Lecture 7 COST MINIMIZATION

7.1. REVIEW

[Constrained optimization problem, unconstrained approach, Lagrange approach…]

7.2. COST MINIMIZATION 1

7.2.1 Some basic results

Consider the case of a firm producing single output $y \in \mathbf{R}$ using n inputs $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbf{R}^n$, x_i denoting the i-th input, $i = 1, \ldots, n$. Let $\mathbf{F} \subset \mathbf{R}^{n+1}$ be the feasible set representing the firm technology, where $(y, \mathbf{x}) \in \mathbf{F}$ means that the production decision $(y, \mathbf{x})$ is technologically feasible. As before, it will be useful to put some structure on the representation of the production technology. Define the production function as

$$g(\mathbf{x}) = \underset{y}{Max} \quad \{y : (y, \mathbf{x}) \in \mathbf{F}\}.$$

Then, express technological feasibility as

$$\mathbf{F} = \{(y, \mathbf{x}): y \leq g(\mathbf{x}): \mathbf{x} \geq 0, y \geq 0\}.$$

For a firm facing *competitive input markets*, let $\mathbf{w} = (w_1, w_2, \ldots, w_n) \in \mathbf{R}^n$ denote the n-vector of input prices, where w_i is the market price for x_i, $i = 1, \ldots, n$. We will assume throughout that prices are positive: $w_i > 0$, $i = 1, \ldots, n$. Then, $C = \mathbf{w} \cdot \mathbf{x} = \Sigma_i\, w_i\, x_i$ is the cost of production for the firm. Then, the smallest feasible cost of producing output $y > 0$ is given by the cost minimization problem

$$\underset{\mathbf{x}}{Min} \quad \{\mathbf{w} \cdot \mathbf{x} : (y, \mathbf{x}) \in \mathbf{F}\}, \text{ or}$$

$$\underset{\mathbf{x}}{Min} \quad \{\mathbf{w} \cdot \mathbf{x} : y \leq g(\mathbf{x}), \mathbf{x} \geq 0\}. \tag{1}$$

For a given output $y > 0$ and under input prices $w > 0$, denote by $\mathbf{x}^c$ *the solution of this minimization problem*. We want to identify and investigate the properties of cost minimizing behavior.

We will make the (intuitive) assumption:

<u>Assumption A1</u>: *Producing a positive output, $y > 0$, requires at least some positive inputs.*

We have the following result:

Under assumption A1, cost minimization implies **technical efficiency,** *where $(y, \mathbf{x}^c)$ is located on the production function, $y = g(\mathbf{x}^c)$.*

This implies that *A1 is a sufficient condition for cost minimizing behavior to be technically efficient*. Alternatively, under such conditions, *observing technical inefficiency (where $y < g(\mathbf{x})$) is inconsistent with cost minimizing behavior*. Two factors can contribute to this situation: 1/ the firm does not have access to the best available technology (suggesting a need to speed up technological adoption for the firm); 2/ the firm management is poor and makes poor use of available resource in the production process (suggesting a need either to replace the management or to improve managerial abilities).

It follows that, under assumption A1, the cost minimization problem can be written equivalently as

$$\underset{\mathbf{x}}{Min} \quad \{\mathbf{w} \cdot \mathbf{x} : y = g(\mathbf{x}), \mathbf{x} \geq 0, \mathbf{x} \in \mathbf{R}^n\}. \tag{2}$$

Denote the solution of this minimization problem by $\mathbf{x}^c(y, \mathbf{w}) = (x_1^c(y, \mathbf{w}), \ldots, x_n^c(y, \mathbf{w}))$. They are the *cost minimizing decision rules* or *cost minimizing input demand functions*. In this section, we have three objectives:

1) to identify cost minimizing behavior $\mathbf{x}^c(y, \mathbf{w})$;
2) to relate cost minimizing behavior to profit maximizing behavior;
3) to investigate the properties of the decision rules $\mathbf{x}^c(y, \mathbf{w})$, providing information on how an economically rational firm would behave under changing input markets conditions (as represented by prices $\mathbf{w}$).

7.2.2 Production Behavior

Earlier, we analyzed the constrained maximization problem $\underset{\mathbf{x}}{Max}\ \{f(\mathbf{x}): h_1(\mathbf{x}) = 0, \ldots h_m(\mathbf{x}) = 0; \mathbf{x} \geq 0; \mathbf{x} \in \mathbf{R}^n\}$, and we know that the same rationale applies to the constrained minimization problem, $\underset{\mathbf{x}}{Min}\ \{f(\mathbf{x}): \mathbf{h}(\mathbf{x}) = 0; \mathbf{x} \geq 0; \mathbf{x} \in \mathbf{R}^n\}$, with opposite signs in the SONCs and SOSCs. Let $f(\mathbf{x}) \equiv \mathbf{w} \cdot \mathbf{x}$ be the objective function (note that it is linear in $\mathbf{x} \in \mathbf{R}^n$). Let $\mathbf{h}(\mathbf{x}) \equiv y - g(\mathbf{x}) = 0$ denote the constraint that production decisions be technically efficient. Assume that the production function $g(\mathbf{x})$ is twice continuous differentiable. Then, cost minimization given in (2) is a standard constrained minimization problem, with a single equality restriction ($m = 1$):

$$\underset{\mathbf{x}}{Min}\ \{\mathbf{w} \cdot \mathbf{x} : y - g(\mathbf{x}) = 0, \mathbf{x} \geq 0, \mathbf{x} \in \mathbf{R}^n\}, \tag{2'}$$

which has for solution $\mathbf{x}^c(y, \mathbf{w}) = (x_1^c(y, \mathbf{w}), \ldots, , x_n^c(y, \mathbf{w}))$.

We are interested in applying the Lagrange approach to problem (2'). We know that it requires the following constraint qualification:

Constraint Qualification (CQ): There exists an input i satisfying $\frac{\partial \mathbf{h}(\mathbf{x}^c)}{\partial x_i} \equiv -\frac{\partial g(\mathbf{x}^c)}{\partial x_i} \neq 0$.

Result 1: *Under cost minimization, finding $x_i^c > 0$ for some input i implies that* $\frac{\partial g(\mathbf{x}^c)}{\partial x_i} \geq 0$.

Proof: To see that, assume that $\frac{\partial g(\mathbf{x}^c)}{\partial x_i} < 0$ for some i where $x_i^c > 0$. Then, reducing x_i^c would be feasible in (1) and would reduce cost (given $w_i > 0$). This would contradict that x_i^c is cost minimizing.

It shows that, whenever any input is used, *cost minimizing behavior rules out the production region where its marginal product is negative*. Alternatively stated, finding evidence that a firm is using an input $x_i > 0$ in the region of negative marginal product (where $\frac{\partial g}{\partial x_i} < 0$) is sufficient information to conclude that the firm is *not* minimizing cost.

Below, we will consider the case where $\mathbf{x}^c$ is an *interior solution* to the cost minimization problem (2) or (2'), i.e. where $x_i^c > 0$ for all inputs, $i = 1, \ldots, n$.

Using the Lagrange approach, define the Lagrangean

$$L(\mathbf{x}, \lambda) = \mathbf{w} \cdot \mathbf{x} + \lambda\, [y - g(\mathbf{x})], \tag{3}$$

where $\lambda \in \mathbf{R}$ is a Lagrange multiplier associated with the constraint $y - g(\mathbf{x}) = 0$.

When the constraint qualification (*CQ*) *holds and* $\mathbf{x}^c$ *is an interior solution*, we know that cost minimizing behavior $\mathbf{x}^c$ implies the *first order necessary condition* (*FONC*)

$$L_x(\mathbf{x}^c, \lambda^c) \equiv [\tfrac{\partial L(\mathbf{x}^c,\lambda^c)}{\partial x_1}, \ldots, \tfrac{\partial L(\mathbf{x}^c,\lambda^c)}{\partial x_n}] \equiv \mathbf{w}^T - \lambda^c g_{\mathbf{x}}(\mathbf{x}^c) = 0, \tag{4a}$$

and

$$L_\lambda(\mathbf{x}^c, \lambda^c) \equiv \tfrac{\partial L(\mathbf{x}^c,\lambda^c)}{\partial \lambda} \equiv y - g(\mathbf{x}^c) = 0. \tag{4b}$$

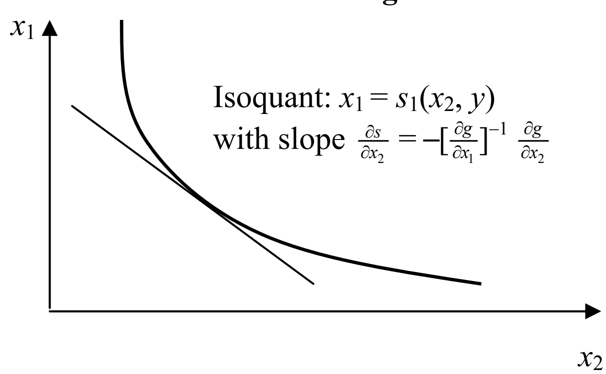
where $\mathbf{w}^T = (w_1, \ldots, w_n)$ is a $(1\times n)$ vector of input prices, and $g_{\mathbf{x}}(\mathbf{x}^c) = [\tfrac{\partial g(\mathbf{x}^c)}{\partial x_1}, \ldots, \tfrac{\partial g(\mathbf{x}^c)}{\partial x_n}]$ is a $(1\times n)$ vector of marginal products. Equation (4a) can be equivalently written as

$\lambda^c g_{\mathbf{x}}(\mathbf{x}^c) = \mathbf{w}^T$, or

$$\lambda^c \tfrac{\partial g(\mathbf{x}^c)}{\partial x_i} = w_i,\ i = 1, \ldots, n. \tag{4a'}$$

Define an *isoquant* as the set $\mathbf{Q}(y) = \{\mathbf{x}: g(\mathbf{x}) = y, \mathbf{x} \geq 0, \mathbf{x} \in \mathbf{R}^n\}$, i.e. as the set of inputs $\mathbf{x}$ generating a given output $y > 0$. For $\mathbf{x} > 0$, the isoquant is characterized by $g(\mathbf{x}) = y$. Consider solving the equation $g(\mathbf{x}) = y$ for x_1, giving $x_1 = s_1(x_2, \ldots, x_n, y)$. Then, for a given $y > 0$, $x_1 = s_1(x_2, \ldots, x_n, y)$ is the equation for an isoquant. Define the *marginal rate of substitution* between x_i and x_1 as the (negative) slope of the isoquant, $-\tfrac{\partial s_1}{\partial x_i}$, $i = 2, \ldots, n$. Given $\tfrac{\partial g}{\partial x_i} > 0$, apply the implicit function theorem to $g(s(x_2, \ldots, x_n, y), x_2, \ldots, x_n) = y$ to obtain $\tfrac{\partial s}{\partial x_i} = -[\tfrac{\partial g}{\partial x_1}]^{-1} \tfrac{\partial g}{\partial x_i}$, $i = 1, \ldots, n$. In general, assuming that $\tfrac{\partial g}{\partial x_j} \neq 0$, it follows that the *marginal rate of substitution between inputs i and j is* $MRS_{ij} = [\tfrac{\partial g}{\partial x_1}]^{-1} \tfrac{\partial g}{\partial x_i}$. See figure 7.1.

Figure 7.1



Given $w_i > 0$, (4a') implies that $\lambda^c \neq 0$ and $\tfrac{\partial g(\mathbf{x}^c)}{\partial x_i} \neq 0$, $i = 1, \ldots, n$. Then, (4a') can be written as

$$\frac{\frac{\partial g(\mathbf{x}^c)}{\partial x_i}}{\frac{\partial g(\mathbf{x}^c)}{\partial x_j}} = \frac{w_i}{w_j}, \text{ for all } i \neq j.$$

But the left-hand side is *the marginal rate of substitution between inputs* x_i *and* x_j, MRS_{ij}. This shows that:

> *The FONC* (*4a*)-(*4b*) *imply that the marginal rate of substitution between any two inputs* x_i *and* x_j, MRS_{ij}, *must be equal to the input price ratio* $\frac{w_i}{w_j}$.

Finally, note that an interior solution implies $\frac{\partial g(\mathbf{x}^c)}{\partial x_i} \geq 0$ from **result 1**. Then, given positive price ($w_i > 0$), equation (4a') yields:

Result 2: *An interior solution to the cost minimization problem implies that the Lagrange multiplier is necessarily positive, $\lambda^c > 0$.*

We know that under (CQ), FONC (4a)-(4b) are *necessary* conditions for $\mathbf{x}^c$ to be an interior solution to the cost minimization problem. Equations (4a)-(4b) constitute a system of $n + 1$ equations in $n + 1$ unknown, $(\mathbf{x}, \lambda) = (x_1, \ldots, x_n, \lambda)$. Consider the values $(\mathbf{x}^c, \lambda^c)$ that solve (4a)-(4b). Under what conditions would we be sure that $\mathbf{x}^c$ is a solution to the cost minimization problem?

In the Lagrange approach, we know that (4a)-(4b) is a *sufficient* condition to identify $\mathbf{x}^c$ as an interior solution to the cost minimization problem (2) or (2') *if* three conditions hold:

1. $\mathbf{h}(\mathbf{x})$ are quasi-convex functions,
2. $\lambda^c \geq 0$,
3. $f(\mathbf{x})$ is convex.

In our cost minimization problem, $f(\mathbf{x}) \equiv \mathbf{w} \cdot \mathbf{x}$ is linear; thus it is convex. In addition, $h(\mathbf{x}) \equiv y - g(\mathbf{x})$ is quasi-convex if and only $g(\mathbf{x})$ is quasi-concave. Finally, from result 2, an interior solution implies that $\lambda^c > 0$. We obtain:

if the production function $g(\mathbf{x})$ is quasi-concave, then (4a)-(4b) is a sufficient condition to identify $\mathbf{x}^c$ as a global interior solution to the cost minimization problem (2) or (2').

Result 3: *Under (CQ) and the quasi-concavity of the production function $g(\mathbf{x})$, the FONC (4a)-(4b) are necessary and sufficient to identify $\mathbf{x}^c$ as an interior solution to the cost minimization problem.*

Note: Recall that a quasi-concave function is not always concave. Thus, the above result does *not require a concave production function*, i.e. it allows for a technology that does *not* exhibit diminishing marginal productivity. However, a concave function is always quasi-concave. Thus, as a special case, under (CQ) and diminishing marginal productivity, the FONC (4a)-(4b) are also necessary and sufficient to identify $\mathbf{x}^c$ as an interior solution to the cost minimization problem.

It can be shown that the quasi-concavity of the production function $g(\mathbf{x})$ implies the second order necessary condition (SONCmin):

$$[\mathbf{s}_b(\mathbf{x}_b^c)^T, \mathbf{I}_{n-1}]\, L_{\mathbf{xx}}(\mathbf{x}^c, \lambda^c) \begin{bmatrix} \mathbf{s}_b(\mathbf{x}_b^c) \\ \mathbf{I}_{n-1} \end{bmatrix} = \text{a symmetric, positive semi-definite matrix,}$$

where $\mathbf{x} = (\mathbf{x}_a, \mathbf{x}_b)$ with $\mathbf{x}_a = x_1$, $\mathbf{x}_b = (x_2, \ldots, x_n)$, and $\mathbf{s}_b = -[g_a]^{-1} g_b$. This means that choosing production points where the quasi-concavity property does not hold would be *inconsistent* with cost minimization. (Note that, since a concave function is always quasi-concave, such points would necessarily be in a region where diminishing marginal productivity does not hold).

What about the second order sufficient condition (SOSCmin, where the matrix $[\mathbf{s}_b(\mathbf{x}_b^c)^T, \mathbf{I}_{n-1}]\, L_{\mathbf{xx}}(\mathbf{x}^c, \lambda^c) \begin{bmatrix} \mathbf{s}_b(\mathbf{x}_b^c) \\ \mathbf{I}_{n-1} \end{bmatrix}$ is now positive-definite)? In general, SONCmin does not imply SOSCmin. Given **result 3**, it may seem that we do not need the SOSCmin. This is true if we are just interested in identifying the cost minimizing *level* $\mathbf{x}^c(y, \mathbf{w})$. However, if we also want to investigate the *properties* of $\mathbf{x}^c(y, \mathbf{w})$, then strengthening SONCmin into SOSCmin can prove useful.

7.3 Properties of the cost minimizing input demand functions $\mathbf{x}^c(y, \mathbf{w})$

As stated in result 3, under (CQ) and the quasi-concavity of the production function $g(\mathbf{x})$, the FONC (4a)-(4b) are necessary and sufficient to identify $\mathbf{x}^c$ as an interior solution to the cost minimization problem. Yet, this solution may not be unique. And even if it is unique, $\mathbf{x}^c(y, \mathbf{w})$ may not be differentiable functions. In this section, we will consider the convenient case where the decision rules $\mathbf{x}^c(y, \mathbf{w})$ are *differentiable*. This will allow us to make use of calculus tools and to discuss behavioral properties in terms of the derivatives $\frac{\partial \mathbf{x}^c}{\partial \mathbf{w}}$ and $\frac{\partial \mathbf{x}^c}{\partial y}$ measuring the marginal impacts of changes of $(\mathbf{w}, y)$ on cost minimizing behavior $\mathbf{x}^c$.

Let $\mathbf{x}^c(y, \mathbf{w}) = (x_1^c(y, \mathbf{w}), \ldots, x_n^c(y, \mathbf{w}))$ and $\lambda^c(y, \mathbf{w})$ be the solution to the system of equations given by FONC (4a)-(4b). Consider applying the *implicit function theorem* to the FONC (4a)-(4b): $\frac{\partial L}{\partial(\mathbf{x},\lambda)} = 0$. We can use the implicit function theorem if the $(n+m)\times(n+m)$ *Hessian matrix* $\frac{\partial^2 L}{\partial(\mathbf{x},\lambda)^2}$ evaluated at $(\mathbf{x}^c, \lambda^c)$ is *non-singular*. This matrix is called the *bordered Hessian* matrix and is denoted by $\mathbf{H}$,

$$\mathbf{H} \equiv \frac{\partial^2 L}{\partial(\mathbf{x},\lambda)^2} \equiv \begin{bmatrix} \frac{\partial^2 L}{\partial x_1^2} & \frac{\partial^2 L}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 L}{\partial x_1 \partial \lambda} \\ \frac{\partial^2 L}{\partial x_2 \partial x_1} & \frac{\partial^2 L}{\partial x_2^2} & \cdots & \frac{\partial^2 L}{\partial x_2 \partial \lambda} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 L}{\partial \lambda \partial x_1} & \frac{\partial^2 L}{\partial \lambda \partial x_2} & \cdots & \frac{\partial^2 L}{\partial \lambda^2} \end{bmatrix} \equiv \begin{bmatrix} \frac{\partial^2 L}{\partial x_1^2} & \frac{\partial^2 L}{\partial x_1 \partial x_2} & \cdots & -\frac{\partial g}{\partial x_1} \\ \frac{\partial^2 L}{\partial x_2 \partial x_1} & \frac{\partial^2 L}{\partial x_2^2} & \cdots & -\frac{\partial g}{\partial x_2} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{\partial g}{\partial x_1} & -\frac{\partial g}{\partial x_2} & \cdots & 0 \end{bmatrix} \text{ evaluated at } (\mathbf{x}^c, \lambda^c).$$

(Note that, under twice continuous differentiability of $g(\mathbf{x})$, the $(n+m)\times(n+m)$ bordered Hessian matrix $\mathbf{H}$ is *symmetric* from Young theorem). From the implicit function theorem, $\det(\mathbf{H}) \neq 0$ is a sufficient condition for $\mathbf{x}^c(y, \mathbf{w})$ and $\lambda^c(y, \mathbf{w})$ to be continuously differentiable functions. We will exploit this condition below.

In general the second order necessary conditions (SONCmin) stated above do *not* imply that $\det(\mathbf{H}) \neq 0$. However the second order sufficient condition (SOSCmin)

$$[\mathbf{s}_b(\mathbf{x}_b^c)^T, \mathbf{I}_{n-1}]\, L_{\mathbf{xx}}(\mathbf{x}^c, \lambda^c) \begin{bmatrix} \mathbf{s}_b(\mathbf{x}_b^c) \\ \mathbf{I}_{n-1} \end{bmatrix} = \text{a symmetric, positive definite matrix,}$$

does imply that $\mathbf{H}$ *is non-singular*, i.e. that $\det(\mathbf{H}) \neq 0$ (e.g., see the determinant rule for checking the SOSCmin). This means that, if we want to apply the implicit function theorem to the FONC (4a)-(4b) to investigate the *derivative properties* of $\mathbf{x}^c(y, \mathbf{w})$, we *need to strengthen the SONCmin into the SOSCmin*.

Note: In general, the second order sufficient condition (SOSCmin)

$$[\mathbf{s}_b(\mathbf{x}_b^c)^T, \mathbf{I}_{n-1}]\, L_{\mathbf{xx}}(\mathbf{x}^c, \lambda^c) \begin{bmatrix} \mathbf{s}_b(\mathbf{x}_b^c) \\ \mathbf{I}_{n-1} \end{bmatrix} = \text{a symmetric, positive definite matrix,}$$

implies that the production function $g(\mathbf{x})$ is *locally strictly quasi-concave*.

To proceed with the analysis of cost minimizing behavior, substitute $\mathbf{x}^c(y, \mathbf{w})$ and $\lambda^c(y, \mathbf{w})$ into the FONC to give

$$L_{\mathbf{x}}(\mathbf{x}^c(y,\mathbf{w}), \lambda^c(y,\mathbf{w}), \mathbf{w}, y) \equiv \mathbf{w} - \lambda \cdot g_{\mathbf{x}}(\mathbf{x}^c(y,\mathbf{w})) = 0, \tag{5a}$$

$$L_{\lambda}(\mathbf{x}^c(y,\mathbf{w}), \lambda^c(y,\mathbf{w}), \mathbf{w}, y) \equiv g(\mathbf{x}^c(y,\mathbf{w})) = 0, \tag{5b}$$

where $L(\mathbf{x}, \lambda, y, \mathbf{w}) \equiv \mathbf{w}^T\mathbf{x} + \lambda \cdot (y - g(\mathbf{x}))$. Differentiating (5a) and (5b) with respect to $\alpha = (\mathbf{w}, y)$ and using the chain rule gives

$$L_{\mathbf{xx}}\mathbf{x}_\alpha^c + L_{\mathbf{x}\lambda}\lambda_\alpha^c + L_{\mathbf{x}\alpha} = 0,$$

$$L_{\lambda\mathbf{x}}\mathbf{x}_\alpha^c + L_{\lambda\lambda}\lambda_\alpha^c + L_{\lambda\alpha} = 0,$$

Where $L_{\mathbf{xx}} = \frac{\partial^2 L}{\partial \mathbf{x}^2}$ is a $(n\times n)$ matrix, $L_{\lambda\mathbf{x}} = (L_{\mathbf{x}\lambda})^T = \frac{\partial^2 L}{\partial\lambda\partial\mathbf{x}}$ is a $(1\times n)$ vector, $L_{\lambda\lambda} = 0$, $\mathbf{x}_\alpha^c = (\mathbf{x}_\mathbf{w}^c, \mathbf{x}_y^c) = \frac{\partial \mathbf{x}^c}{\partial(\mathbf{w},y)}$ is a $n\times(n+1)$ matrix, and $\lambda_\alpha^c = (\lambda_\mathbf{w}^c, \lambda_y^c) = \frac{\partial \lambda^c}{\partial(\mathbf{w},y)}$ is a $1\times(n+1)$ vector, all expressions evaluated at $\mathbf{x}^c$, λ^c. Using the definition of the Hessian matrix $\mathbf{H}$, this can be written more compactly as

$$\mathbf{H} \cdot \begin{bmatrix} \frac{\partial \mathbf{x}^c}{\partial \mathbf{w}} & \frac{\partial \mathbf{x}^c}{\partial y} \\ \frac{\partial \lambda^c}{\partial \mathbf{w}} & \frac{\partial \lambda^c}{\partial y} \end{bmatrix} = -\begin{bmatrix} \frac{\partial^2 L}{\partial \mathbf{x}\partial \mathbf{w}} & \frac{\partial^2 L}{\partial \mathbf{x}\partial y} \\ \frac{\partial^2 L}{\partial \lambda\partial \mathbf{w}} & \frac{\partial^2 L}{\partial \lambda\partial y} \end{bmatrix},$$

or

$$\mathbf{H} \cdot \begin{bmatrix} \mathbf{x}_\boldsymbol{\alpha}^c \\ \lambda_\boldsymbol{\alpha}^c \end{bmatrix} = -\begin{bmatrix} L_{\mathbf{x}\boldsymbol{\alpha}} \\ L_{\lambda\boldsymbol{\alpha}} \end{bmatrix}, \tag{6}$$

where $\boldsymbol{\alpha} = (w_1, \ldots, w_n, y)$. *Under SOSCmin*, $\det(\mathbf{H}) \neq 0$ and equation (6) can be written as

$$\begin{bmatrix} \mathbf{x}_\boldsymbol{\alpha}^c \\ \lambda_\boldsymbol{\alpha}^c \end{bmatrix} = -\mathbf{H}^{-1} \cdot \begin{bmatrix} L_{\mathbf{x}\boldsymbol{\alpha}} \\ L_{\lambda\boldsymbol{\alpha}} \end{bmatrix}, \tag{7a}$$

which are the *classical comparative static results*. Note that this result can be obtained simply by applying the implicit function theorem to FONC in (5a)-(5b). In our cost minimization problem, $L = \mathbf{w} \cdot \mathbf{x} + \lambda\,(y - g(\mathbf{x}))$, implying that $L_{\mathbf{xw}} = \mathbf{I}_n$ = an identity matrix of dimension n, $L_{\mathbf{x}y} = 0$, $L_{\lambda\mathbf{w}} = 0$, and $L_{\lambda y} = 1$. It follows that $\begin{bmatrix} L_{\mathbf{x}\boldsymbol{\alpha}} \\ L_{\lambda\boldsymbol{\alpha}} \end{bmatrix} = \mathbf{I}_{n+1}$, implying from (7a) that

$$\begin{bmatrix} \mathbf{x}_\boldsymbol{\alpha}^c \\ \lambda_\boldsymbol{\alpha}^c \end{bmatrix} \equiv \begin{bmatrix} \mathbf{x}_\mathbf{w}^c & \mathbf{x}_y^c \\ \lambda_\mathbf{w}^c & \lambda_y^c \end{bmatrix} = -\mathbf{H}^{-1}. \tag{7b}$$

Equation (7b) reflects all the implications of cost minimizing behavior for the input demand functions $\mathbf{x}^c$*, and for* λ^c. However, it is in a form that is not very convenient either for empirical work, or for intuitive reasoning.

7.4 Symmetry Restrictions for $\mathbf{x}^c(y, \mathbf{w})$

We know that the matrix $\mathbf{H}$ is symmetric (from Young theorem, $g(\mathbf{x})$ being assumed twice continuously differentiable). This means that $\mathbf{H}^{-1}$ is also symmetric. Then, equation (7b) implies the following important result:

The (n+1)×(n+1) matrix $\begin{bmatrix} \mathbf{x}^c_{\mathbf{w}} & \mathbf{x}^c_y \\ \lambda^c_{\mathbf{w}} & \lambda^c_y \end{bmatrix}$ *is* ***symmetric***.

- First, this implies that the $(n\times n)$ matrix $\mathbf{x}^c_{\mathbf{w}} \equiv \begin{bmatrix} \frac{\partial x_1^c}{\partial w_1} & \frac{\partial x_1^c}{\partial w_2} & \cdots & \frac{\partial x_1^c}{\partial w_n} \\ \frac{\partial x_2^c}{\partial w_1} & \frac{\partial x_2^c}{\partial w_2} & \cdots & \frac{\partial x_2^c}{\partial w_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial x_n^c}{\partial w_1} & \frac{\partial x_n^c}{\partial w_2} & \cdots & \frac{\partial x_n^c}{\partial w_n} \end{bmatrix}$ is symmetric. This generates $n\times(n-1)/2$ *symmetry restrictions on the effects of prices on cost minimizing input demand functions*: $\frac{\partial x_i^c}{\partial w_j} = \frac{\partial x_j^c}{\partial w_i}$, for all $i \neq j$. *These symmetry restrictions are implied by cost minimization*. This means that finding empirical evidence *against* these symmetry restrictions is sufficient to conclude that the associated decision rules (choosing input choices conditionally on output y) are *inconsistent with cost minimizing behavior*.

- Second, this implies that $\mathbf{x}_y{}^c = (\lambda_{\mathbf{w}}{}^c)^T$, i.e. that $\frac{\partial x_i^c}{\partial y} = \frac{\partial \lambda^c}{\partial w_i}$, $i = 1, \ldots, n$. This result will prove useful later in the class, after we provide an economic interpretation of the Lagrange multiplier λ^c.

7.5 Constraint restrictions

Under assumption A1, we know that the cost minimizing solution $\mathbf{x}^c(y, \mathbf{w})$ must necessarily satisfy the technical efficiency constraint: $y = g(\mathbf{x})$. Thus, $y = g(\mathbf{x}^c(y, \mathbf{w}))$ always. Differentiating this expression with respect to $(\mathbf{w}, y)$ and using the chain rule, we obtain

$$\frac{\partial g}{\partial \mathbf{x}} \frac{\partial \mathbf{x}^c}{\partial \mathbf{w}} = 0,$$

and

$$\frac{\partial g}{\partial \mathbf{x}} \frac{\partial \mathbf{x}^c}{\partial y} = 1,$$

or equivalently,

$$g_{\mathbf{x}}(\mathbf{x}^c)\, \mathbf{x}_{\mathbf{w}}{}^c = 0, \tag{8a}$$

and

$$g_{\mathbf{x}}(\mathbf{x}^c)\, \mathbf{x}_y{}^c = 1. \tag{8b}$$

Equations (8a) and (8b) are implied by cost minimization under assumption A1. Thus, under A1, finding evidence against these relationships can be interpreted as evidence against technical efficiency, and thus against cost minimizing behavior.

Equation (8a) can be written as $\sum_i \frac{\partial g}{\partial x_i} \frac{\partial x_i^c}{\partial w_j} = 0$, for $j = 1, \ldots, n$. It states that the weighted sum of the effects of price w_j across all inputs i (with $\frac{\partial g}{\partial x_i}$ as weights) must be zero. With $\frac{\partial g}{\partial x_i} > 0$, it implies that, if some of these price effects are negative, others must be positive.

Equation (8b) can be written as $\sum_i \frac{\partial g}{\partial x_i}\frac{\partial x_i^c}{\partial y} = 1$. It implies that the weighted sum of the output effects $\frac{\partial x_i^c}{\partial y}$ (with $\frac{\partial g}{\partial x_i}$ as weight) must equal one. This does not rule out the possibility that $\frac{\partial x_i^c}{\partial y}$ can be negative for some i. However, it indicates that *on average* these output effects $\frac{\partial x_i^c}{\partial y}$ are expected to be positive: producing a larger output y tends to require more inputs *on the average*.

7.6 Sign restrictions for $\mathbf{x}^c$

Alternative characterizations of SOSCmin:
Recall that, for our cost minimization problem, the SOSCmin involves the positive definiteness of the matrix

$$[\mathbf{s}_b(\mathbf{x}_b^c)^T, \mathbf{I}_{n-1}]\cdot L_{\mathbf{xx}}(\mathbf{x}^c,\lambda^c)\cdot\begin{bmatrix}\mathbf{s}_b(\mathbf{x}_b^c)\\ \mathbf{I}_{n-1}\end{bmatrix},$$

where $\mathbf{x} = (\mathbf{x}_a, \mathbf{x}_b)$ with $\mathbf{x}_a = x_1$, $\mathbf{x}_b = (x_2, \ldots, x_n)$, and $\mathbf{s}_b = -[g_a]^{-1} g_b$, with $g_a \equiv \frac{\partial g(\mathbf{x}^c)}{\partial \mathbf{x}_a}$ being a scalar, and $g_b \equiv \frac{\partial g(\mathbf{x}^c)}{\partial \mathbf{x}_b}$ being a $1\times(n-1)$ vector evaluated at $(\mathbf{x}^c, \lambda^c)$. Thus, SOSCmin can be expressed as

$$\mathbf{v}_b^T\cdot[-g_b^T\cdot g_a^{-1}, \mathbf{I}_{n-1}]\cdot L_{\mathbf{xx}}(\mathbf{x}^c,\lambda^c)\cdot\begin{bmatrix}-g_a^{-1}\cdot g_b\\ \mathbf{I}_{n-1}\end{bmatrix}\cdot\mathbf{v}_b > 0 \text{ for all } \mathbf{v}_b \neq 0,$$

where $\mathbf{v}_b$ is a $(n-1)\times 1$ vector.

Define $\mathbf{v}_a \equiv -\mathbf{v}_b^T g_b^T (g_a^T)^{-1}$, (which can be written as $g_a \mathbf{v}_a + g_b \mathbf{v}_b = 0$, or $g_{\mathbf{x}}\cdot\begin{bmatrix}\mathbf{v}_a\\ \mathbf{v}_b\end{bmatrix} = 0$). Then, SOSCmin becomes

$$[\mathbf{v}_a^T, \mathbf{v}_b^T]\cdot L_{\mathbf{xx}}(\mathbf{x}^c,\lambda^c)\cdot\begin{bmatrix}\mathbf{v}_a\\ \mathbf{v}_b\end{bmatrix} > 0 \text{ for all } \mathbf{v}_b \neq 0 \text{ satisfying } g_a \mathbf{v}_a + g_b \mathbf{v}_b = 0. \tag{9a}$$

Using the definition of the *bordered Hessian* $\mathbf{H} \equiv \frac{\partial^2 L}{\partial(\mathbf{x},\lambda)^2}\begin{bmatrix}L_{\mathbf{xx}} & -(g_{\mathbf{x}})^T\\ -g_{\mathbf{x}} & 0\end{bmatrix}$ (evaluated at $(\mathbf{x}^c, \lambda^c)$), the SOSCmin (9a) can be alternatively written as

$$[\mathbf{v}_a^T, \mathbf{v}_b^T, \mathbf{v}_c^T]\cdot\mathbf{H}\cdot\begin{bmatrix}\mathbf{v}_a\\ \mathbf{v}_b\\ \mathbf{v}_c\end{bmatrix} > 0 \text{ for all } \mathbf{v} = \begin{bmatrix}\mathbf{v}_a\\ \mathbf{v}_b\\ \mathbf{v}_c\end{bmatrix} \text{ such that } \mathbf{v}_b \neq 0 \text{ and } g_{\mathbf{x}}\cdot\begin{bmatrix}\mathbf{v}_a\\ \mathbf{v}_b\end{bmatrix} = 0, \tag{9b}$$

where $\mathbf{v}_c$ is a scalar. This alternative expression for SOSCmin can be interpreted in terms of the "positive definiteness of the bordered Hessian $\mathbf{H}$, *subject to constraint*, $g_{\mathbf{x}}\cdot\begin{bmatrix}\mathbf{v}_a\\ \mathbf{v}_b\end{bmatrix} = 0$," where $g_{\mathbf{x}} = (g_a, g_b)$ is a $(1\times n)$ vector.

The characterization of the SOSCmin given in (9b) will be useful to obtain the following important result:

The (n×n) matrix $\mathbf{x}^c_{\mathbf{w}} = \frac{\partial \mathbf{x}^c}{\partial \mathbf{w}}$ *is symmetric negative semi-definite.*

Proof: The proof of symmetry was provided earlier. To prove the negative semi-definiteness, choose $\begin{bmatrix} \mathbf{v}_a \\ \mathbf{v}_b \\ \mathbf{v}_c \end{bmatrix} = \begin{bmatrix} \mathbf{x}^c_{\mathbf{w}} \\ \lambda^c_{\mathbf{w}} \end{bmatrix} \cdot \mathbf{u}$ in (9b), where $\mathbf{u}$ is a (n×1) vector. Note from (8a) that this choice always satisfies the constraint $g_{\mathbf{x}} \cdot \begin{bmatrix} \mathbf{v}_a \\ \mathbf{v}_b \end{bmatrix} = 0$ for all $\mathbf{u}$. Then, expression (9b) implies

$$\mathbf{u}^T \cdot [(\mathbf{x}^c_{\mathbf{w}})^T, (\lambda^c_{\mathbf{w}})^T] \cdot \mathbf{H} \cdot \begin{bmatrix} \mathbf{x}^c_{\mathbf{w}} \\ \lambda^c_{\mathbf{w}} \end{bmatrix} \cdot \mathbf{u} \geq 0 \text{ for all } \mathbf{u}.$$

But $[(\mathbf{x}_{\mathbf{w}}{}^c)^T\ (\lambda_{\mathbf{w}}{}^c)^T]\ \mathbf{H} = [-\mathbf{I}_n, 0]$ from (7b). It follows that

$$\mathbf{u}^T\ \mathbf{x}_{\mathbf{w}}\ \mathbf{u} \leq 0 \text{ for all } \mathbf{u},$$

implying that $\mathbf{x}_{\mathbf{w}}$ is a negative semi-definite matrix.

The negative semi-definiteness of $\mathbf{x}_{\mathbf{w}}$ implies that $\frac{\partial x_i^c}{\partial w_i} \leq 0$, for $i = 1, \ldots, n$. This gives the intuitive result that *cost-minimizing input demand functions are downward sloping*. It means that, under cost minimization, increasing an input price provides a disincentive for the firm to use the corresponding input. This is implied by cost minimization. It follows that any evidence indicating that input price has a positive effect on input demand (conditional on output y) would be sufficient to conclude that the associated decision rule is *inconsistent* with cost minimization.

7.7 The homogeneity property

Recall our cost minimization problem

$$\underset{\mathbf{x}}{Min} \quad \{\mathbf{w} \cdot \mathbf{x} : (y, \mathbf{x}) \in \mathbf{F}\},$$

with $\mathbf{x}^c(y, \mathbf{w})$ as solution. Now, consider the modified problem where all prices are rescaled by a constant $k > 0$,

$$\underset{\mathbf{x}}{Min} \quad \{k \cdot \mathbf{w} \cdot \mathbf{x} : (y, \mathbf{x}) \in \mathbf{F}\},$$

which has for solution $\mathbf{x}^c(y, k\,\mathbf{w})$. Note this modified problem can be alternatively written as

$$k\,[\underset{\mathbf{x}}{Min} \quad \{k \cdot \mathbf{w} \cdot \mathbf{x} : (y, \mathbf{x}) \in \mathbf{F}\}],$$

which has for solution $\mathbf{x}^c(y, \mathbf{w})$. It follows that

$$\mathbf{x}^c(y, k\,\mathbf{w}) = \mathbf{x}^c(y, \mathbf{w}) \text{ for any } k > 0.$$

This shows that cost minimizing decision rules $\mathbf{x}^c(y, \mathbf{w})$ are *homogeneous of degree zero in input prices* $\mathbf{w}$. It means that *cost-minimizing behavior is invariant to a proportional change in all input prices* (i.e., it responds only to changes in *relative prices*).

This homogeneity property imposes restrictions on behavior. To see that, under the continuous differentiability of $\mathbf{x}^c(y, \mathbf{w})$, apply *Euler theorem* to the functions $\mathbf{x}^c(y, \mathbf{w})$ that are homogeneous of degree zero in $\mathbf{w}$. This gives:

$$\frac{\partial \mathbf{x}^c}{\partial \mathbf{w}}\ \mathbf{w} = 0, \tag{10}$$

where $\mathbf{w}$ is a (n×1) vector, or

$$\sum_j \frac{\partial x_i^c}{\partial w_j} w_j = 0,\ i = 1, \ldots, n,$$

or, for $x_i^* > 0$,

$$\sum_j \frac{\partial x_i^c}{\partial w_j} \frac{w_j}{x_i^c} = 0,\ i = 1, \ldots, n,$$

or

$$\sum_j \frac{\partial \ln(x_i^c)}{\partial \ln(w_j)} = 0,\ i = 1, \ldots, n,$$

where $\frac{\partial \ln(x_i^c)}{\partial \ln(w_j)} = \frac{\partial x_i^c}{\partial w_j} \frac{w_j}{x_i^c}$ is the elasticity of x_i^c with respect to the j-th input price w_j. (Note: A price elasticity can be conveniently interpreted as measuring the percentage change in a quantity due to a one percent change in price).

It follows that, *for each input demand, the sum of price elasticities across all input prices* **w** *must be zero*. This generates *n homogeneity restrictions* (one for each input demand), as implied by cost minimization.

7.8 The two input case (n = 2)

Here, we consider the special case of a two-input firm, where n = 2. Of course, all the general results obtained above apply: downward sloping input demand, symmetry restrictions, homogeneity restrictions, etc. However, the two input case gives additional results.

When $n = 2$, $g_{\mathbf{xx}} = \frac{\partial^2 g}{\partial \mathbf{x}^2} = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix}$, where $g_{ij} \equiv \frac{\partial^2 g}{\partial x_i \partial x_j}$. Under twice continuous differentiability of $g(\mathbf{x})$, $g_{12} = g_{21}$ (from Young theorem). Then, the bordered Hessian is

$\mathbf{H} = \begin{bmatrix} -\lambda g_{11} & -\lambda g_{12} & -g_1 \\ -\lambda g_{12} & -\lambda g_{22} & -g_2 \\ -g_1 & -g_2 & 0 \end{bmatrix}$, where $\det(\mathbf{H}) = \lambda\, [g_{11}\, (g_2)^2 + g_{22}\, (g_1)^2 - 2 g_{12}\, g_1\, g_2]$. Note that

$$\mathbf{H}^{-1} = \frac{1}{\det(\mathbf{H})} \cdot \begin{bmatrix} -(g_2)^2 & g_1 g_2 & \lambda\,(g_{12} g_2 - g_{22} g_1) \\ g_1 g_2 & -(g_1)^2 & -\lambda\,(g_{11} g_2 - g_{12} g_1) \\ \lambda\,(g_{12} g_2 - g_{22} g_1) & -\lambda\,(g_{11} g_2 - g_{12} g_1) & \lambda^2\,(g_{11} g_{22} - (g_{12})^2) \end{bmatrix}.$$

(Check that $\mathbf{H}\,\mathbf{H}^{-1} = \mathbf{I}_3$).

From (7b), we obtain

$$\frac{\partial x_1^c}{\partial w_1} = \frac{1}{\det(\mathbf{H})} (g_2)^2 < 0,$$

which is negative as expected, since $\det(\mathbf{H}) < 0$ from *SOSCmin*.

Similarly, from (7b), we have

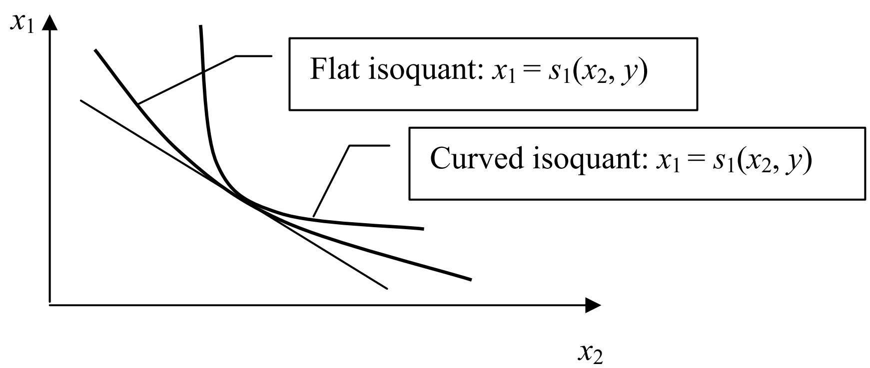
$$\frac{\partial x_1^c}{\partial w_2} = -\frac{1}{\det(\mathbf{H})} g_1 g_2 > 0,$$

since $g_i > 0$ under positive prices. This implies that $\frac{\partial x_1^c}{\partial w_2}$ ($= \frac{\partial x_2^c}{\partial w_1}$ by symmetry) is always positive in the two input case.

Note: This result is specific to the two-input case and does not necessarily hold if there are more than two inputs. Indeed, with $n = 2$, the homogeneity restriction (10) gives $\frac{\partial x_1^c}{\partial w_1} w_1 + \frac{\partial x_1^c}{\partial w_2} w_2 = 0$. Given $\frac{\partial x_1^c}{\partial w_1} < 0$, and positive prices **w**, the *homogeneity restriction* implies that $\frac{\partial x_1^c}{\partial w_2} = -\frac{\partial x_1^c}{\partial w_1} \frac{w_1}{w_2} > 0$. Such a result does not necessarily follow when $n > 2$.

Here, det(**H**) is close to zero if the isoquant $x_1 = s_1(x_2, y)$ is "flat." Alternatively, |det(**H**)| can become large when the isoquant is "curved." With det(**H**) being in the denominator in the above expressions for $\frac{\partial x_i^c}{\partial w_j}$, this suggests that the effects of input prices on cost minimizing input demand $\mathbf{x}^c$ tend to be small under "curved isoquant", but can become large under "flat isoquant." See figure 7.2. Such price effects thus depend on the nature of the underlying technology. We will revisit this issue later.

Figure 7.2



7.9 The general case

So far, we have limited our analysis to the case of a firm producing a single output. We now consider the general case where a multi-output firm produces *m outputs using n inputs*. Letting $\mathbf{y} = (y_1, \ldots, y_m)$ be the vector of outputs, the production technology is given by the feasible set $(\mathbf{y}, \mathbf{x}) \in \mathbf{F}$. Then, cost-minimizing behavior for a multi-output firm is given by

$$\underset{\mathbf{x}}{Min} \ \{\mathbf{w} \cdot \mathbf{x} : (\mathbf{y}, \mathbf{x}) \in \mathbf{F}\}$$, which has for solution $\mathbf{x}^c(\mathbf{y}, \mathbf{w})$.

In this context, all the properties we derived above still apply.
In particular, $\mathbf{x}^c(\mathbf{y}, \mathbf{w})$ is homogeneous of degree zero in prices **w** (implying that only relative prices matter).
And, if differentiable, the cost minimizing input demand functions satisfy the property:

$\mathbf{x}^c_{\mathbf{w}} = \frac{\partial \mathbf{x}^c}{\partial \mathbf{w}}$ is a $(n \times n)$ symmetric, negative semi-definite matrix,

which implies that

- the input demand functions are *downward sloping*: $\frac{\partial x_i^c}{\partial w_i} \leq 0$, $i = 1, \ldots, n$,
- $n\times(n-1)/2$ *symmetry restrictions* hold: $\frac{\partial x_i^c}{\partial w_j} = \frac{\partial x_j^c}{\partial w_i}$, for all $i \neq j$.

In general, the output effects $\frac{\partial x_i^c}{\partial \mathbf{y}}$ can be positive or negative: they depend on the nature of the underlying technology. We will revisit this issue later.