

Lecture 7: More about Cost Functions

7.1 Summary of properties of cost functions

Let f be a monotonic production function. The associated cost function $c(w, y)$ is

- continuous
- concave in w
- monotone nondecreasing in (w, y)
- homogeneous of degree one in w, that is, $c(\lambda w, y) = \lambda c(w, y)$ for $\lambda > 0$.

Moreover, if $\hat{x}(w, y)$ is the **conditional factor demand**, then

$$\frac{\partial c(w, y)}{\partial w_i} = \hat{x}_i(w, y).$$

7.2 Cost minimization

Mathematically the cost minimization problem can be formulated as follows.

$$\underset{x}{\text{minimize}} \;\; w \cdot x \quad \text{subject to } f(x) \geqslant y, \;\; x \geqq 0,$$

where $w \gg 0$ and $y > 0$.

It is clear that if f is monotonic, we may replace the condition $f(x) \geqslant y$ by $f(x) - y = 0$ without changing the solution. Let $\hat{x}(w, y)$ solve this problem, and assume that $\hat{x} \gg 0$. The Lagrangean for this minimization problem is

$$w \cdot x - \lambda\Big(f(x) - y\Big).$$

The gradient of the constraint function (with respect to x) is just $f'(\hat{x})$, which is not zero. Therefore by the Lagrange Multiplier Theorem, there is a Lagrange multiplier $\hat{\lambda}$ (depending on w, y) so that locally the first order conditions

$$w_i - \hat{\lambda}(w, y) f_i\Big(\hat{x}(w, y)\Big) = 0, \quad i = 1, \dots, n, \tag{1}$$

where $f_i(x) = \frac{\partial f(x)}{\partial x_i}$, and the constraint

$$y - f\Big(\hat{x}(w, y)\Big) = 0 \tag{2}$$

hold for all w, y. Note that (1) implies that $\hat\lambda > 0$.

The second order condition is that

$$\hat\lambda \sum_{i=1}^{n}\sum_{j=1}^{n} f_{ij}(\hat x) v_i v_j \leqslant 0, \tag{3}$$

for all $v \in \boldsymbol{R}^{\mathrm{n}}$ satisfying

$$f'(\hat x)\cdot v = \sum_{i=1}^{n} f_i(\hat x) v_i = 0.$$

Using the **method of implicit differentiation** with respect to each w_j on (1) yields:

$$\delta_{ij} - \frac{\partial\hat\lambda}{\partial w_j} f_i(\hat x) - \hat\lambda \sum_{k=1}^{n} f_{ik}(\hat x)\frac{\partial \hat x_k}{\partial w_j} = 0, \qquad \begin{matrix} i = 1,\dots,n \\ j = 1,\dots,n \end{matrix}, \tag{4}$$

where δ_{ij} is the Kronecker delta,

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j. \end{cases}$$

Differentiating (1) with respect to y yields

$$-\frac{\partial\hat\lambda}{\partial y} f_i(\hat x) - \hat\lambda \sum_{k=1}^{n} f_{ik}(\hat x)\frac{\partial \hat x_k}{\partial y} = 0, \quad i = 1,\dots,n, \tag{5}$$

Now differentiate (2) with respect to each w_j to get

$$-\sum_{k=1}^{n} f_k(\hat x)\frac{\partial \hat x_k}{\partial w_j} = 0, \quad j = 1,\dots,n, \tag{6}$$

and with respect to y to get

$$-\sum_{k=1}^{n} f_k(\hat x)\frac{\partial \hat x_k}{\partial y} + 1 = 0. \tag{7}$$

We can rearrange equations (4) through (7) into one gigantic matrix equation:

$$\begin{bmatrix} \hat\lambda f_{11} & \dots & \hat\lambda f_{1n} & f_1 \\ \vdots & & \vdots & \vdots \\ \vdots & & \vdots & \vdots \\ \hat\lambda f_{n1} & \dots & \hat\lambda f_{nn} & f_n \\ f_1 & \dots & f_n & 0 \end{bmatrix} \begin{bmatrix} \frac{\partial \hat x_1}{\partial w_1} & \cdots & \frac{\partial \hat x_1}{\partial w_n} & \frac{\partial \hat x_1}{\partial y} \\ \vdots & & \vdots & \vdots \\ \vdots & & \vdots & \vdots \\ \frac{\partial \hat x_n}{\partial w_1} & \dots & \frac{\partial \hat x_n}{\partial w_n} & \frac{\partial \hat x_n}{\partial y} \\ \frac{\partial \hat\lambda}{\partial w_1} & \dots & \frac{\partial \hat\lambda}{\partial w_n} & \frac{\partial \hat\lambda}{\partial y} \end{bmatrix} = \begin{bmatrix} 1 & 0 & \dots\dots & 0 & 0 \\ 0 & \ddots & & \vdots & \vdots \\ \vdots & & \ddots & 0 & \vdots \\ 0 & \dots\dots & 0 & 1 & 0 \\ 0 & \dots & \dots & 0 & 1 \end{bmatrix}. \tag{8}$$

To see where this comes from, break up the $(n{+}1)\times(n{+}1)$ matrix equation into four blocks. The upper left $n \times n$ block comes from (4). The upper right $n \times 1$

$$\left[\begin{array}{c|c} \ddots & \vdots \\ (4) & (5) \\ \ddots & \vdots \\ \hline \dots (6) \dots & (7) \end{array}\right]$$

Figure 7.1. The blocks in the matrix version of equations (4) through (7).

block comes from (5). The lower left $1 \times n$ block comes from (6), and finally the lower right 1×1 block is just (7). This tells us is that

$$\begin{bmatrix} \frac{\partial \hat{x}_1}{\partial w_1} & \cdots & \frac{\partial \hat{x}_1}{\partial w_n} & \frac{\partial \hat{x}_1}{\partial y} \\ \vdots & & \vdots & \vdots \\ \vdots & & \vdots & \vdots \\ \frac{\partial \hat{x}_n}{\partial w_1} & \cdots & \frac{\partial \hat{x}_n}{\partial w_n} & \frac{\partial \hat{x}_n}{\partial y} \\ \frac{\partial \hat{\lambda}}{\partial w_1} & \cdots & \frac{\partial \hat{\lambda}}{\partial w_n} & \frac{\partial \hat{\lambda}}{\partial y} \end{bmatrix} = \begin{bmatrix} \hat{\lambda} f_{11} & \dots & \hat{\lambda} f_{1n} & f_1 \\ \vdots & & \vdots & \vdots \\ \vdots & & \vdots & \vdots \\ \hat{\lambda} f_{n1} & \dots & \hat{\lambda} f_{nn} & f_n \\ f_1 & \dots & f_n & 0 \end{bmatrix}^{-1}. \tag{9}$$

So the second order conditions and Theorems 6.7.1 and 6.7.2 imply that the $n \times n$ matrix

$$\begin{bmatrix} \frac{\partial \hat{x}_1}{\partial w_1} & \cdots & \frac{\partial \hat{x}_1}{\partial w_n} \\ \vdots & & \vdots \\ \frac{\partial \hat{x}_n}{\partial w_1} & \cdots & \frac{\partial \hat{x}_n}{\partial w_n} \end{bmatrix}$$

is negative semidefinite of rank $n-1$, being the upper left block of the inverse of a bordered matrix that is negative definite under constraint. (See my notes on quadratic forms [2].) It follows therefore that

$$\frac{\partial \hat{x}_i}{\partial w_i} \leqslant 0 \quad i = 1, \dots, n.$$

Note that this approach provides us with conditions under which the cost function is twice continuously differentiable. It follows from (9) that if the bordered Hessian is invertible, the Implicit Function Theorem tells us that $\hat{x}$ and $\hat{\lambda}$ are C^1 functions of w and y (since f is C^2). On the other hand, if $\hat{x}$ and $\hat{\lambda}$ are C^1 functions of w and y, then (8) implies that the bordered Hessian is invertible. In either case, the marginal cost $\frac{\partial c}{\partial y} = \hat{\lambda}$, is a C^1 function of w and y, so the cost function is C^2, which is hard to establish by other means.

Returning now to (9), note that since the Hessian is a symmetric matrix, we have a number of **reciprocity** results. Namely:

$$\frac{\partial \hat{x}_i}{\partial w_j} = \frac{\partial \hat{x}_j}{\partial w_i} \qquad \begin{array}{l} i = 1, \dots, n \\ j = 1, \dots, n, \end{array}$$

and

$$\frac{\partial \hat{x}_i}{\partial y} = \frac{\partial \hat{\lambda}}{\partial w_i} = \frac{\partial^2 c}{\partial w_i \partial y}.$$

7.3 The marginal cost function

Define the cost function c by

$$c(w, y) = \sum_{k=1}^{n} w_k \hat{x}_k(w, y).$$

Then

$$\frac{\partial c(w, y)}{\partial y} = \sum_{i=k}^{n} w_k \frac{\partial \hat{x}_k(w, y)}{\partial y},$$

and

$$\frac{\partial^2 c(w, y)}{\partial y^2} = \sum_{i=k}^{n} w_k \frac{\partial^2 \hat{x}_k(w, y)}{\partial y^2}. \tag{10}$$

From (1), we have $w_k = \hat{\lambda} f_k(\hat{x})$, so

$$\frac{\partial c(w, y)}{\partial y} = \hat{\lambda} \sum_{k=1}^{n} f_i(\hat{x}) \frac{\partial \hat{x}_k(w, y)}{\partial y} = \hat{\lambda}, \tag{11}$$

where the second equality is just (7).

> That is, *the Lagrange multiplier $\hat{\lambda}$ is the marginal cost.*

Now let's see whether the marginal cost is increasing or decreasing as a function of y. Differentiating (7) with respect to y yields

$$\sum_{j=1}^{n} \left(\frac{\partial \hat{x}_j}{\partial y} \sum_{i=1}^{n} f_{ij}(\hat{x}) \frac{\partial \hat{x}_i}{\partial y} + f_j(\hat{x}) \frac{\partial^2 \hat{x}_j}{\partial y^2} \right) = 0,$$

or rearranging,

$$\sum_{j=1}^{n} f_j(\hat{x}) \frac{\partial^2 \hat{x}_j}{\partial y^2} = -\sum_{i=1}^{n} \sum_{j=1}^{n} f_{ij}(\hat{x}) \frac{\partial \hat{x}_i}{\partial y} \frac{\partial \hat{x}_j}{\partial y}. \tag{12}$$

From (10) and (1) we have that the left-hand side of (12) is $\frac{1}{\hat{\lambda}} \frac{\partial^2 c}{\partial y^2}$. What is the right-hand side?

Fix w and consider the curve $y \mapsto \hat{x}(y)$. This is called an **expansion path**. It traces out the optimal input combination as a function of the level of output. The tangent line to this curve at $\hat{x}$ is just $\{\hat{x} + \alpha v : \alpha \in \boldsymbol{R}\}$, where

$$v_i = \frac{\partial \hat{x}_i}{\partial y}.$$

Write the output along this tangent line, $f(\hat{x}+\alpha v)$, as a function $\hat{f}$ of α. That is, $\hat{f}(\alpha) = f(\hat{x}+\alpha v)$. By the chain rule,

$$\hat{f}'(\alpha) = \sum_{j=1}^{n} f_j(\hat{x}+\alpha v)v_j,$$

and

$$\hat{f}''(\alpha) = \sum_{i=1}^{n}\sum_{j=1}^{n} f_{ij}(\hat{x}+\alpha v)v_i v_j,$$

so

$$\hat{f}''(0) = \sum_{i=1}^{n}\sum_{j=1}^{n} f_{ij}(\hat{x})\frac{\partial \hat{x}_i}{\partial y}\frac{\partial \hat{x}_j}{\partial y}.$$

Thus (12) can be written as

$$\frac{\partial^2 c}{\partial y^2} = -\hat{\lambda}\hat{f}''(0).$$

In other words (12) asserts that *the slope of the marginal cost curve is increasing (that is, the cost function is a locally convex function of y) when the production function is locally concave on the line tangent to the expansion path*, and vice-versa.

7.4 Average cost and elasticity of scale

Recall that a production function f exhibits **constant returns to scale** if $f(\alpha x) = \alpha f(x)$ for all $\alpha > 0$. It exhibits **increasing returns to scale** if $f(\alpha x) > \alpha f(x)$ for $\alpha > 1$, and **decreasing returns to scale** if $f(\alpha x) < \alpha f(x)$ for $\alpha > 1$. If f is **homogeneous of degree k**, that is, if

$$f(\alpha x) = \alpha^k f(x),$$

then the returns to scale are decreasing, constant, or increasing, as $k < 1$, $k = 1$, or $k > 1$. Define

$$h(\alpha, x) = f(\alpha x).$$

The **elasticity of scale** $e(x)$ of the production function at x is defined to be

$$D_1 h(1, x)\frac{1}{f(x)} = f'(x)\cdot x/f(x),$$

where D_1 denotes the partial derivative with respect to the first argument α, and which Varian [3] writes as

$$\left.\frac{df(\alpha x)}{d\alpha}\frac{\alpha}{f(x)}\right|_{\alpha=1}.$$

If f is homogeneous of degree k, then $e(x) = k$, as

$$D_1 h(\alpha, x) = k\alpha^{k-1} f(x).$$

Even if f is not homogeneous, following Varian, we can express the elasticity of scale in terms of the marginal and average cost functions, at least for points x that minimize cost uniquely for some (y, w):

$$\begin{aligned}
e\Big(\hat{x}(y,w)\Big) &= f'(\hat{x}) \cdot \hat{x}/f(\hat{x}) \\
&= f'(\hat{x}) \cdot \hat{x}/y && \text{as } y = f\Big(\hat{x}(y,w)\Big) \\
&= \frac{w}{\hat{\lambda}} \cdot \hat{x}/y && \text{by the first order condition } w = \hat{\lambda} f'(\hat{x}) \\
&= \frac{c(y,w)/y}{D_y c(y,w)} && \text{as } c(y,w) = w \cdot \hat{x}(y,w), \text{ and by (11) } \hat{\lambda} = D_y c(y,w) \\
&= \mathrm{AC}(y)/\,\mathrm{MC}(y).
\end{aligned}$$

Holding w fixed, and writing the cost simply as a function of y,

$$\frac{d}{dy}\mathrm{AC}(y) = \frac{d}{dy}\frac{c(y)}{y} = \frac{c'(y)y - c(y)}{y^2} = \frac{1}{y}\left(c'(y) - \frac{c(y)}{y}\right) = \frac{1}{y}\left(\mathrm{MC}(y) - \mathrm{AC}(y)\right).$$

Thus

$$\mathrm{AC}'(y) > 0 \iff \mathrm{MC}(y) > \mathrm{AC}(y) \iff e(\hat{x}) < 1.$$

7.5 Average cost and constant returns to scale

If f exhibits constant returns to scale, then:

- the conditional input demand functions $\hat{x}(w, y)$ are homogenous of degree 1 in y.
- Marginal cost = average cost.
- For a price-taking profit maximizer, price = marginal cost = average cost, so profit is zero.
- If price is less than marginal cost, then the optimal output is zero. If price is equal to marginal cost, then every level of output maximizes profit, which is zero. If price is greater than marginal cost, then the profit function is unbounded, so no profit maximizer exists.

7.6 Recovering the Production Function from a Cost Function

We already know from the support function that the input requirement set for y is

$$\{x : w \cdot x \geqslant c(w, y) \text{ for all } w \in \boldsymbol{R}^{\mathrm{n}}_{+}\}.$$

But there is often another way to get a nicer expression for the production function using the envelope theorem.

7.6.1 Example Consider the cost function (with two inputs)

$$c(w, y) = y(w_1^\sigma + w_2^\sigma)^{1/\sigma}.$$

By the envelope theorem

$$\frac{\partial c}{\partial w_i} = y\frac{1}{\sigma}(w_1^\sigma + w_2^\sigma)^{\frac{1-\sigma}{\sigma}}\sigma w_i^{\sigma-1} = x_i^*,$$

where $x^*(w, y)$ is the cost minimizing input vector. We can eliminate w_1 and w_2 and solve for y as a function of x_1 and x_2. Here's the trick: exponentiate the above equality to the

$$\rho = \frac{\sigma}{\sigma - 1}$$

power to get

$$y^\rho(w_1^\sigma + w_2^\sigma)^{-1}w_i^\sigma = x_i^\rho$$

and sum over i to get

$$x_1^\rho + x_2^\rho = y^\rho(w_1^\sigma + w_2^\sigma)^{-1}(w_1^\sigma + w_2^\sigma) = y^\rho,$$

which gives the production function

$$y = (x_1^\rho + x_2^\rho)^{1/\rho}.$$

This called the constant elasticity of substitution production function, or the **Arrow–Chenery–Minhas–Solow production function**, see [1]. □

7.6.2 Example Given the cost function

$$c(w, y) = y\sum_{i=1}^n \alpha_i w_i$$

By the envelope theorem

$$\frac{\partial c}{\partial w_i} = \alpha_i y = x_i^*,$$

where $x^*(w, y)$ is the cost minimizing input vector. This implies that the cost minimizing point x^* is independent of w! Thus

$$y = \frac{x_i^*}{\alpha_i}, \quad i = 1, \dots, n.$$

Using the support function approach to finding the input requirement set, we see that it is $\{x : x \geqq x^*\}$, so that the production function is

$$y = \min_i \frac{x_i}{\alpha_i}.$$

This sort of production function is a **Leontief production function**. □

References

[1] K. J. Arrow, H. B. Chenery, B. S. Minhas, and R. M. Solow. 1961. Capital-labor substitution and economic efficiency. *Review of Economics and Statistics* 43(3):225–250. http://www.jstor.org/stable/1927286

[2] K. C. Border. 2001. More than you wanted to know about quadratic forms. http://www.its.caltech.edu/~kcborder/Notes/QuadraticForms.pdf

[3] H. R. Varian. 1992. *Microeconomic analysis*, 3d. ed. New York: W. W. Norton & Co.