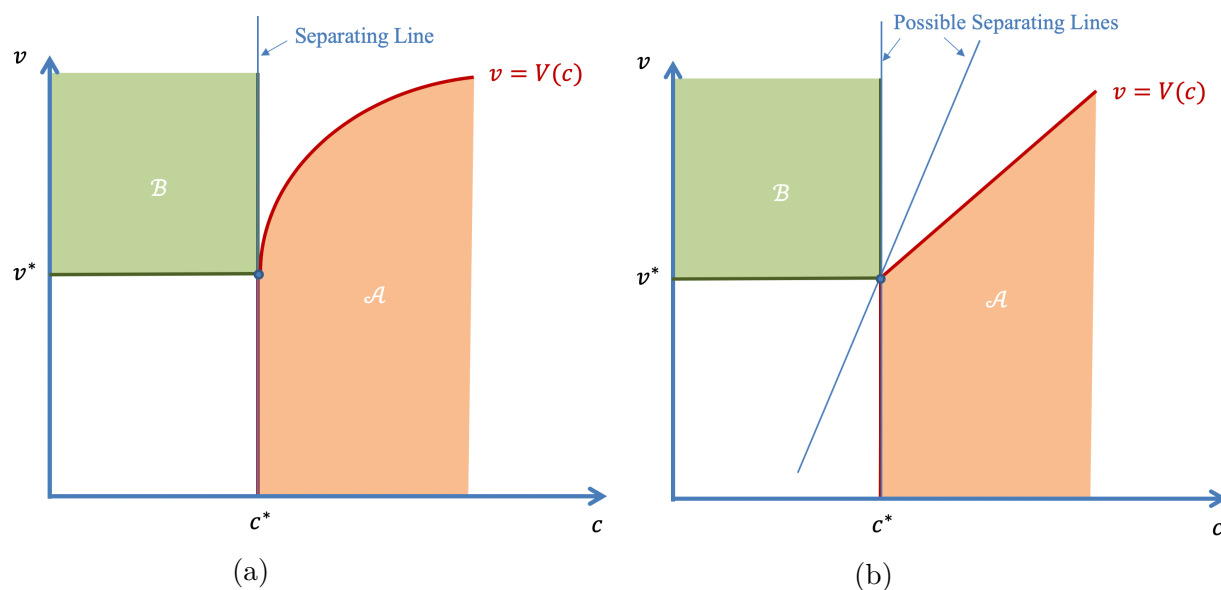


Figure 7.3: Failure of Constraint Qualification

The two cases differ in the behavior of $V(c)$ as c approaches c^* .

- (i) In case 7.3a, the marginal revenue product of the resource goes to infinity as c approaches c^* . And only a vertical separating line exists.
- (ii) In case 7.3b, the marginal revenue product is finite. And except for the vertical separating line, there also exist separating lines with finite slopes (positive ι).

We would like to ensure a positive ι so that the marginal revenue product of a resource is finite. And we do this by ensuring the existence of c such that $c < c^*$. Due to the existence of case (ii) above, such conditions are only sufficient but not necessary.

Claim. *If there exists an x^o such that $G(x^o) \ll c^*$ and $F(x^o)$ is defined, then $\iota > 0$.*

This requirement is the *constraint qualification* for the concave programming problem. It is sometimes called the *Slater condition*.

Intuitively, for a scalar c , such a condition works since by construction, $(G(x^o), F(x^o)) \in \mathcal{A}$ and $(G(x^o), F(x^o))$ is a point to the left of c^* . Thus, the separating line cannot have an infinite slope at c^* .

We will next prove that the Slater condition implies $\iota > 0$ in general.

Proof. We prove by contradiction. Suppose that the condition holds but $\iota = 0$.

Then, on one hand, $\lambda_i \geq 0$ for all i and $\lambda_i > 0$ for at least one i .

Therefore, by $G(x^o) \ll c^* \iff G^i(x^o) < c_i^*$, we have

$$\implies \lambda(G(x^o) - c^*) = \sum_{i=1}^m \lambda_i(G^i(x^o) - c_i^*) < 0. \quad (7.7)$$

On the other hand, $(G(x^o), F(x^o)) \in \mathcal{A}$ since revenue of $F(x^o)$ can be generated using the input vector $G(x^o)$. Therefore, by the separation property,

$$\begin{aligned} -\lambda G(x^o) &\stackrel{\underbrace{=}_{\iota=0}}{=} \iota F(x^o) - \lambda G(x^o) \stackrel{\underbrace{\leq}_{\text{separation property}}}{\leq} \iota v^* - \lambda c^* \stackrel{\underbrace{=}_{\iota=0}}{=} -\lambda c^* \\ \implies \lambda(G(x^o) - c^*) &\geq 0. \end{aligned} \quad (7.8)$$

(7.7) and (7.8) contradict each other. This contradiction forces us to conclude that the initial supposition $\iota = 0$ must be wrong. \square

Normalization. The separation property (7.6) is unaffected if we multiply by b , ι and λ_i by the same positive number. Once we can be sure that $\iota \neq 0$, we can choose a scale to make $\iota = 1$. In economic terms, ι and λ constitute a system of shadow prices, ι for revenue and λ for the inputs. Only relative prices matter for economic decisions, in setting $\iota = 1$, we are choosing revenue to be the numéraire. We will adopt this normalization henceforth.

Shadow Price Interpretation of λ . Observe that by the separation property (7.6), for all $(c, v) \in \mathcal{A}$,

$$v - \lambda c \leq v^* - \lambda c^*.$$

That is, (c^*, v^*) achieves the maximum value of $(v - \lambda c)$ among all points $(c, v) \in \mathcal{A}$. If we interpret λ as the vector of shadow prices of inputs, then $(v - \lambda c)$ is the profit that accrues when a producer uses inputs c to produce revenue v . Since all points in \mathcal{A} represents feasible production plans, the result says that a profit-maximizing producer will pick (c^*, v^*) . This means that the producer need not be aware that in fact the availability of inputs is limited to c^* . He may think that he is free to choose any c but ends up choosing the right c^* . It is the prices λ that brings home to him the scarcity.

The principle behind this interpretation is general and important: *constrained choice can be converted into unconstrained choice if the proper scarcity costs or shadow values of the constraints are netted out of the criterion function.* As it will become clear later, this is the most important feature of Lagrange's Method in concave programming.

Generalized Marginal Products. For any c , the point $(c, V(c))$ is in \mathcal{A} . So by the separation property, we have

$$\begin{aligned} V(c) - \lambda c &\leq V(c^*) - \lambda c^*, \\ \text{or } V(c) - V(c^*) &\leq \lambda(c - c^*). \end{aligned} \quad (7.9)$$

This looks very much like the concavity property (7.1). If $V(c)$ is differentiable, then by Proposition 7.A.1, concavity of $V(c)$ means

$$V(c) - V(c^*) \leq V_c(c^*)(c - c^*). \quad (7.10)$$

(7.9) and (7.10) suggest $\lambda = V_c(c^*)$, and confirm our interpretation of λ as shadow prices. However, the problem is that V may not be differentiable. Let us consider a general point $(c, V(c))$ with its associated multiplier vector λ .¹ Compare this with a neighboring point where only the i^{th} input is increase: $(c + he^i, V(c + he^i))$, where h is a positive scalar and e^i is a vector with its i^{th} component equal to 1 and all others 0. Then, (7.9) becomes

$$\begin{aligned} V(c + he^i) - V(c) &\leq \lambda he^i = h\lambda_i \\ \implies \frac{[V(c + he^i) - V(c)]}{h} &\leq \lambda_i. \end{aligned} \quad (7.11)$$

We will show that by the concavity of V , the left-hand side of (7.11) is a non-increasing function of h . To see this, consider two points $(c + he^i, V(c + he^i))$ and $(c + \alpha he^i, V(c + \alpha he^i))$ for some $h > 0$ and $\alpha \in (0, 1)$. Then by concavity of V ,

$$\begin{aligned} V(c + \alpha he^i) &\geq \alpha V(c + he^i) + (1 - \alpha)V(c) \\ \implies V(c + \alpha he^i) - V(c) &\geq \alpha [V(c + he^i) - V(c)] \\ \implies \frac{V(c + \alpha he^i) - V(c)}{\alpha h} &\geq \frac{V(c + he^i) - V(c)}{h} \end{aligned} \quad (7.12)$$

¹The asterisks, having served the purpose of distinguishing a particular point in the (c, v) space for separation, will be dropped from now on.

Since $\alpha h < h$, (7.12) implies that the left-hand side of (7.11), namely, $\frac{V(c+he^i)-V(c)}{h}$ is non-increasing in h .

Graphically, it is simply the slope of the chord. See Figure 7.4. It is not hard to see that the slope is larger when h decreases.

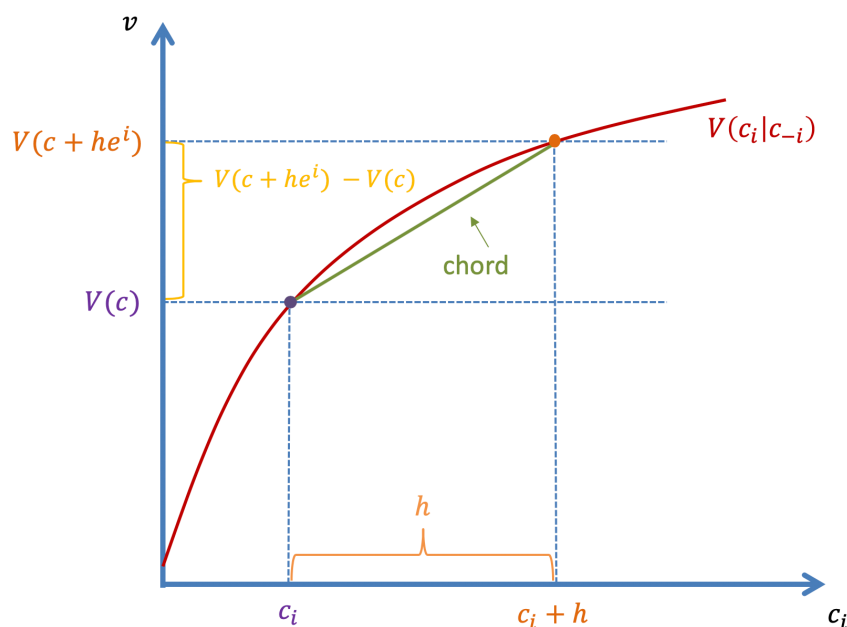


Figure 7.4: Convavity of $V(c)$

Therefore, the left-hand side of (7.11) must attain the maximum as h goes to zero from positive values. This limit is defined as the “rightward” partial derivative of V with respect to the i^{th} coordinate of c : $V_i^+(c)$. Therefore, (7.11) implies $V_i^+(c) \leq \lambda_i$.

Similarly, we could repeat the analysis for $h < 0$. Now, (7.9) implies

$$\begin{aligned} V(c + he^i) - V(c) &\leq \lambda he^i = h\lambda_i \\ \implies \frac{[V(c + he^i) - V(c)]}{h} &\geq \lambda_i. \end{aligned} \quad (7.13)$$

Taking the limit from the negative values of h gives the “leftward” partial derivative $V_i^-(c)$. This proves $V_i^-(c) \geq \lambda_i$.

Combining the two, we have

$$V_i^-(c) \geq \lambda_i \geq V_i^+(c). \quad (7.14)$$

This result generalizes the notion of diminishing marginal returns and relates the multipliers to these generalized marginal products.

Figure 7.5 illustrates this for the case of a scalar c .

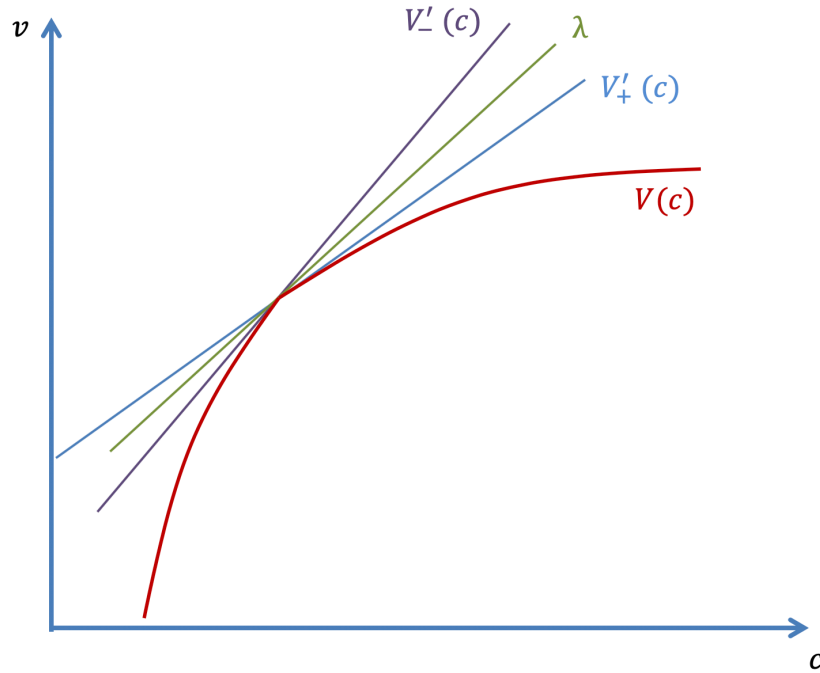


Figure 7.5: Generalized Marginal Products

Choice Variables. So far the vector of choice variables x has been kept in the background. Let's now consider it explicitly. The point $(G(x^*), F(x^*)) \in \mathcal{A}$ since revenue of $F(x^*)$ can be generated using the input vector $G(x^*)$. The separation property gives

$$F(x^*) - \lambda G(x^*) \leq V(c) - \lambda c \underset{F(x^*)=V(c)}{\implies} \lambda [c - G(x^*)] \leq 0.$$

That is, $\sum_{i=1}^m \lambda_i [c_i - G^i(x^*)] \leq 0$. Since $\lambda_i \geq 0$ and $G^i(x) \leq c_i$ for all i , we have $\lambda_i [c_i - G^i(x^*)] \geq 0$ for all i . Therefore,

$$\lambda_i [c_i - G^i(x^*)] = 0. \quad (7.15)$$

This is just the notion of *complementary slackness* we have learned before.

Finally, for any x , the point $(G(x), F(x)) \in \mathcal{A}$ since revenue of $F(x)$ can be generated using the input vector $G(x)$. Recognizing (7.15), the separation property gives

$$F(x) - \lambda G(x) \underset{\text{separation property}}{\leq} V(c) - \lambda c \underset{F(x^*)=V(c) \text{ and (7.15)}}{=} F(x^*) - \lambda G(x^*)$$

for all x .

That is, x^* maximizes $F(x) - \lambda G(x)$ without any constraints. This means that the shadow prices allow us to convert the original constrained revenue-maximization problem into an unconstrained profit-maximization problem.

All of the above reasoning can now be summarized into the basic theorem of this section:

Theorem 7.1 (Necessary Conditions for Concave Programming). *Suppose that F is a concave function and G is a vector convex function, and that there exists an x^o satisfying $G(x^o) \ll c$. If x^* maximizes $F(x)$ subject to $G(x) \leq c$, then there is a row vector λ such that*

- (i) x^* maximizes $F(x) - \lambda G(x)$ without any constraints, and
- (ii) $\lambda \geq 0$, $G(x^*) \leq c$ with complementary slackness.

Note that Theorem 7.1 does not require F and G to have derivatives. But if the functions are differentiable, then we have the first-order necessary conditions for the maximization problem (i):

$$F_x(x^*) - \lambda G_x(x^*) = 0. \tag{7.16}$$

In terms of the Lagrangian $\mathcal{L}(x, \lambda)$, (7.16) becomes $\mathcal{L}_x(x^*, \lambda) = 0$. This is just the condition of Lagrange's Theorem with Inequality Constraints. We could further add the non-negativity constraints on x , and get Kuhn-Tucker Theorem.²

There is one respect in which concave programming goes beyond the general Lagrange or Kuhn-Tucker conditions. The first-order necessary conditions (7.16) are not sufficient to ensure maximum. In general, there was no claim that x^* maximized the Lagrangian. However, when F is concave and G is convex, part (i) of Theorem 7.1 is easily transformed into $\mathcal{L}(x, \lambda) \leq \mathcal{L}(x^*, \lambda)$ for all x , so x^* does maximize the Lagrangian. Therefore, our interpretation of Lagrange's method as converting the constrained revenue-maximization into unconstrained profit-maximization must be confined to the case of concave programming.

²One way do to this is to recognize $x \geq 0$ or $-x \leq 0$, as another n inequality constraints, associate with them an n -dimensional multiplier μ and repeat the previous analysis.

Sufficiency. The first-order necessary conditions are *sufficient* to yield a true maximum in the concave programming problem. The argument proceeds in two parts.

1. Suppose x^* satisfies (i) and (ii) in Theorem 7.1. Then, for any feasible x , we have

$$\begin{aligned}
 F(x^*) - \lambda G(x^*) &\underset{\text{(i)}}{\geq} F(x) - \lambda G(x) \\
 \implies F(x^*) - \lambda c &\underset{\text{(ii) complementary slackness: } \lambda[c-G(x^*)]=0}{\geq} F(x) - \lambda G(x) \\
 \implies F(x^*) &\geq F(x) + \lambda[c - G(x)] \underset{x \text{ is feasible: } G(x) \leq c}{\geq} F(x).
 \end{aligned}$$

Thus, x^* maximizes $F(x)$ subject to $G(x) \leq c$.

2. Suppose x^* satisfies the first-order condition (7.16). Since F is concave, G is convex, and $\lambda \geq 0$, then $F - \lambda G$ is concave. Then, by Proposition 7.A.1,

$$[F(x) - \lambda G(x)] - [F(x^*) - \lambda G(x^*)] \underset{\text{Proposition 7.A.1: Concavity}}{\leq} [F_x(x) - \lambda G_x(x)](x - x^*) \underset{(7.16)}{=} 0.$$

Therefore,

$$F(x) - \lambda G(x) \leq F(x^*) - \lambda G(x^*),$$

or x^* maximizes $F(x) - \lambda G(x)$ without any constraints.

This result is summarized into the theorem below:

Theorem 7.2 (Sufficient Conditions for Concave Programming). *If x^* and λ are such that*

- (i) x^* maximizes $F(x) - \lambda G(x)$ without any constraints, and
- (ii) $\lambda \geq 0$, $G(x^*) \leq c$ with complementary slackness,

then x^ maximizes $F(x)$ subject to $G(x) \leq c$. If $F - \lambda G$ is concave (for which in turn it suffices to have F concave and G convex), then*

$$F_x(x^*) - \lambda G_x(x^*) = 0 \tag{7.16}$$

implies (i) above.

Note that no constraint qualification appears in the sufficient conditions.

7.C. Quasi-concave Programming

In the separation approach of Chapter 6, F was merely quasi-concave and each component constraint function in G was quasi-convex. In this chapter, the stronger assumption of concavity and convexity has been made so far. In fact, the weaker assumptions of quasi-concavity and quasi-convexity make little difference to the necessary conditions. They yield sufficient conditions like the ones above for concave programming, but only in the presence of some further technical conditions that are quite complex to establish. For interested students, please refer to the paper “Arrow and Enthoven (1961). Quasi-concave Programming. *Econometrica*, 779-800.”

We will discuss only a limited version of quasi-concave programming, namely, the one where the objective function is quasi-concave and the constraint function is linear:³

$$\begin{aligned} \max_x F(x) & \qquad \qquad \qquad \text{(MP1)} \\ \text{s.t. } px & \leq b, \end{aligned}$$

where p is a row vector and b is a number.

Recall the definition of Quasiconcavity:

Definition 6.B.3 (Quasi-concave Function). *A function $f : \mathcal{S} \rightarrow \mathbb{R}$, defined on a convex set $\mathcal{S} \subset \mathbb{R}^N$, quasi-concave if the set $\{x | f(x) \geq c\}$ is convex for all $c \in \mathbb{R}$, or equivalently, if $f(\alpha x^a + (1 - \alpha)x^b) \geq \min\{f(x^a), f(x^b)\}$, for all x^a, x^b and for all $\alpha \in [0, 1]$.*

First, we need to establish some property of quasi-concave function, relating to the derivatives. For a quasi-concave objective function F , suppose $F(x^b) \geq F(x^a)$. Then,

$$F((1 - \alpha)x^a + \alpha x^b) \geq F(x^a), \tag{7.17}$$

for all $\alpha \in [0, 1]$. Let $h(\alpha) = F((1 - \alpha)x^a + \alpha x^b) = F(x^a + \alpha(x^b - x^a))$. Then, (7.17) becomes

$$h(\alpha) \geq h(0) \implies \frac{h(\alpha) - h(0)}{\alpha} \geq 0. \tag{7.18}$$

³The mirror-image case of a linear objective and a quasi-convex constraint can be treated in the same way.

By the definition of derivative,

$$\lim_{\alpha \rightarrow 0} \left[\frac{h(\alpha) - h(0)}{\alpha} \right] = h'(0).$$

Since (7.18) holds when $\alpha \rightarrow 0$, we have

$$h'(0) \geq 0. \tag{7.19}$$

On the other hand, by chain rule,

$$\begin{aligned} h'(\alpha) &= F_x(x^a + \alpha(x^b - x^a))(x^b - x^a) \\ \implies h'(0) &= F_x(x^a)(x^b - x^a) \end{aligned} \tag{7.20}$$

(7.19) and (7.20), we have

$$F_x(x^a)(x^b - x^a) \geq 0. \tag{7.21}$$

This holds for all x^a, x^b such that $F(x^b) \geq F(x^a)$.

Now consider the maximization problem (MP1). The first-order necessary conditions are

$$F_x(x^*) - \lambda p = 0 \tag{7.22}$$

$$px^* \leq b \text{ and } \lambda \geq 0, \text{ with complementary slackness}$$

We claim that (7.22) is also sufficient when $\lambda > 0$ and the constraint is binding.⁴ Formally,

Claim. *If F is continuous and quasi-concave, x^* and $\lambda > 0$ satisfy the first-order necessary conditions, then x^* solves the quasi-concave programming problem.*

Proof. Suppose that there exists x such that $F(x) > F(x^*) \equiv v^*$. We will show that x is not feasible, that is, $px > b$.

By (7.21), $F(x) > F(x^*)$ implies

$$F_x(x^*)(x - x^*) \geq 0. \tag{7.23}$$

⁴Appendix B provides an example of a spurious stationary point where (7.22) holds with $\lambda = 0$.

Substituting (7.22) into (7.23), we have

$$\lambda p(x - x^*) \geq 0 \underset{\lambda > 0}{\implies} p(x - x^*) \geq 0 \quad \text{or} \quad px \geq px^* \underset{\text{constraint binding}}{=} b.$$

In other words, the upper contour set of $F(x)$ for the value v^* is contained in the half-space on or above the constraint line. To complete the proof, we need to rule out $px = b$.

Since F is continuous and $F(x) > F(x^*)$, x is an interior point of the upper contour set of $F(x)$ for the value v^* .⁵ Then if $px = b$, there exists y such that $F(y) > F(x^*)$ and $py < b$, leading to a contradiction with the previous result that $F(x) > F(x^*)$ implies $px \geq b$.⁶ Therefore, we must have $px > b$. This completes the proof. \square

Figure 7.6 illustrate the quasi-concave programming problem with a linear constraint.

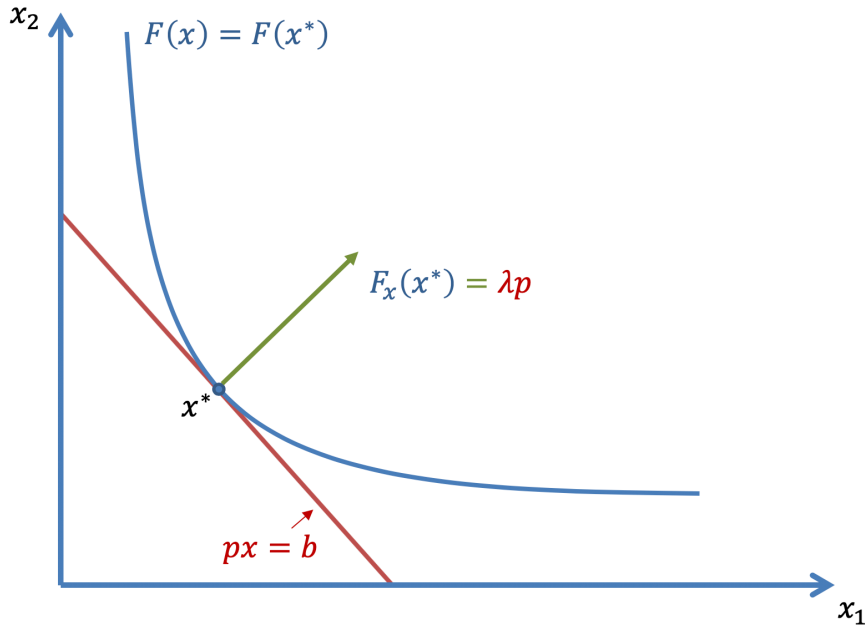


Figure 7.6: Quasi-concave Objective and Linear Constraint

$F_x(x^*)$ is normal to the contour of $F(x)$ at x^* . p is normal to the constraint $px = b$ at x^* . The usual tangency condition is equivalent to the normal vectors being parallel.

Equation (7.22) expresses this, with the constant of proportionality equal to λ .

⁵Continuity of $F(x)$ means that for any $\varepsilon > 0$, there exists $\delta > 0$, such that for all y satisfying $\|y - x\| < \delta$, we have $\|F(y) - F(x)\| < \varepsilon$. We choose $\varepsilon \in (0, F(x) - F(x^*))$. Then, by continuity of $F(x)$, we can find δ such that for all y satisfying $\|y - x\| < \delta$, we have $F(y) \in (F(x) - \varepsilon, F(x) + \varepsilon)$. Therefore, $F(y) > F(x) - \varepsilon > F(x^*)$. Thus, x is an interior point of $F(x) > F(x^*)$.

⁶The detailed argument goes as follows. Consider δ as defined in footnote 5. By footnote 5, for all y satisfying $\|y - x\| < \delta$, we have $F(y) > F(x^*)$. Since x is not an interior point of $px \geq b$, then there exists y such that $\|y - x\| < \delta$ and $py < b$. Therefore, we can find some y such that $F(y) > F(x^*)$ but $py < b$.

7.D. Uniqueness

The above sufficient conditions for concave as well as quasi-concave programming are *weak* in the sense that they establish that no other feasible choice x can do better than x^* . They do not rule out the existence of other feasible choices that yield $F(x) = F(x^*)$. In other words, they do not establish the uniqueness of the optimum.

As discussed in Chapter 6, a strengthening of the concept of concavity or quasi-concavity gives uniqueness.

Definition 7.D.1 (Strictly Concave Function). *A function $f : \mathcal{S} \rightarrow \mathbb{R}$, defined on a convex set $\mathcal{S} \subset \mathbb{R}^N$, is strictly concave if*

$$f(\alpha x^a + (1 - \alpha)x^b) > \alpha f(x^a) + (1 - \alpha)f(x^b), \quad (7.24)$$

for all $x^a, x^b \in \mathcal{S}$ and for all $\alpha \in (0, 1)$.

Claim. *If the objective function F in the concave programming problem is strictly concave, then the maximizer x^* is unique.*

Proof. We prove by contradiction. Suppose that $x^{*'}$ is another solution. Then, $F(x^*) = F(x^{*'}) = v^*$, and $G(x^*) \leq c, G(x^{*'}) \leq c$. Now consider $\alpha x^* + (1 - \alpha)x^{*'}$.

(i) $\alpha x^* + (1 - \alpha)x^{*'}$ is feasible since for each i , the convexity of G^i implies

$$G^i(\alpha x^* + (1 - \alpha)x^{*'}) \underbrace{\leq}_{\text{convexity}} \alpha G^i(x^*) + (1 - \alpha)G^i(x^{*'}) \underbrace{\leq}_{\text{feasibility of } x^* \text{ and } x^{*'}} \alpha c_i + (1 - \alpha)c_i = c_i.$$

(ii) $\alpha x^* + (1 - \alpha)x^{*'}$ yields higher value than v^* since the strict concavity of F implies

$$F(\alpha x^* + (1 - \alpha)x^{*'}) \underbrace{>}_{\text{strict concavity}} \alpha F(x^*) + (1 - \alpha)F(x^{*'}) = \alpha v^* + (1 - \alpha)v^* = v^*.$$

Therefore, we have found a feasible choice $\alpha x^* + (1 - \alpha)x^{*'}$ which yields higher value than v^* . This contradicts with the fact the x^* and $x^{*'}$ are optimal. Therefore, the initial supposition must be wrong and strict concavity of F implies the uniqueness of the maximizer.

7.E. Examples

Example 7.1: Linear Programming.

An important special case of concave programming is the theory of *linear programming*.

Here the objective and constraint functions are linear:

$$\begin{aligned} \max_x F(x) &\equiv ax && \text{(Primal)} \\ \text{s.t. } G(x) &\equiv Bx \leq c \text{ and } x \geq 0, \end{aligned}$$

where a is an n -dimensional row vector and B an m -by- n matrix. Now

$$F_x(x) = a \text{ and } G_x(x) = B.$$

When the constraint functions are linear, no constraint qualification is needed.⁷

All conditions of concave programming are fulfilled, and the Kuhn-Tucker conditions are both necessary and sufficient.

The Lagrangian is

$$\mathcal{L}(x, \lambda) = ax + \lambda[c - Bx]. \quad (7.25)$$

The optimum x^* and λ^* ⁸ satisfy the Kuhn-Tucker conditions:

$$a - \lambda^* B \leq 0, \quad x^* \geq 0, \quad \text{with complementary slackness,} \quad (7.26)$$

$$c - Bx^* \geq 0, \quad \lambda^* \geq 0, \quad \text{with complementary slackness.} \quad (7.27)$$

(7.26) and (7.27) contain 2^{m+n} combinations of patterns of equations and inequalities.

As a special feature of the linear programming problem, if k of the constraints in (7.27) hold with equality, then exactly $(n - k)$ non-negativity constraints in (7.26) should bind.

When this is the case, the corresponding equations for λ is also of the correct number m .

⁷Loosely, *Constraint Qualification* is to ensure that the first order approximation works. If *Constraint Qualification* fails, there will be a discrepancy between the original problem and the linearly approximated one. When the constraints are already linear, there is no need to find linear approximations to them, so the issue does not arise.

⁸In this problem, we will have occasion to consider λ as variables. Therefore, the optimal value is denoted as λ^* .

Next, consider a new linear programming problem:

$$\begin{aligned} \max_y & -yc && \text{(Dual)} \\ \text{s.t.} & -yB \leq -a \text{ and } y \geq 0, \end{aligned}$$

where y is a m -dimensional row vector and the vectors a , c and the matrix B are exactly as before.

We introduce a column vector μ of multipliers and define the Lagrangian:

$$\mathcal{L}(x, \lambda) = -yc + [-a + yB]\mu. \tag{7.28}$$

The optimum y^* and μ^* satisfy the necessary and sufficient Kuhn-Tucker conditions:

$$-c + B\mu^* \leq 0, \quad y^* \geq 0, \quad \text{with complementary slackness,} \tag{7.29}$$

$$-a + y^*B \geq 0, \quad \mu^* \geq 0, \quad \text{with complementary slackness.} \tag{7.30}$$

(7.29) is exactly the same as (7.27) and (7.30) is exactly the same as (7.26), if we replace y^* by λ^* and μ^* by x^* . In other words, the optimum x^* and λ^* solve the new problem.

The new problem is said to be *dual* to the original, which is then called the *primal* problem in the pair. This captures an important economic relationship between prices and quantities in economics.

We interpret the primal problem as follows:

$$\begin{aligned} \max_x & \underbrace{a}_{\text{output prices}} \underbrace{x}_{\text{output quantities}} \\ \text{s.t.} & \underbrace{Bx}_{\text{inputs for producing } x} \leq \underbrace{c}_{\text{input supplies}} \text{ and } x \geq 0, \end{aligned}$$

Solving the primal problem, we get x^* and λ^* . λ^* is the vector of shadow prices of the inputs. Rewriting the dual problem in terms of λ , we know from the previous analysis that λ^* solves the dual problem.

$$\lambda^* = \min_{\lambda} \{\lambda c \mid \lambda B \geq a \text{ and } \lambda \geq 0\}$$

Thus, the shadow prices minimize the cost of the input c .

Note that the j^{th} component of λB is $\sum_i \lambda_i B_{ij}$, which is the cost of the bundle of inputs needed to produce one unit of good j , calculated using the shadow prices. The constraint $\sum_i \lambda_i B_{ij} \geq a_j$ means that the input cost of good j is at least as great as the unit value of output of good j . This is true for all good j . In other words, the shadow prices of inputs ensure that no good can make a strictly positive profit – a standard “competitive” condition in economics.

Complementary slackness in (7.26) ensures that

- (i) If the unit cost of production of j , $\sum_i \lambda_i B_{ij}$, exceeds its price a_j , then $x_j = 0$. That is, if the production of j would entail a loss when calculated using the shadow prices, then good j would not be produced.
- (ii) If good j is produced in positive quantity, $x_j > 0$, then the unit cost exactly equals the price, $\sum_i \lambda_i B_{ij} = a_j$. That is, the profit is exactly 0.

This can be summarized by observing that complementary slackness in (7.26) and (7.27) imply

$$\begin{aligned} [a - \lambda^* B]x^* = 0 &\implies ax^* = \lambda^* Bx^* \\ \lambda^*[c - Bx^*] = 0 &\implies \lambda^* c = \lambda^* Bx^* \end{aligned}$$

Combining the two, we have

$$ax^* = \lambda^* c \tag{7.31}$$

This says that the value of the optimum output equals the cost of the factor supplies evaluated at the shadow prices. This result can be interpreted as the familiar circular flow of income, that is, national product equals national income.

Finally, it is easy to check that if we take the dual problem as our starting-point and go through the mechanical steps to finding its dual, we return to the primal. In other words, duality is reflexive.

This is the essence of the duality theory of linear programming. One final remark is that we took the optimum x^* as our starting point, however, the solution may not exist, because the constraints may be mutually inconsistent, or they may define an unbounded feasible set. This issue beyond our discussion here and is left to more advanced texts.

Example 7.2: Failure of Profit-maximizing.

For a scalar x , consider the following maximization problem:

$$\begin{aligned} \max_x F(x) &\equiv e^x \\ \text{s. t. } G(x) &\equiv x \leq 1. \end{aligned}$$

$F(x)$ is increasing, and the maximum occurs at $x = 1$. See Figure 7.7 below.

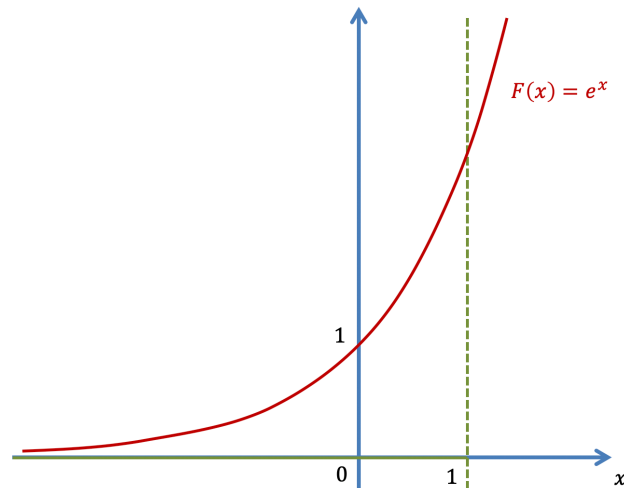


Figure 7.7: Convex F

Kuhn-Tucker Theorem applies. The Lagrangian is

$$\mathcal{L}(x, \lambda) = e^x + \lambda(1 - x).$$

Kuhn-Tucker necessary conditions are

$$\partial \mathcal{L}(x, \lambda) / \partial x = e^x - \lambda = 0;$$

$$\partial \mathcal{L}(x, \lambda) / \partial \lambda = 1 - x \geq 0 \text{ and } \lambda \geq 0, \text{ with complementary slackness.}$$

The solution is

$$\begin{cases} x^* = 1 \\ \lambda = e. \end{cases}$$

However, $x = 1$ does not maximize $F(x) - \lambda G(x)$ without constraints. In fact, $e^x - ex$ can be made arbitrarily large by increasing x beyond 1. Here, Lagrange's method does not convert the original constrained maximization problem into an unconstrained profit-maximization problem. The difficulty is that F is not concave.

Appendix A

Proposition 7.A.1 (Concave Function). *A differentiable function $f : \mathcal{S} \rightarrow \mathbb{R}$, defined on a convex set $\mathcal{S} \subset \mathbb{R}^N$, is concave if and only if*

$$f_x(x^a)(x^b - x^a) \geq f(x^b) - f(x^a), \quad (7.1)$$

for all $x^a, x^b \in \mathcal{S}$.

Proof. First, from Definition 6.B.5, for all $x^a, x^b \in \mathcal{S}$ and all $\alpha \in [0, 1]$, we have

$$\begin{aligned} f((1 - \alpha)x^a + \alpha x^b) &\geq (1 - \alpha)f(x^a) + \alpha f(x^b) \\ \implies f(x^a + \alpha(x^b - x^a)) - f(x^a) &\geq f(x^a) + \alpha(f(x^b) - f(x^a)) - f(x^a) \\ \implies \frac{f(x^a + \alpha(x^b - x^a)) - f(x^a)}{\alpha} &\geq f(x^b) - f(x^a) \end{aligned} \quad (7.32)$$

Let $h(\alpha) = f(x^a + \alpha(x^b - x^a))$. Then, the left-hand side of (7.32) becomes

$$\frac{h(\alpha) - h(0)}{\alpha}.$$

By the definition of derivative,

$$\lim_{\alpha \rightarrow 0} \left[\frac{h(\alpha) - h(0)}{\alpha} \right] = h'(0).$$

Since (7.32) holds when $\alpha \rightarrow 0$, we have

$$h'(0) \geq f(x^b) - f(x^a). \quad (7.33)$$

On the other hand, by chain rule,

$$\begin{aligned} h'(\alpha) &= f_x(x^a + \alpha(x^b - x^a))(x^b - x^a) \\ \implies h'(0) &= f_x(x^a)(x^b - x^a) \end{aligned} \quad (7.34)$$

(7.33) and (7.34), we have

$$f_x(x^a)(x^b - x^a) \geq f(x^b) - f(x^a),$$

which is (7.1).

Next, we look at the other direction. Let $x^a, x^b \in \mathcal{S}$, and define $x^c = (1 - \alpha)x^a + \alpha x^b$ for

some $\alpha \in [0, 1]$. Since \mathcal{S} is convex, $x^c \in \mathcal{S}$. Since (7.1) holds for all values in \mathcal{S} , we have

$$f_x(x^c)(x^a - x^c) \geq f(x^a) - f(x^c) \quad (7.35)$$

$$\text{and } f_x(x^c)(x^b - x^c) \geq f(x^b) - f(x^c) \quad (7.36)$$

Multiply (7.35) by $(1 - \alpha)$ and (7.36) by α , and adding the two, we have

$$\begin{aligned} & (1 - \alpha)f_x(x^c)(x^a - x^c) + \alpha f_x(x^c)(x^b - x^c) \geq (1 - \alpha)(f(x^a) - f(x^c)) + \alpha(f(x^b) - f(x^c)) \\ \implies & f_x(x^c) \left[(1 - \alpha)x^a + \alpha x^b - x^c \right] \geq (1 - \alpha)f(x^a) + \alpha f(x^b) - f(x^c) \\ \implies & 0 \geq (1 - \alpha)f(x^a) + \alpha f(x^b) - f((1 - \alpha)x^a + \alpha x^b) \\ \implies & f((1 - \alpha)x^a + \alpha x^b) \geq (1 - \alpha)f(x^a) + \alpha f(x^b), \end{aligned}$$

which corresponds to Definition 6.B.5. □

Appendix B

Consider the following maximization problem:

$$\begin{aligned} & \max_x (x - 1)^3 \\ & \text{s. t. } x - 2 \leq 0. \end{aligned}$$

By the first-order necessary conditions, we obtain two candidate solutions:

$$\left\{ \begin{array}{l} x = 1 \\ \lambda = 0 \end{array} \right. \quad \text{and} \quad \left\{ \begin{array}{l} x = 2 \\ \lambda = 3 \end{array} \right.$$

The first solution is a spurious stationary point, which is not a solution for maximum.