Chapter 4. Maximum Theorem, Implicit Function Theorem and Envelope Theorem*

This chapter will cover three key theorems: the maximum theorem (or the theorem of maximum), the implicit function theorem, and the envelope theorem. The maximum theorem studies the continuity of the optimizer and optimum, the implicit function theorem studies the differentiablity of the optimizer, and the envelope theorem studies the differentiablity of the optimum, all with respect to a group of parameters. The maximum theorem was first stated and proven by Claude Berge (1959; 1963, p. 116). Augustin-Louis Cauchy is credited with the first rigorous form of the implicit function theorem. As to the envelope theorem, the clues of inventors are not very clear; see Löfgren (2011) for a summary of its history.

In the previous two chapters, we are "**moving along** a curve", that is, the objective function f and the feasible set G are given (or the parameter value is given) and we are searching over $\mathbf{x}$ along f to find the optimal $\mathbf{x}^*$. On the contrary, we are "**shifting** a curve" in this chapter, i.e., we change the parameter values to shift f and G, and check how the optimizer and optimum respond to such shifting.

Related materials of this chapter can be found in Chapter 19.1, 19.2 and 19.4 of Simon and Blume (1994) and Chapter 9 of Sundaram (1996).

1 The Maximum Theorem

Often in economics we are not so much interested in what the solution to a particular maximization problem is but rather wish to know how the solution to a parameterized problem depends on the parameters. Thus in our example of utility maximization in Section 1.1 of Chapter 2 we might be interested not so much in what the solution to the maximization problem is when $p_1 = 2$, $p_2 = 7$, and $y = 25$, but rather in how the solution depends on p_1, p_2, and y. (That is, we might be interested in the demand function.) Sometimes we shall also be interested in how the maximized function depends on the parameters - in the example how the maximized utility depends on p_1, p_2, and y.

This raises a number of questions. In order for us to speak meaningfully of a demand *function* it should be the case that the maximization problem has a unique solution. Further, we would like to know if the “demand” function is continuous - or even if it is differentiable. Consider again the

*Email: pingyu@hku.hk

equality-constrained maximization problem, but this time let us explicitly add some parameters.

$$\begin{aligned} &\max_{x_1,\cdots,x_n} \; f(x_1,\cdots,x_n,a_1,\cdots,a_k) \\ &\text{s.t. } \; g_1(x_1,\cdots,x_n,a_1,\cdots,a_k) = c_1, \\ &\qquad\qquad \vdots \\ &\qquad g_m(x_1,\cdots,x_n,a_1,\cdots,a_k) = c_m. \end{aligned} \tag{1}$$

The uniqueness problem is studied in the last chapter. Now let $v(a_1,\cdots,a_k)$ be the maximized value of f when the parameters are $(a_1,\cdots,a_k)$. Let us suppose that the problem is such that the solution is unique and that $(x_1^*(a_1,\cdots,a_k),\cdots,x_n^*(a_1,\cdots,a_k))$ are the values that maximize the function f when the parameters are $\mathbf{a} \equiv (a_1,\cdots,a_k)'$; then

$$v(a_1,\cdots,a_k) = f(x_1^*(a_1,\cdots,a_k),\cdots,x_n^*(a_1,\cdots,a_k),a_1,\cdots,a_k).$$

(Notice however that the function v is uniquely defined even if there is not a unique maximizer.)

The Maximum Theorem gives conditions on the problem under which the function v and the functions $x_1^*,\cdots,x_n^*$ are continuous in $\mathbf{a}$. The constraints in the problem (1) define a set of *feasible* vectors $\mathbf{x}$ over which the function f is to be maximized. Let us call this set $G(a_1,\cdots,a_k)$, i.e.,

$$G(a_1,\cdots,a_k) = \{(x_1,\cdots,x_n) \mid g_j(x_1,\cdots,x_n,a_1,\cdots,a_k) = c_j, \; \forall \; j\}.$$

Now we can restate the problem as

$$\begin{aligned} &\max_{x_1,\cdots,x_n} \; f(x_1,\cdots,x_n,a_1,\cdots,a_k) \\ &\text{s.t. } \; (x_1,\cdots,x_n) \in G(a_1,\cdots,a_k). \end{aligned} \tag{2}$$

Notice that both the function f and the feasible set G depend on the parameters $\mathbf{a}$, i.e., both may change as $\mathbf{a}$ changes. The Maximum Theorem requires both that the function f be continuous as a function of $\mathbf{x}$ and $\mathbf{a}$ and that the feasible set $G(a_1,\cdots,a_k)$ change continuously as $\mathbf{a}$ changes. We already know what it means for f to be continuous but the notion of what it means for a set to change continuously is less elementary. We call G a **set valued function** or a **correspondence**. G associates with any vector $(a_1,\cdots,a_k)$ a subset of the vectors $(x_1,\cdots,x_n)$. The following two definitions define what we mean by a correspondence being continuous. First we define what it means for two sets to be close.

Definition 1 *Two sets of vectors A and B are **within ϵ of each other** if for any vector x in one set there is a vector x' in the other set such that $x' \in B_\epsilon(x)$.*

We can now define the continuity of the correspondence G in essentially the same way that we define the continuity of a single valued function.

Definition 2 *The correspondence G is **continuous** at $(a_1,\cdots,a_k)$ if $\forall\ \epsilon > 0$, $\exists\ \delta > 0$ such that if $(a_1',\cdots,a_k')$ is within δ of $(a_1,\cdots,a_k)$ then $G(a_1',\cdots,a_k')$ is within ϵ of $G(a_1,\cdots,a_k)$.*

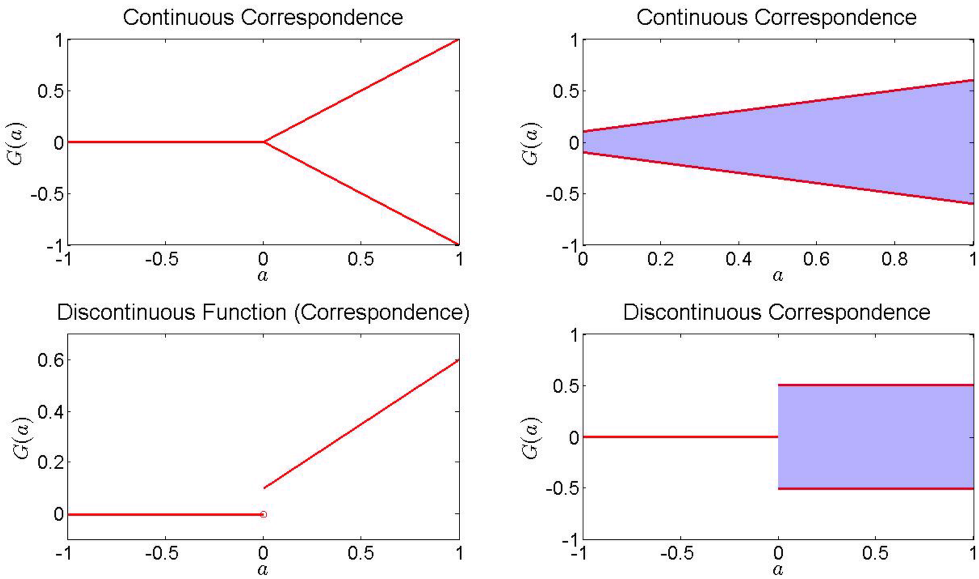

Figure 1: Continuous and Discontinuous Correspondence

Remark 1 *If G is a function, then the continuity of G as a correspondence is equivalent to the continuity as a function. Figure 1 shows some cases where G is continuous and discontinuous.*

It is, unfortunately, not the case that the continuity of the functions g_j necessarily implies the continuity of the feasible set. (Exercise 1 asks you to construct a counterexample.)

We are now in a position to state the Maximum Theorem. We assume that f is a continuous function, that G is a continuous correspondence, and that for any $(a_1, \cdots, a_k)$ the set $G(a_1, \cdots, a_k)$ is compact. The Weierstrass Theorem thus guarantees that there is a solution to the maximization problem (2) for any $(a_1, \cdots, a_k)$.

Theorem 1 (The Maximum Theorem) *Suppose that $f(x_1, \cdots, x_n, a_1, \cdots, a_k)$ is continuous (in $(x_1, \cdots, x_n, a_1, \cdots, a_k)$), that $G(a_1, \cdots, a_k)$ is a continuous correspondence, and that for any $(a_1, \cdots, a_k)$ the set $G(a_1, \cdots, a_k)$ is compact. Then*

(i) *$v(a_1, \cdots, a_k)$ is continuous, and*

(ii) *if $(x_1^*(a_1, \cdots, a_k), \cdots, x_n^*(a_1, \cdots, a_k))$ are (single valued) functions then they are also continuous.*

Our maximum theorem assumes $\mathbf{x}^*(a_1, \cdots, a_k)$ to be single valued, which can be achieved by our uniqueness theorem from Chapter 3. If $\mathbf{x}^*(a_1, \cdots, a_k)$ is set valued, then we can show $\mathbf{x}^*(a_1, \cdots, a_k)$ is a **upper hemicontinuous** correspondence, where the concept of upper hemicontinuity is not introduced to avoid further complication. Nevertheless, a single-valued upper hemicontinuous correspondence must be a continuous function; this is the second result of the maximum theorem.

Exercise 1 *Show by example that even if the functions g_j are continuous the correspondence G may not be continuous. [Hint: Use the case $n = m = k = 1$.]*

2 The Implicit Function Theorem

In Chapter 2 we said things like: "Now we have three equations in x_1^*, x_2^*, and the new artificial or auxiliary variable λ. Again we can, perhaps, solve these equations for x_1^*, x_2^*, and λ." In this section we examine the question of when we *can* solve a system of n equations to give n of the variable in terms of the others. Let us suppose that we have n endogenous variables $x_1, \ldots, x_n$, m exogenous variables or parameters, $b_1, \ldots, b_m$, and n equations or equilibrium conditions

$$\begin{aligned} f_1(x_1, \cdots, x_n, b_1, \cdots, b_m) &= 0, \\ f_2(x_1, \cdots, x_n, b_1, \cdots, b_m) &= 0, \\ &\vdots \\ f_n(x_1, \cdots, x_n, b_1, \cdots, b_m) &= 0, \end{aligned} \tag{3}$$

or, using vector notation,

$$\mathbf{f}(\mathbf{x}, \mathbf{b}) = \mathbf{0},$$

where $\mathbf{f} : \mathbb{R}^{n+m} \to \mathbb{R}^n$, $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{b} \in \mathbb{R}^m$, and $\mathbf{0} \in \mathbb{R}^n$.

When can we solve this system to obtain functions giving each x_i as a function of $b_1, \ldots, b_m$? As we'll see below we only give an incomplete answer to this question, but first let's look at the case that the function f is a linear function. Suppose that our equations are

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n + c_{11}b_1 + \cdots + c_{1m}b_m &= 0, \\ a_{21}x_1 + \cdots + a_{2n}x_n + c_{21}b_1 + \cdots + c_{2m}b_m &= 0, \\ &\vdots \\ a_{n1}x_1 + \cdots + a_{nn}x_n + c_{n1}b_1 + \cdots + c_{nm}b_m &= 0. \end{aligned}$$

We can write this, in matrix notation, as

$$[\mathbf{A} \mid \mathbf{C}] \begin{pmatrix} \mathbf{x} \\ \mathbf{b} \end{pmatrix} = \mathbf{0},$$

where $\mathbf{A}$ is an $n \times n$ matrix, $\mathbf{C}$ is an $n \times m$ matrix, $\mathbf{x}$ is an $n \times 1$ (column) vector, and $\mathbf{b}$ is an $m \times 1$ vector. This we can rewrite as

$$\mathbf{A}\mathbf{x} + \mathbf{C}\mathbf{b} = \mathbf{0},$$

and solve this to give

$$\mathbf{x} = -\mathbf{A}^{-1}\mathbf{C}\mathbf{b}.$$

And we can do this as long as the matrix $\mathbf{A}$ can be inverted, that is, as long as the matrix $\mathbf{A}$ is of

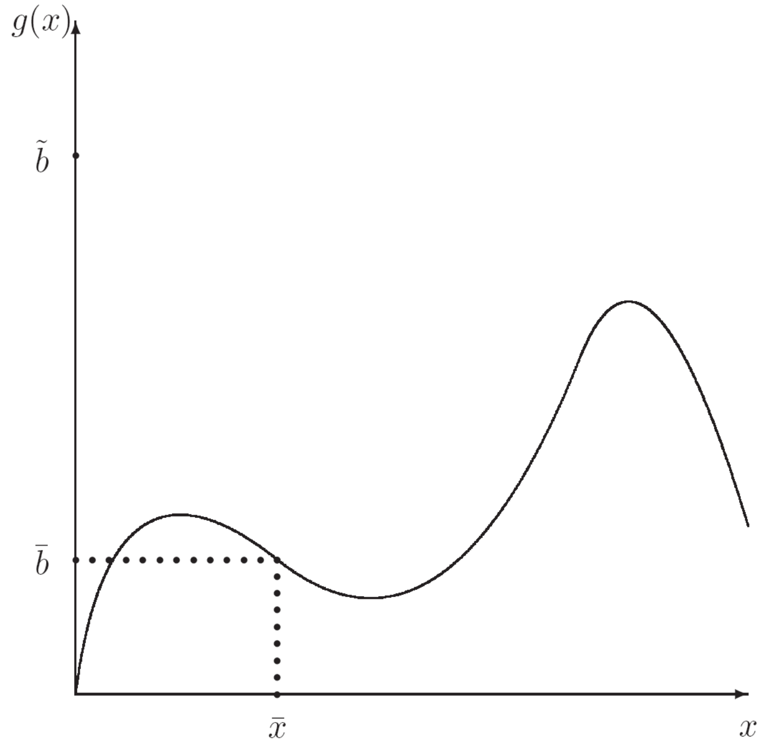

Figure 2: Intuition for the Implicit Function Theorem

full rank. As a result,

$$\frac{\partial \mathbf{x}}{\partial \mathbf{b}'} = -\mathbf{A}^{-1}\mathbf{C}.$$

Our answer to the general question in which the function f may not be linear is that if there are some values $(\bar{\mathbf{x}}, \bar{\mathbf{b}})$ for which $f(\bar{\mathbf{x}}, \bar{\mathbf{b}}) = \mathbf{0}$ then if, when we take a linear approximation to f we can solve the approximate linear system as we did above, then we can solve the true nonlinear system, at least in a neighborhood of $(\bar{\mathbf{x}}, \bar{\mathbf{b}})$. By this last phrase we mean that if $\mathbf{b}$ is not close to $\bar{\mathbf{b}}$ we may not be able to solve the system, and that for a particular value of $\mathbf{b}$ there may be many values of $\mathbf{x}$ that solve the system, but there is only one close to $\bar{\mathbf{x}}$.

To see why we can't, in general, do better than this consider the equation $f : \mathbb{R}^2 \to \mathbb{R}$ given by $f(x, b) = g(x) - b$, where the function g is graphed in Figure 2. Notice that the values $(\bar{x}, \bar{b})$ satisfy the equation $f(x, b) = 0$. For all values of b close to $\bar{b}$ we can find a unique value of x close to $\bar{x}$ such that $f(x, b) = 0$. However, (1) for each value of b there are other values of x far away from $\bar{x}$ that also satisfy $f(x, b) = 0$, and (2) there are values of b, such as $\tilde{b}$ for which there are no values of x that satisfy $f(x, b) = 0$.

Let us consider again the system of equations (3).

Theorem 2 (Implicit Function Theorem) *Suppose that* $\mathbf{f} : \mathbb{R}^{n+m} \to \mathbb{R}^n$ *is a* C^1 *function on* *an open set $A \subset \mathbb{R}^{n+m}$ and that $(\bar{\mathbf{x}}, \bar{\mathbf{b}})$ in A is such that $f(\bar{\mathbf{x}}, \bar{\mathbf{b}}) = \mathbf{0}$. Suppose also that*

$$\frac{\partial \mathbf{f}(\bar{\mathbf{x}}, \bar{\mathbf{b}})}{\partial \mathbf{x}'} = \begin{pmatrix} \frac{\partial f_1(\bar{\mathbf{x}}, \bar{\mathbf{b}})}{\partial x_1} & \cdots & \frac{\partial f_1(\bar{\mathbf{x}}, \bar{\mathbf{b}})}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n(\bar{\mathbf{x}}, \bar{\mathbf{b}})}{\partial x_1} & \cdots & \frac{\partial f_n(\bar{\mathbf{x}}, \bar{\mathbf{b}})}{\partial x_n} \end{pmatrix}$$

is of full rank. Then there are open sets $A_1 \subset \mathbb{R}^n$ and $A_2 \subset \mathbb{R}^m$ with $\bar{\mathbf{x}} \in A_1$, $\bar{\mathbf{b}} \in A_2$ and $A_1 \times A_2 \subset A$ such that for each $\mathbf{b}$ in A_2 there is exactly one $\mathbf{g}(\mathbf{b})$ in A_1 such that $\mathbf{f}(\mathbf{g}(\mathbf{b}), \mathbf{b}) = \mathbf{0}$. Moreover, $\mathbf{g} : A_2 \to A_1$ is a C^1 function and

$$\left[\frac{\partial \mathbf{g}(\mathbf{b})}{\partial \mathbf{b}'}\right]_{n\times m} = -\left[\frac{\partial \mathbf{f}(\mathbf{g}(\mathbf{b}), \mathbf{b})}{\partial \mathbf{x}'}\right]^{-1}_{n\times n} \left[\frac{\partial \mathbf{f}(\mathbf{g}(\mathbf{b}), \mathbf{b})}{\partial \mathbf{b}'}\right]_{n\times m},$$

or more compactly,

$$D_{\mathbf{b}}\mathbf{g}(\mathbf{b}) = -\left[D_{\mathbf{x}}\mathbf{f}(\mathbf{g}(\mathbf{b}), \mathbf{b})\right]^{-1} D_{\mathbf{b}}\mathbf{f}(\mathbf{g}(\mathbf{b}), \mathbf{b}).$$

How to understand this theorem? For $(\mathbf{x}, \mathbf{b})$ in a neighborhood of $(\bar{\mathbf{x}}, \bar{\mathbf{b}})$ such that $\mathbf{f}(\mathbf{x}, \mathbf{b}) = \mathbf{0}$ with $\mathbf{x} = \mathbf{g}(\mathbf{b})$, by totally differentiating the system of equations with respect to $\mathbf{b}$, we have

$$\frac{\partial \mathbf{f}(\mathbf{g}(\mathbf{b}), \mathbf{b})}{\partial \mathbf{x}'}\frac{\partial \mathbf{g}(\mathbf{b})}{\partial \mathbf{b}'} + \frac{\partial \mathbf{f}(\mathbf{g}(\mathbf{b}), \mathbf{b})}{\partial \mathbf{b}'} = \mathbf{0}_{n\times m},$$

so

$$\frac{\partial \mathbf{g}(\mathbf{b})}{\partial \mathbf{b}'} = -\left[\frac{\partial \mathbf{f}(\mathbf{g}(\mathbf{b}), \mathbf{b})}{\partial \mathbf{x}'}\right]^{-1} \left[\frac{\partial \mathbf{f}(\mathbf{g}(\mathbf{b}), \mathbf{b})}{\partial \mathbf{b}'}\right].$$

In summary, the implicit function theorem (IFT) is a *local* result, rather than a *global* result.[1] The IFT does not provide conditions to guarantee the existence of $(\bar{\mathbf{x}}, \bar{\mathbf{b}})$ such that $f(\bar{\mathbf{x}}, \bar{\mathbf{b}}) = \mathbf{0}$; rather, it provides conditions such that if such an $(\bar{\mathbf{x}}, \bar{\mathbf{b}})$ exists, then we can also uniquely solve $f(\mathbf{x}, \mathbf{b}) = \mathbf{0}$ in its neighborhood. So the most important application of the IFT is to obtain $\frac{\partial \mathbf{g}(\mathbf{b})}{\partial \mathbf{b}'}$ rather than guarantee the existence or uniqueness of the solution; this is so-called **comparative statics**.

Suppose the equality-constrained problem we are considering is

$$\begin{aligned} &\max_{\mathbf{x}} f(\mathbf{x}, \mathbf{c}) \\ &\text{s.t. } \mathbf{h}(\mathbf{x}, \mathbf{d}) = \mathbf{0}, \end{aligned}$$

where $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{c} \in \mathbb{R}^L$, $\mathbf{d} \in \mathbb{R}^M$, and $\mathbf{h} \in \mathbb{R}^K$. We can also assume there is some overlap between $\mathbf{c}$ and $\mathbf{d}$. Form the Lagrangian

$$\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \mathbf{c}, \mathbf{d}) = f(\mathbf{x}, \mathbf{c}) + \boldsymbol{\lambda} \cdot \mathbf{h}(\mathbf{x}, \mathbf{d});$$

[1] For a global result, see Gale and Nikaido (1965) and Berry et al. (2013).

then the FOCs are

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial \mathbf{x}}(\mathbf{x}^*, \boldsymbol{\lambda}^*, \mathbf{c}, \mathbf{d}) &= \frac{\partial f}{\partial \mathbf{x}}(\mathbf{x}^*, \mathbf{c}) + \frac{\partial \mathbf{h}(\mathbf{x}^*, \mathbf{d})'}{\partial \mathbf{x}}\boldsymbol{\lambda}^* = \mathbf{0}, \\ \frac{\partial \mathcal{L}}{\partial \boldsymbol{\lambda}}(\mathbf{x}^*, \boldsymbol{\lambda}^*, \mathbf{c}, \mathbf{d}) &= \mathbf{h}(\mathbf{x}^*, \mathbf{d}) = \mathbf{0}.\end{aligned}$$

Casting in the notation of the IFT,

$$\mathbf{f}(\mathbf{x}, \mathbf{b}) = \begin{pmatrix} \frac{\partial f}{\partial \mathbf{x}}(\mathbf{x}, \mathbf{c}) + \frac{\partial \mathbf{h}(\mathbf{x},\mathbf{d})'}{\partial \mathbf{x}}\boldsymbol{\lambda} \\ \mathbf{h}(\mathbf{x}, \mathbf{d}) \end{pmatrix} = D_{\mathbf{x},\boldsymbol{\lambda}}\mathcal{L}(\mathbf{x}^*, \boldsymbol{\lambda}^*, \mathbf{c}, \mathbf{d})',$$

where $\mathbf{x} \sim (\mathbf{x}, \boldsymbol{\lambda})$, $\mathbf{b} \sim (\mathbf{c}, \mathbf{d})$, $n \sim n + K$, and $m \sim (L + M)$, i.e., there are $(n + K)$ unknowns $(\mathbf{x}, \boldsymbol{\lambda})$ and $(L + M)$ parameters. Suppose f and $\mathbf{h}$ are C^2 functions; then $\mathbf{f}(\mathbf{x}, \mathbf{b})$ is a C^1 function as required by the IFT. Suppose that $(\mathbf{x}^*, \boldsymbol{\lambda}^*)$ satisfies

$$\begin{pmatrix} \frac{\partial f}{\partial \mathbf{x}}(\mathbf{x}^*, \mathbf{c}) + \frac{\partial \mathbf{h}(\mathbf{x}^*,\mathbf{d})'}{\partial \mathbf{x}}\boldsymbol{\lambda}^* \\ \mathbf{h}(\mathbf{x}^*, \mathbf{d}) \end{pmatrix} = \mathbf{0},$$

and

$$\begin{pmatrix} D^2_{\mathbf{x}} f(\mathbf{x}^*, \mathbf{c}) + D^2_{\mathbf{x}}\left(\mathbf{h}(\mathbf{x}^*, \mathbf{d})'\boldsymbol{\lambda}^*\right) & \left(D_{\mathbf{x}}\mathbf{h}(\mathbf{x}^*, \mathbf{d})\right)' \\ D_{\mathbf{x}}\mathbf{h}(\mathbf{x}^*, \mathbf{d}) & \mathbf{0}_{K\times K} \end{pmatrix}$$

is of full rank, where $D^2_{\mathbf{x}}\left(\mathbf{h}(\mathbf{x}^*, \mathbf{d})'\boldsymbol{\lambda}^*\right) = \sum_{k=1}^{K} \lambda^*_k \left(D^2_{\mathbf{x}} h_k(\mathbf{x}^*, \mathbf{d})\right)$. Then the assumptions of the IFT are satisfied, and we can apply the theorem to have

$$\frac{\partial \left(\mathbf{x}^{*\prime}, \boldsymbol{\lambda}^{*\prime}\right)'}{\partial \left(\mathbf{c}', \mathbf{d}'\right)} = -\begin{pmatrix} D^2_{\mathbf{x}} f(\mathbf{x}^*, \mathbf{c}) + D^2_{\mathbf{x}}\left(\mathbf{h}(\mathbf{x}^*, \mathbf{d})'\boldsymbol{\lambda}^*\right) & \left(D_{\mathbf{x}}\mathbf{h}(\mathbf{x}^*, \mathbf{d})\right)' \\ D_{\mathbf{x}}\mathbf{h}(\mathbf{x}^*, \mathbf{d}) & \mathbf{0}_{K\times K} \end{pmatrix}^{-1} \begin{pmatrix} \frac{\partial^2 f}{\partial \mathbf{x}\partial \mathbf{c}'}(\mathbf{x}^*, \mathbf{c}) & \frac{\partial \mathbf{h}(\mathbf{x}^*,\mathbf{d})'\boldsymbol{\lambda}^*}{\partial \mathbf{x}\partial \mathbf{d}'} \\ \mathbf{0}_{K\times L} & \frac{\partial \mathbf{h}(\mathbf{x}^*,\mathbf{d})}{\partial \mathbf{d}'} \end{pmatrix},$$

where $(\mathbf{x}^*, \boldsymbol{\lambda}^*)$ are treated as functions of $(\mathbf{c}, \mathbf{d})$. It is interesting to observe that $D^2_{\mathbf{x}} f(\mathbf{x}^*, \mathbf{c}) + D^2_{\mathbf{x}}\left(\mathbf{h}(\mathbf{x}^*, \mathbf{d})'\boldsymbol{\lambda}^*\right)$ is exactly $D^2_{\mathbf{x}}\mathcal{L}(\mathbf{x}^*, \boldsymbol{\lambda}^*, \mathbf{c}, \mathbf{d})$ in Section 3 of Chapter 3, the first matrix on the right hand side is $D^2_{\mathbf{x},\boldsymbol{\lambda}}\mathcal{L}(\mathbf{x}^*, \boldsymbol{\lambda}^*, \mathbf{c}, \mathbf{d})$, and the second matrix is $\frac{\partial^2 \mathcal{L}(\mathbf{x}^*,\boldsymbol{\lambda}^*,\mathbf{c},\mathbf{d})}{\partial(\mathbf{x}',\boldsymbol{\lambda}')'\partial(\mathbf{c}',\mathbf{d}')}$.

We use the following example to illustrate the application of the IFT in the comparative static analysis.

Example 1 *The seller of a product pays a proportional tax at a flat rate $\theta \in (0, 1)$. Hence, the effective price received by the seller is $(1 - \theta) P$, where P is the market price for the good. Market supply and demand are given by the differentiable functions*

$$\begin{aligned} Q^d &= D(P),\ \text{with } D'(\cdot) < 0 \\ Q^s &= S((1 - \theta) P),\ \text{with } S'(\cdot) > 0 \end{aligned}$$

and equilibrium requires market clearing, that is, $Q^s = Q^d$. Analyze, graphically and analytically, the effects of a decrease in the tax rate on the quantity transacted and the equilibrium price.

Solution: Market clearing requires

$$S((1-\theta)\,P) = D(P). \tag{4}$$

This equation implicitly defines the equilibrium price as a function $P^* = P(\theta)$ of the parameter θ. Substituting the solution function $P(\cdot)$ back into (4), we have the identity

$$S\left[(1-\theta)\,P(\theta)\right] = D\left[P(\theta)\right].$$

Applying the IFT directly with $f(P,\theta) = D\left[P(\theta)\right] - S\left[(1-\theta)\,P(\theta)\right]$, we have

$$P'(\theta) = \frac{-PS'(\cdot)}{D'(\cdot) - (1-\theta)\,S'(\cdot)} = \frac{(-)}{(-)} > 0.$$

Next, the quantity transacted in equilibrium is given by $Q^* = D\left[P(\theta)\right]$, and therefore

$$\frac{dQ^*}{d\theta} = D'(P^*)\,P'(\theta) < 0.$$

Graphically, a reduction in the tax rate increases the effective price received by sellers for any given market price; these are therefore willing to sell any given quantity at a lower market price. Hence the supply curve shifts down. The equilibrium price falls, and the equilibrium quantity increases, as shown in Figure 3. □

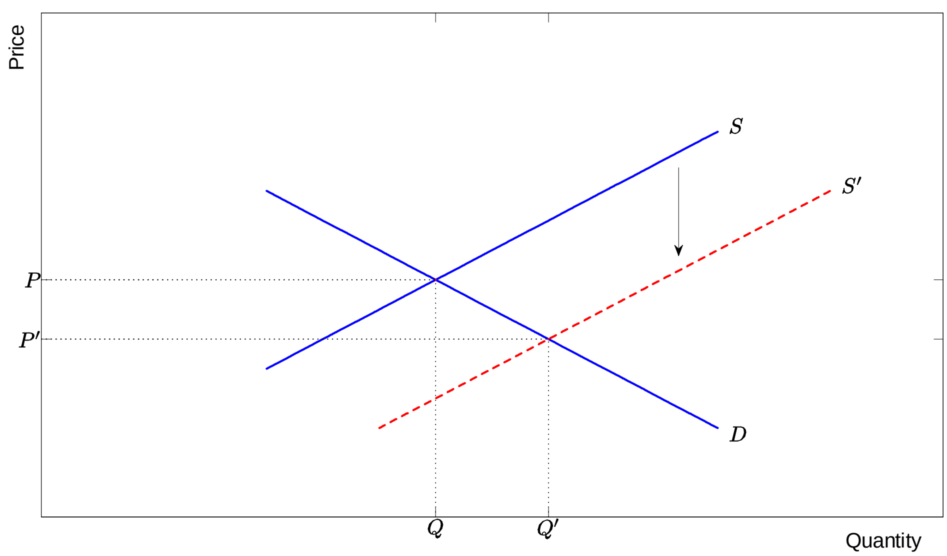

Figure 3: Effect of a Tax Reduction

Exercise 2 *Consider the general utility maximization problem*

$$\max_{x_1,\cdots,x_n} u(x_1,\cdots,x_n)$$
$$s.t.\ p_1x_1+\cdots+p_nx_n = w.$$

Suppose that for some price vector $\bar{\mathbf{p}}$ *the maximization problem has a utility maximizing bundle* $\bar{\mathbf{x}}$*. Find conditions on the utility function such that in a neighborhood of* $(\bar{\mathbf{x}},\bar{\mathbf{p}})$ *we can solve for the demand functions* $\mathbf{x}(\mathbf{p})$*. Find the derivatives of the demand functions,* $\partial\mathbf{x}/\partial\mathbf{p}'$*.*

Exercise 3 *Now suppose that there are only two goods and the utility function is given by*

$$u(x_1,x_2) = (x_1)^{\frac{1}{3}}(x_2)^{\frac{2}{3}}.$$

Solve this utility maximization problem, as you learned to do in Chapter 2, and then differentiate the demand functions that you find to find the partial derivative with respect to p_1*,* p_2*, and* w *of each demand function. Also find the same derivatives using the method of the previous exercise.*

3 The Envelope Theorem

In this section we examine a theorem that is particularly useful in the study of consumer and producer theory. There is in fact nothing mysterious about this theorem. You will see that the proof of this theorem is simply calculation and a number of substitutions. Moreover the theorem has a very clear intuition. It is this: Suppose we are at a maximum (in an unconstrained problem) and we change the data of the problem by a very small amount. Now both the solution of the problem and the value at the maximum will change. However at a maximum the function is flat (the first derivative is zero). Thus when we want to know by how much the maximized value has changed it does not matter (very much) whether or not we take account of how the maximizer changes or not. See Figure 4. The intuition for a constrained problem is similar and only a little more complicated.

To motivate our discussion of the Envelope Theorem we will first consider a particular case, viz., the relation between short and long run average cost curves. Recall that, in general we assume that the average cost of producing some good is a function of the amount of the good to be produced. The short run average cost function is defined to be the function which for any quantity, Q, gives the average cost of producing that quantity, taking as given the scale of operation, i.e., the size and number of plants and other fixed capital which we assume cannot be changed in the short run (whatever that is). The long run average cost function on the other hand gives, as a function of Q, the average cost of producing Q units of the good, with the scale of operation selected to be the optimal scale for that level of production.

That is, if we let the scale of operation be measured by a single variable k, say, and we let the short run average cost of producing Q units when the scale is k be given by $SRAC(Q, k)$ and the

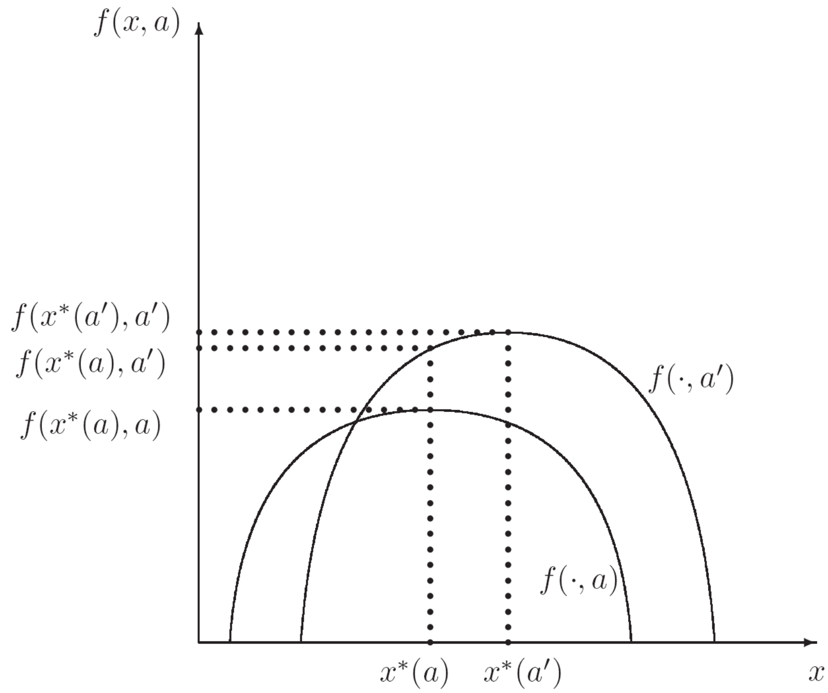

Figure 4: Intuition for the Envelope Theorem

long run average cost of producing Q units by $LRAC(Q)$ then we have

$$LRAC(Q) = \min_k SRAC(Q, k).$$

Let us denote, for a given value Q, the optimal level of k by $k(Q)$. That is, $k(Q)$ is the value of k that minimizes the right hand side of the above equation.

Graphically, for any fixed level of k the short run average cost function can be represented by a curve (normally assumed to be U-shaped) drawn in two dimensions with quantity on the horizontal axis and cost on the vertical axis. Now think about drawing one short run average cost curve for each of the (infinite) possible values of k. One way of thinking about the long run average cost curve is as the "bottom" or *envelope* of these short run average cost curves. Suppose that we consider a point on this long run or envelope curve. What can be said about the slope of the long run average cost curve at this point. A little thought should convince you that it should be the same as the slope of the short run curve through the same point. (If it were not then that short run curve would come below the long run curve, a contradiction.) That is,

$$\frac{d\ LRAC(Q)}{dQ} = \frac{\partial\ SRAC(Q, k(Q))}{\partial Q}.$$

See Figure 5.

The envelope theorem is a general statement of the result of which this is a special case. We

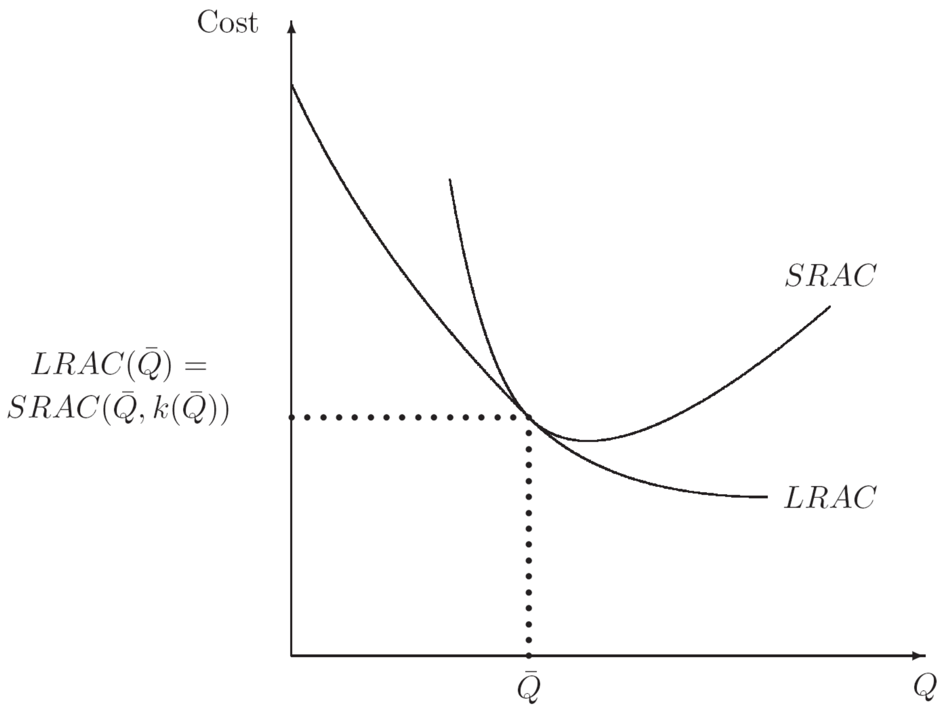

Figure 5: The Relationship Between SRAC and LRAC

will consider not only cases in which Q and k are vectors, but also cases in which the maximization or minimization problem includes some constraints.

Let us consider again the maximization problem (1). Let $\mathcal{L}(x_1, \cdots, x_n, \lambda_1, \cdots, \lambda_m; a_1, \cdots, a_k)$ be the Lagrangian function:

$$\mathcal{L}(x_1, \cdots, x_n, \lambda_1, \cdots, \lambda_m; a_1, \cdots, a_k) = f(x_1, \cdots, x_n, a_1, \cdots, a_k) + \sum_{j=1}^{m} \lambda_j (c_j - g_j(x_1, \cdots, x_n, a_1, \cdots, a_k)).$$

Let $(x_1^*(a_1, \cdots, a_k), \cdots, x_n^*(a_1, \cdots, a_k))$ and $(\lambda_1^*(a_1, \cdots, a_k), \cdots, \lambda_m^*(a_1, \cdots, a_k))$ be the values of x and λ that solve this problem. Now let

$$v(a_1, \cdots, a_k) = f(x_1^*(a_1, \cdots, a_k), \cdots, x_n^*(a_1, \cdots, a_k), a_1, \cdots, a_k).$$

That is, $v(a_1, \cdots, a_k)$ is the maximized value of the function f when the parameters are $(a_1, \cdots, a_k)$. The envelope theorem says that the derivative of v is equal to the derivative of $\mathcal{L}$ at the maximizing values of x and λ. Or, more precisely

Theorem 3 (The Envelope Theorem) *If all functions are defined as above and the problem is* *such that the functions x^* and λ^* are well defined, then*

$$\begin{aligned}\frac{\partial v}{\partial a_h}(a_1,\cdots,a_k) =& \frac{\partial \mathcal{L}}{\partial a_h}(x_1^*(a_1,\cdots,a_k),\cdots,x_n^*(a_1,\cdots,a_k),\lambda_1^*(a_1,\cdots,a_k),\cdots,\lambda_m^*(a_1,\cdots,a_k);a_1,\cdots,a_k)\\ =& \frac{\partial f}{\partial a_h}(x_1^*(a_1,\cdots,a_k),\cdots,x_n^*(a_1,\cdots,a_k),a_1,\cdots,a_k)\\ &-\sum_{j=1}^m \lambda_j^*(a_1,\cdots,a_k)\frac{\partial g_h}{\partial a_h}(x_1^*(a_1,\cdots,a_k),\cdots,x_n^*(a_1,\cdots,a_k),a_1,\cdots,a_k)\end{aligned}$$

for all h.

In order to show the advantages of using matrix and vector notation we shall restate the theorem in that notation before returning to give a proof of the theorem. (In proving the theorem we shall return to using mainly scalar notation.)

$$\begin{aligned}\frac{\partial v}{\partial \mathbf{a}}(\mathbf{a}) &= \frac{\partial \mathcal{L}}{\partial \mathbf{a}}(\mathbf{x}^*(\mathbf{a}),\boldsymbol{\lambda}^*(\mathbf{a});\mathbf{a})\\ &= \frac{\partial f}{\partial \mathbf{a}}(\mathbf{x}^*(\mathbf{a}),\mathbf{a}) - \frac{\partial \mathbf{g}(\mathbf{x}^*(\mathbf{a}),\mathbf{a})'}{\partial \mathbf{a}}\boldsymbol{\lambda}^*(\mathbf{a}).\end{aligned}$$

Proof. From the definition of the function v we have

$$v(a_1,\cdots,a_k) = f(x_1^*(a_1,\cdots,a_k),\cdots,x_n^*(a_1,\cdots,a_k),a_1,\cdots,a_k).$$

Thus

$$\frac{\partial v}{\partial a_h}(\mathbf{a}) = \frac{\partial f}{\partial a_h}(\mathbf{x}^*(\mathbf{a}),\mathbf{a}) + \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\mathbf{x}^*(\mathbf{a}),\mathbf{a})\frac{\partial x_i^*}{\partial a_h}(\mathbf{a}). \tag{5}$$

Now, from the FOCs we have

$$\frac{\partial f}{\partial x_i}(\mathbf{x}^*(\mathbf{a}),\mathbf{a}) - \sum_{j=1}^m \lambda_j(\mathbf{a})\frac{\partial g_j}{\partial x_i}(\mathbf{x}^*(\mathbf{a}),\mathbf{a}) = 0.$$

Or

$$\frac{\partial f}{\partial x_i}(\mathbf{x}^*(\mathbf{a}),\mathbf{a}) = \sum_{j=1}^m \lambda_j(\mathbf{a})\frac{\partial g_j}{\partial x_i}(\mathbf{x}^*(\mathbf{a}),\mathbf{a}). \tag{6}$$

Also, since $\mathbf{x}^*(\mathbf{a})$ satisfies the constraints we have, for each j

$$g_j(x_1^*(\mathbf{a}),\cdots,x_n^*(\mathbf{a}),a_1,\cdots,a_k) = c_j.$$

And, since this holds as an identity, we may differentiate both sides with respect to a_h giving

$$\sum_{i=1}^n \frac{\partial g_j}{\partial x_i}(\mathbf{x}^*(\mathbf{a}),\mathbf{a})\frac{\partial x_i^*}{\partial a_h}(\mathbf{a}) + \frac{\partial g_j}{\partial a_h}(\mathbf{x}^*(\mathbf{a}),\mathbf{a}) = 0.$$

Or

$$\sum_{i=1}^{n} \frac{\partial g_j}{\partial x_i}(\mathbf{x}^*(\mathbf{a}), \mathbf{a}) \frac{\partial x_i^*}{\partial a_h}(\mathbf{a}) = -\frac{\partial g_j}{\partial a_h}(\mathbf{x}^*(\mathbf{a}), \mathbf{a}). \tag{7}$$

Substituting (6) into (5) gives

$$\frac{\partial v}{\partial a_h}(\mathbf{a}) = \frac{\partial f}{\partial a_h}(\mathbf{x}^*(\mathbf{a}), \mathbf{a}) + \sum_{i=1}^{n} \left[\sum_{j=1}^{m} \lambda_j(\mathbf{a}) \frac{\partial g_j}{\partial x_i}(\mathbf{x}^*(\mathbf{a}), \mathbf{a}) \right] \frac{\partial x_i^*}{\partial a_h}(\mathbf{a}).$$

Changing the order of summation gives

$$\frac{\partial v}{\partial a_h}(\mathbf{a}) = \frac{\partial f}{\partial a_h}(\mathbf{x}^*(\mathbf{a}), \mathbf{a}) + \sum_{j=1}^{m} \lambda_j(\mathbf{a}) \left[\sum_{i=1}^{n} \frac{\partial g_j}{\partial x_i}(\mathbf{x}^*(\mathbf{a}), \mathbf{a}) \frac{\partial x_i^*}{\partial a_h}(\mathbf{a}) \right]. \tag{8}$$

And now substituting (7) into (8) gives

$$\frac{\partial v}{\partial a_h}(\mathbf{a}) = \frac{\partial f}{\partial a_h}(\mathbf{x}^*(\mathbf{a}), \mathbf{a}) - \sum_{j=1}^{m} \lambda_j(\mathbf{a}) \frac{\partial g_j}{\partial a_h}(\mathbf{x}^*(\mathbf{a}), \mathbf{a}),$$

which is the required result. ■

Exercise 4 *Rewrite this proof using matrix notation. Go through your proof and identify the dimension of each of the vectors or matrices you use. For example,* $D_{\mathbf{x}} f$ *is a* $1 \times n$ *vector, and* $D_{\mathbf{x}}\mathbf{g}$ *is an* $m \times n$ *matrix.*

In microeconomics, Hotelling's Theorem, Hicks-Slutsky equations and Roy's Theorem in consumer theorem, Hotelling's lemma in production theory, and Shephard's lemma in both consumer and production theory are all straightforward applications of the envelope theorem.

3.1 Interpretation of the Lagrange Multiplier

A straightforward corollary of the envelope theorem is the meaning of the Lagrange multiplier. We have actually already interpreted the Lagrange multiplier intuitively as the penalty on violating the constraint in Section 1.1 of Chapter 2; here is a more rigorous statement. Treating c_j, $j = 1, \cdots, m$, also as parameters, then by the envelope theorem,

$$\frac{\partial v}{\partial c_j}(a_1, \cdots, a_k, c_1, \cdots, c_m) = \lambda_j^*(a_1, \cdots, a_k, c_1, \cdots, c_m).$$

How to understand this result? Think of f as the profit function of a firm, the g_j equation as the resource constraint, and c_j as the amount of input j available to the firm. In this situation, $\frac{\partial v}{\partial c_j}(a_1, \cdots, a_k, c_1, \cdots, c_m)$ represents the change in the optimal profit resulting from availability of one more unit of input j. Alternatively, it tells the maximum amount the firm would be willing to pay to acquire another unit of input j. For this reason, λ_j^* is often called the **internal value** or

imputed value, or more frequently, the **shadow price** of input j. It may be a more important index to the firm than the external market price of input j.

We consider only the equality-constrained problem in this chapter. All the results can be extended to the inequality-constrained problems. For example, in the envelope theorem, if some inequality constraint is not binding, then the corresponding multiplier is zero, and we can consider only the binding constraints without loss of generality.