5330 Advanced Microeconomic Theory

Lecture: The Producer

Professor Paul Segerstrom

Introduction

This lecture is mainly based on

- Geoffrey Jehle and Philip Reny (2011), *Advanced Microeconomic Theory*, chapter 3, and any good mathematics textbook.

Production

- Production is the process of transforming inputs into outputs. The fundamental reality firms must contend with in this process is *technological feasibility*.
- We consider firms producing only a single product from many inputs and describe the firm's technology in terms of a **production function**.
- We denote the amount of output by y and the amount of input i by x_i, so that with n inputs, the entire vector of inputs is denoted by $\mathbf{x} \equiv (x_1, \ldots, x_n)$, where $\mathbf{x} \geq \mathbf{0}$ and $y \geq 0$.
- A production function simply describes for each vector of inputs the amount of output that can be produced.

- The production function f is therefore a mapping from $\mathbb{R}^n_+$ into $\mathbb{R}_+$ and is written as $y = f(\mathbf{x})$.
- We shall maintain the following assumption on the production function f.

  ASSUMPTION: The production function $f : \mathbb{R}^n_+ \to \mathbb{R}_+$ is continuous, strictly increasing and strictly quasiconcave on $\mathbb{R}^n_+$ and $f(\mathbf{0}) = 0$.

- Here, quasiconcavity means that any convex combination of two input vectors can produce at least as much output as one of the original two.

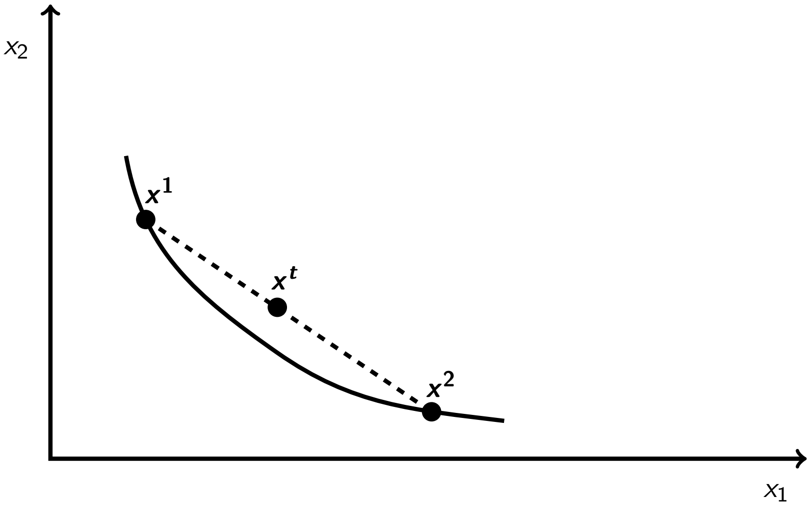

Figure: Illustration of a strictly quasiconcave production function.

- When the production function is differentiable, its partial derivative $\partial f(\mathbf{x})/\partial x_i$ is called the **marginal product** of input i.
- If f is strictly increasing and everywhere differentiable, then $\partial f(\mathbf{x})/\partial x_i > 0$ for "almost all" input vectors. We will often assume for simplicity that strict inequality always holds.
- For any fixed level of output y, the set of input vectors producing y units of output is called the y-level **isoquant**

$$Q(y) \equiv \{\mathbf{x} \geq \mathbf{0} \,|\, f(\mathbf{x}) = y\}.$$

- An analog to the marginal rate of substitution in consumer theory is the **marginal rate of technical substitution (MRTS)** in producer theory. This measures the rate at which one input can be substituted for another without changing the amount of output produced.
- Formally, the marginal rate of technical substitution between inputs i and j when the current input vector is $\mathbf{x}$ is defined as the ratio of marginal products:

$$MRTS_{ij}(\mathbf{x}) \equiv \frac{\partial f(\mathbf{x})/\partial x_i}{\partial f(\mathbf{x})/\partial x_j}.$$

- To justify this, suppose $n = 2$. Then the (absolute value of the) slope of an isoquant curve is the marginal rate of technical substitution.
- Let $x_2 = g(x_1)$ be the function describing an isoquant curve. It follows that for all x_1, $f(x_1, g(x_1)) =$ constant. Its derivative with respect to x_1 must equal zero:

$$\frac{\partial f(x_1, x_2)}{\partial x_1} + \frac{\partial f(x_1, x_2)}{\partial x_2} g'(x_1) = 0.$$

  Thus, the marginal rate of technical substitution is

$$MRTS_{12}(\mathbf{x}) \equiv -g'(x_1) = \frac{\partial f(\mathbf{x})/\partial x_1}{\partial f(\mathbf{x})/\partial x_2} > 0.$$

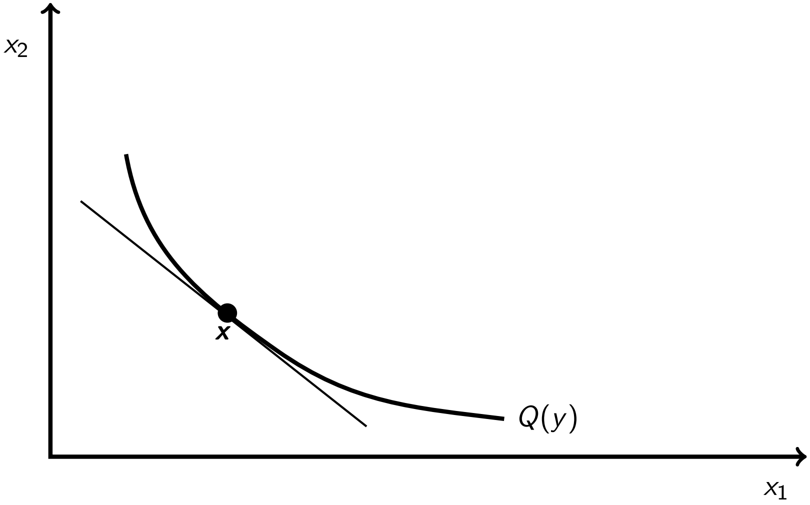

Figure: The marginal rate of technical substitution measures the |slope| of an isoquant. The elasticity of substitution measures the curvature of an isoquant.

- Between two inputs x_i and x_j, holding all other inputs and the level of output constant, the *elasticity of substitution* σ is defined as the percentage change in the input proportions x_j/x_i associated with a 1 percent change in the MRTS between them.

- **DEFINITION (The Elasticity of Substitution)**: For a production function $f(\mathbf{x})$, the elasticity of substitution between inputs i and j at the point $\mathbf{x}$ is defined as

$$\sigma_{ij} \equiv \frac{d \ln(x_j/x_i)}{d \ln(f_i(\mathbf{x})/f_j(\mathbf{x}))} = \frac{d(x_j/x_i)}{x_j/x_i} \frac{f_i(\mathbf{x})/f_j(\mathbf{x})}{d(f_i(\mathbf{x})/f_j(\mathbf{x}))}$$

  where f_i and f_j are the marginal products of inputs i and j.

- Thinking about the $n = 2$ case, when the production function is strictly quasiconcave and $MRTS_{12} > 0$ increases by 1 percent, x_2 increases, x_1 decreases and x_2/x_1 increases. So $\sigma_{12} > 0$.

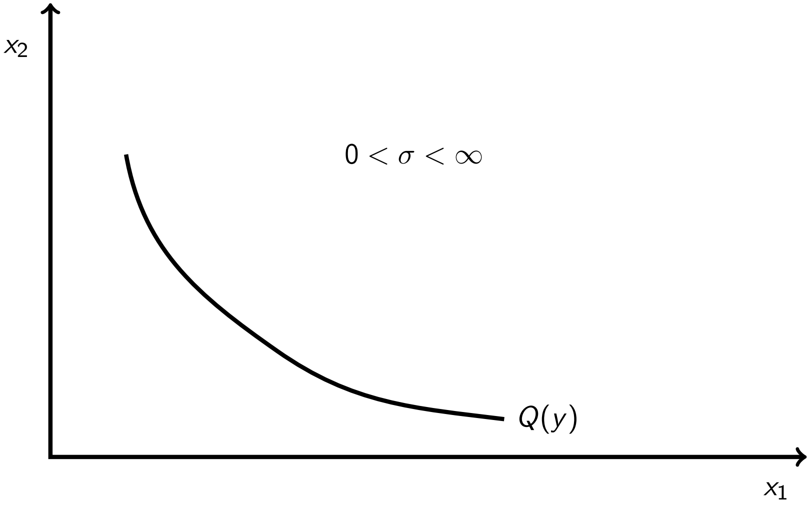

Figure: σ is finite but larger than zero, indicating less that perfect substitutability.

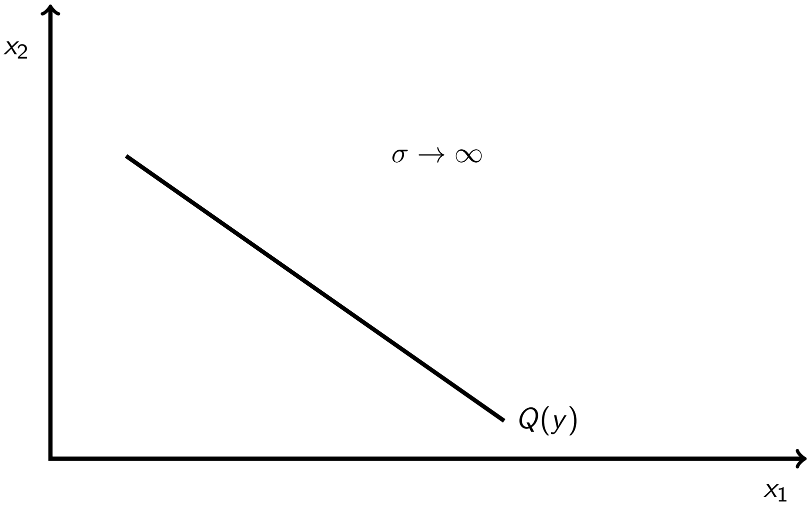

Figure: σ is infinite and there is perfect substitutability between inputs.

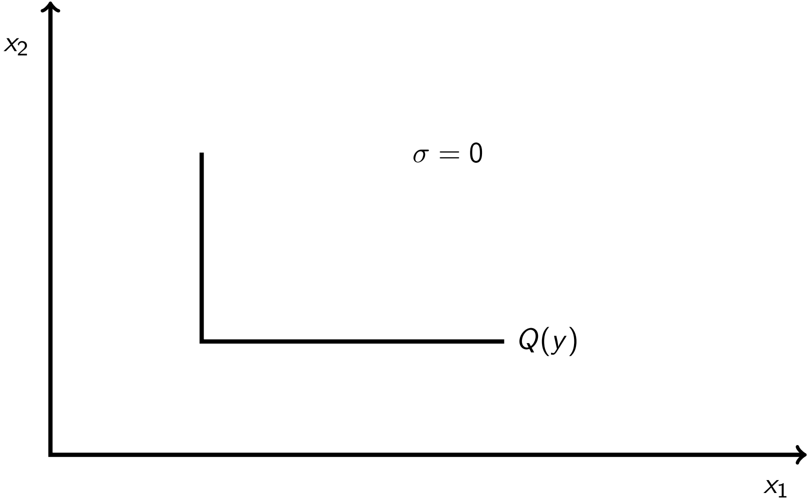

Figure: σ is zero and there is no substitutability between inputs.

An Example: CES Production Function

- We are familiar with the CES utility function from demand theory. It is time to see where this name comes from by considering the CES production function,

$$y = f(x_1, x_2) \equiv (x_1^\rho + x_2^\rho)^{1/\rho}$$

where $-\infty < \rho < 1$ and $\rho \neq 0$.

- To calculate the elasticity of substitution, first we calculate the marginal rate of technical substitution:

$$MRTS_{12}(\mathbf{x}) \equiv \frac{\partial f(\mathbf{x})/\partial x_1}{\partial f(\mathbf{x})/\partial x_2} = \frac{(1/\rho)(x_1^\rho + x_2^\rho)^{(1/\rho)-1}\rho x_1^{\rho-1}}{(1/\rho)(x_1^\rho + x_2^\rho)^{(1/\rho)-1}\rho x_2^{\rho-1}}$$

$$= \left(\frac{x_1}{x_2}\right)^{\rho-1}$$

- Taking logs of both sides yields

$$\ln MRTS_{12}(\mathbf{x}) = (\rho - 1)\ln\left(\frac{x_1}{x_2}\right) = (1-\rho)\ln\left(\frac{x_2}{x_1}\right)$$

  and then totally differentiating

$$d\ln MRTS_{12}(\mathbf{x}) = (1-\rho)\, d\ln\left(\frac{x_2}{x_1}\right).$$

- Thus the elasticity of substitution is

$$\sigma_{12} \equiv \frac{d\ln(x_2/x_1)}{d\ln(f_1(\mathbf{x})/f_2(\mathbf{x}))} = \frac{1}{1-\rho},$$

  which is a constant: hence the initials CES, which stands for **constant elasticity of substitution**.

- $$\sigma_{12} \equiv \frac{d \ln(x_2/x_1)}{d \ln(f_1(\mathbf{x})/f_2(\mathbf{x}))} = \frac{1}{1-\rho}.$$

  With the CES form, the degree of substitutability between inputs is always the same, regardless of the level of output or input proportions.
- As ρ increases and converges to one, σ_{12} increases and converges to $+\infty$ (perfect substitutability between inputs and linear isoquants).
- As ρ decreases and converges to $-\infty$, σ_{12} decreases and converges to zero (no substitutability between inputs and L-shaped isoquants).

- It is easy to verify that

$$y = \left(\sum_{i=1}^{n} \alpha_i x_i^{\rho} \right)^{1/\rho} \quad \text{where} \quad \sum_{i=1}^{n} \alpha_i = 1$$

  is a CES form with $\sigma_{ij} = \frac{1}{1-\rho}$ for all $i \neq j$.
- It can be shown that as $\rho \to 0$, $\sigma_{ij} \to 1$ and this CES form reduces to the Cobb-Douglas form

$$y = \prod_{i=1}^{n} x_i^{\alpha_i}.$$

  So, for example, the Cobb-Douglas production function $y = A x_1^{\alpha} x_2^{1-\alpha}$ has constant elasticity of substitution equal to one.

- $$y = \left(\sum_{i=1}^{n} \alpha_i x_i^{\rho} \right)^{1/\rho} \quad \text{where } \sum_{i=1}^{n} \alpha_i = 1$$

  is a CES form with $\sigma_{ij} = \frac{1}{1-\rho}$ for all $i \neq j$.
- As $\rho \to -\infty$, $\sigma_{ij} \to 0$, giving the Leontief form as a limiting case, where

  $$y = \min\{x_1, \ldots, x_n\}$$

  with L-shaped isoquants.
- The following continuum of goods utility function is also CES with $\sigma = \frac{1}{1-\rho}$ and is used Melitz (2003)

  $$u = \left(\int_{\omega \in \Omega} x(\omega)^{\rho} \, d\omega \right)^{1/\rho}.$$

A Second Example: Cobb-Douglas Production Function

- The Cobb-Douglas production function in its simplest form is

$$y = f(x_1, x_2) \equiv x_1^\alpha x_2^{1-\alpha} \quad \text{where} \quad 0 < \alpha < 1.$$

- First we calculate the marginal rate of technical substitution:

$$MRTS_{12}(\mathbf{x}) \equiv \frac{\partial f(\mathbf{x})/\partial x_1}{\partial f(\mathbf{x})/\partial x_2} = \frac{\alpha x_1^{\alpha-1} x_2^{1-\alpha}}{x_1^\alpha (1-\alpha) x_2^{-\alpha}} = \frac{\alpha}{1-\alpha} \frac{x_2}{x_1}$$

- Note that the slope of an isoquant only depends on the input ratio x_2/x_1, and that the CES form has the same property.

- Taking logs of both sides and then totally differentiating yields

$$\ln MRTS_{12}(\mathbf{x}) = \ln\left(\frac{\alpha}{1-\alpha}\right) + \ln\left(\frac{x_2}{x_1}\right)$$

$$d\ln MRTS_{12}(\mathbf{x}) = d\ln\left(\frac{x_2}{x_1}\right)$$

- Thus the elasticity of substitution is

$$\sigma_{12} \equiv \frac{d\ln(x_2/x_1)}{d\ln(f_1(\mathbf{x})/f_2(\mathbf{x}))} = 1.$$

- The Cobb-Douglas production has the same elasticity of substitution (equal to one) for all output levels y and all input ratios x_2/x_1. It is a constant elasticity of substitution production function, as was earlier claimed.

- Looking at the CES form

$$y = \left(\sum_{i=1}^{n} \alpha_i x_i^{\rho} \right)^{1/\rho} \quad \text{where } \sum_{i=1}^{n} \alpha_i = 1 \text{ and } \sigma_{ij} = \frac{1}{1-\rho}$$

  and noting that as $\rho \to 0$, $\sigma_{ij} \to 1$, it is not surprising that the CES form converges to the Cobb-Douglas form

$$y = \prod_{i=1}^{n} x_i^{\alpha_i}.$$

- There is convergence in the shapes of the isoquants.

- All CES production functions (including the limiting cases of Cobb-Douglas and Leontief) are members of the class of linear homogeneous production functions, and these are important in theoretical and applied work.
- **DEFINITION**: A real-valued function $f(\mathbf{x})$ is called homogeneous of degree k if
$$f(t\mathbf{x}) = t^k f(\mathbf{x}) \;\; \text{for all} \;\; t > 0.$$
- **THEOREM (Shephard)**: Let $f(\mathbf{x})$ be a production function satisfying the previously-stated assumptions and suppose it is homogeneous of degree $k \in (0, 1]$. Then $f(\mathbf{x})$ is a concave function.

- **Proof**: Suppose first that $k = 1$. Take any $\mathbf{x}^1 \gg \mathbf{0}$ and $\mathbf{x}^2 \gg \mathbf{0}$ and let $y^1 = f(\mathbf{x}^1)$ and $y^2 = f(\mathbf{x}^2)$. Then $y^1, y^2 > 0$ because $f(\mathbf{0}) = 0$ and f is strictly increasing.
- Therefore, because f is homogeneous of degree one,

$$f\left(\frac{\mathbf{x}^1}{y^1}\right) = \frac{y^1}{y^1} = 1 = f\left(\frac{\mathbf{x}^2}{y^2}\right) = \frac{y^2}{y^2}.$$

- Because f is quasiconcave,

$$f\left(t\frac{\mathbf{x}^1}{y^1} + (1-t)\frac{\mathbf{x}^2}{y^2}\right) \geq 1 \text{ for all } t \in [0,1].$$

- Now let $t^* \equiv y^1/(y^1 + y^2)$ and $1 - t^* = y^2/(y^1 + y^2)$. Then

$$f\left(t^*\frac{\mathbf{x}^1}{y^1} + (1-t^*)\frac{\mathbf{x}^2}{y^2}\right) = f\left(\frac{\mathbf{x}^1}{y^1 + y^2} + \frac{\mathbf{x}^2}{y^1 + y^2}\right) \geq 1$$

- $$f\left(\frac{\mathbf{x}^1}{y^1+y^2}+\frac{\mathbf{x}^2}{y^1+y^2}\right)\geq 1$$

  Again invoking the linear homogeneity of f gives

  $$f\left(\frac{\mathbf{x}^1}{y^1+y^2}+\frac{\mathbf{x}^2}{y^1+y^2}\right)=\frac{1}{y^1+y^2}f(\mathbf{x}^1+\mathbf{x}^2)\geq 1$$

  so

  $$f(\mathbf{x}^1+\mathbf{x}^2)\geq y^1+y^2=f(\mathbf{x}^1)+f(\mathbf{x}^2).$$

  Since this last equation holds for all $\mathbf{x}^1,\mathbf{x}^2\gg\mathbf{0}$, the continuity of f implies that this equation hold for all $\mathbf{x}^1,\mathbf{x}^2\geq\mathbf{0}$.

- To complete the proof for the $k = 1$ case, consider any two vectors $\mathsf{x}^1 \geq \mathbf{0}$ and $\mathsf{x}^2 \geq \mathbf{0}$ and any $t \in [0, 1]$.
- Recall that linear homogeneity ensures that

$$f(t\mathsf{x}^1) = t\, f(\mathsf{x}^1)$$

$$f((1-t)\mathsf{x}^2) = (1-t)\, f(\mathsf{x}^2).$$

- We conclude that

$$f(t\mathsf{x}^1 + (1-t)\mathsf{x}^2) \geq f(t\mathsf{x}^1) + f((1-t)\mathsf{x}^2) = t\, f(\mathsf{x}^1) + (1-t)\, f(\mathsf{x}^2).$$

- Suppose now that f is homogeneous of degree $k \in (0,1)$. Then $f^{1/k}$ is homogeneous of degree one

$$f(t\mathbf{x}) = t^k f(\mathbf{x})$$

$$f(t\mathbf{x})^{1/k} = [t^k f(\mathbf{x})]^{1/k} = t\, f(\mathbf{x})^{1/k}$$

and satisfies the standard assumptions (continuous, strictly increasing, strictly quasiconcave). Hence, by what we have just proven, $f^{1/k}$ is concave. But then $f = [f^{1/k}]^k$ is concave since $k < 1$.

$\square$

- **DEFINITION (Returns to Scale)**: A production function $f(\mathbf{x})$ has the property of:

  1. Constant returns to scale if $f(t\mathbf{x}) = t\,f(\mathbf{x})$ for all $t > 0$ and all $\mathbf{x}$.
  2. Increasing returns to scale if $f(t\mathbf{x}) > t\,f(\mathbf{x})$ for all $t > 1$ and all $\mathbf{x}$.
  3. Decreasing returns to scale if $f(t\mathbf{x}) < t\,f(\mathbf{x})$ for all $t > 1$ and all $\mathbf{x}$.

Cost

- We will assume throughout that firms are perfectly competitive on their input markets and that therefore they face input prices.
- Let $\mathbf{w} = (w_1, \ldots, w_n) \geq \mathbf{0}$ be a vector of prevailing market prices at which the firm can buy inputs $\mathbf{x} = (x_1, \ldots, x_n) \geq \mathbf{0}$
- Because the firm is a profit maximizer, it will choose to produce some level of output while using that input vector requiring the smallest money outlay.

- **DEFINITION (The Cost Function)**: The cost function, defined for all input prices $\mathbf{w} \gg \mathbf{0}$ and all output levels $y \in f(\mathbb{R}^n_+)$, is the minimum-value function

$$c(\mathbf{w}, y) \equiv \min_{\mathbf{x} \in \mathbb{R}^n_+} \mathbf{w} \cdot \mathbf{x} \quad \text{s.t.} \quad f(\mathbf{x}) \geq y.$$

  If $x(\mathbf{w}, y)$ solves the cost-minimization problem, then

$$c(\mathbf{w}, y) = \mathbf{w} \cdot \mathbf{x}(\mathbf{w}, y).$$

- To apply the Kuhn-Tucker Theorem, we rewrite the problem as

$$-c(\mathbf{w}, y) \equiv \max_{\mathbf{x} \in \mathbb{R}^n_+} -\mathbf{w} \cdot \mathbf{x} \quad \text{s.t.} \quad y - f(\mathbf{x}) \leq 0.$$

- Because f is strictly increasing, the constraint will always be binding at a solution.
- Consequently, the cost minimization problem is equivalent to

$$-c(\mathbf{w}, y) \equiv \max_{\mathbf{x} \in \mathbb{R}^n_+} -\mathbf{w} \cdot \mathbf{x} \quad \text{s.t.} \quad y - f(\mathbf{x}) = 0.$$

- Let $\mathbf{x}^*$ denote a solution to this problem. To keep things simple, assume that $\mathbf{x}^* \gg \mathbf{0}$, and that f is differentiable at $\mathbf{x}^*$ with $\nabla f(\mathbf{x}^*) \gg \mathbf{0}$.
- Then the Lagrangian function is

$$\mathcal{L}(\mathbf{x}, \lambda) \equiv -\mathbf{w} \cdot \mathbf{x} - \lambda \left[y - f(\mathbf{x}) \right]$$

- $$\mathcal{L}(\mathbf{x}, \lambda) \equiv -\mathbf{w} \cdot \mathbf{x} - \lambda [y - f(\mathbf{x})]$$

  and the first-order conditions are

  $$\frac{\partial \mathcal{L}}{\partial x_i} = -w_i + \lambda \frac{\partial f}{\partial x_i} = 0, \quad i = 1, \ldots, n$$

  $$\frac{\partial \mathcal{L}}{\partial \lambda} = -y + f(\mathbf{x}) = 0$$

- Thus, by Lagrange's Theorem, there exists a $\lambda^* \in \mathbb{R}$ such that

  $$w_i = \lambda^* \frac{\partial f(\mathbf{x}^*)}{\partial x_i}, \quad i = 1, \ldots, n.$$

- $$w_i = \lambda^* \frac{\partial f(\mathbf{x}^*)}{\partial x_i}, \quad i = 1, \ldots, n.$$

  Because $w_i > 0$ for all i, we divide these equation by each other to obtain
  $$\frac{\partial f(\mathbf{x}^*)/\partial x_i}{\partial f(\mathbf{x}^*)/\partial x_j} = \frac{w_i}{w_j}.$$
  Thus, cost minimization implies that the marginal rate of substitution between any two inputs is equal to the ratio of their prices.

- From the first-order conditions, it is clear the solution depends on the parameters $\mathbf{w}$ and y. Moreover, because $\mathbf{w} \gg \mathbf{0}$ and f is strictly quasiconcave, the solution exists and is unique.
- So we can write the solution as $\mathbf{x}^* \equiv \mathbf{x}(\mathbf{w}, y)$ and refer to it as the firm's **conditional input demand**.

An Example: CES Production

- Suppose the firm's technology is the two-input CES form $f(x_1, x_2) \equiv (x_1^\rho + x_2^\rho)^{1/\rho}$ where $-\infty < \rho < 1$ and $\rho \neq 0$. Its cost minimization problem is then

$$-c(\mathbf{w}, y) \equiv \max_{x_1 \geq 0,\, x_2 \geq 0} -(w_1 x_1 + w_2 x_2) \quad \text{s.t.} \quad y - (x_1^\rho + x_2^\rho)^{1/\rho} \leq 0.$$

- Assuming $y > 0$ and an interior solution, the Lagrangian function is

$$\mathcal{L}(x_1, x_2, \lambda) \equiv -(w_1 x_1 + w_2 x_2) - \lambda \left[y - (x_1^\rho + x_2^\rho)^{1/\rho} \right]$$

and the first-order conditions are

$$\frac{\partial \mathcal{L}}{\partial x_i} = -w_i + \lambda (1/\rho)(x_1^\rho + x_2^\rho)^{(1/\rho)-1} \rho x_i^{\rho-1} = 0, \quad i = 1, 2$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = -y + (x_1^\rho + x_2^\rho)^{1/\rho} = 0$$

- The first-order Lagrangian conditions reduce to 2 equations in 2 unknowns

$$\frac{w_1}{w_2} = \left(\frac{x_1}{x_2}\right)^{\rho-1}$$

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

- Solving for x_1 yields

$$x_1 = x_2 \left(\frac{w_1}{w_2}\right)^{1/(\rho-1)}$$

$$y = \left(x_2^\rho \left(\frac{w_1}{w_2}\right)^{\rho/(\rho-1)} + x_2^\rho\right)^{1/\rho} \left[\left(\frac{w_2}{w_2}\right)^{\rho/(\rho-1)}\right]^{1/\rho}$$

$$y = \frac{x_2}{w_2^{1/(\rho-1)}} \left(w_1^{\rho/(\rho-1)} + w_2^{\rho/(\rho-1)}\right)^{1/\rho}$$

- So, rearranging gives the conditional input demands

$$x_2 = yw_2^{1/(\rho-1)}\left(w_1^{\rho/(\rho-1)} + w_2^{\rho/(\rho-1)}\right)^{-1/\rho}$$

$$x_1 = x_2\left(\frac{w_1}{w_2}\right)^{1/(\rho-1)} = yw_1^{1/(\rho-1)}\left(w_1^{\rho/(\rho-1)} + w_2^{\rho/(\rho-1)}\right)^{-1/\rho}$$

and the cost function

$$c(\mathbf{w}, y) = w_1 x_1(\mathbf{w}, y) + w_2 x_2(\mathbf{w}, y)$$

$$= \left(w_1 y w_1^{1/(\rho-1)} + w_2 y w_2^{1/(\rho-1)}\right)\left(w_1^{\rho/(\rho-1)} + w_2^{\rho/(\rho-1)}\right)^{-1/\rho}$$

$$= y\left(w_1^{\rho/(\rho-1)} + w_2^{\rho/(\rho-1)}\right)\left(w_1^{\rho/(\rho-1)} + w_2^{\rho/(\rho-1)}\right)^{-1/\rho}$$

$$= y\left(w_1^{\rho/(\rho-1)} + w_2^{\rho/(\rho-1)}\right)^{(\rho-1)/\rho}$$

- To compare the cost function with the expenditure function, consider their definitions.
- Expenditure function:

$$e(\mathbf{p}, u) \equiv \min_{\mathbf{x} \in \mathbb{R}^n_+} \mathbf{p} \cdot \mathbf{x} \quad \text{s.t.} \quad u(\mathbf{x}) \geq u$$

- Cost function:

$$c(\mathbf{w}, y) \equiv \min_{\mathbf{x} \in \mathbb{R}^n_+} \mathbf{w} \cdot \mathbf{x} \quad \text{s.t.} \quad f(\mathbf{x}) \geq y.$$

- Mathematically, the two functions are identical. Consequently, for every theorem we proved about expenditure functions, there is an equivalent theorem for cost functions.

- **THEOREM (Properties of the Cost Function)**: If $f(\cdot)$ is continuous and strictly increasing, then $c(\mathbf{w}, y)$ is

  1. Zero when $y = 0$.
  2. Continuous on its domain.
  3. For all $\mathbf{w} \gg \mathbf{0}$, strictly increasing and unbounded above in y.
  4. Increasing in $\mathbf{w}$.
  5. Homogeneous of degree 1 in $\mathbf{w}$.
  6. Concave in $\mathbf{w}$.

  If, in addition, $f(\cdot)$ is strictly quasiconcave, we have

  7. Shephard's lemma: $c(\mathbf{w}, y)$ is differentiable in $\mathbf{w}$ and

$$\frac{\partial c(\mathbf{w}^0, y^0)}{\partial w_i} = x_i(\mathbf{w}^0, y^0), \quad i = 1, \ldots, n.$$

- **THEOREM (Properties of Conditional Input Demands)**: Suppose the production function satisfies the previously-stated assumptions (continuous, strictly increasing, strictly quasiconcave) and that the associated cost function is twice continuously differentiable. Then

  1. $\mathbf{x}(\mathbf{w}, y)$ is homogeneous of degree zero in $\mathbf{w}$.
  2. The substitution matrix

$$\begin{pmatrix} \frac{\partial x_1(\mathbf{w},y)}{\partial w_1} & \cdots & \frac{\partial x_1(\mathbf{w},y)}{\partial w_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial x_n(\mathbf{w},y)}{\partial w_1} & \cdots & \frac{\partial x_n(\mathbf{w},y)}{\partial w_n} \end{pmatrix}$$

  is symmetric and negative semidefinite. In particular, the negative semi-definiteness property implies that $\partial x_i(\mathbf{w}, y)/\partial w_i \leq 0$ for all i.

- **Proof**: The second part of the theorem has already been proved when we studied the properties of Hicksian demand systems, but the first part is new, so here is a proof.
- Since the cost function is homogeneous of degree one in $\mathbf{w}$, the partial derivative of the cost function with respect to any w_i is homogeneous of degree zero in $\mathbf{w}$.
- Here, we are using the theorem

  THEOREM (Partial Derivatives of Homogeneous Functions):
  If $f(\mathbf{x})$ is homogeneous of degree k, its partial derivatives are homogeneous of degree $k-1$.

- Thus, by Shephard's lemma,

$$\frac{\partial c(\mathbf{w}^0, y^0)}{\partial w_i} = x_i(\mathbf{w}^0, y^0), \quad i = 1, \ldots, n,$$

  $x_i(\mathbf{w}, y)$ is homogeneous of degree zero in $\mathbf{w}$ for all i, so $\mathbf{x}(\mathbf{w}, y)$ is homogeneous of degree zero in $\mathbf{w}$.

$\square$

Duality in Production

- Given a production function $y = f(\mathbf{x})$ describing a firm's technology, we can calculate the firm's corresponding cost function $c(\mathbf{w}, y)$.
- The question naturally arises, can we do this in reverse? Given a firm's cost function $c(\mathbf{w}, y)$, can we calculate the firm's production function?
- More generally, if we simply construct some function of $\mathbf{w}$ and y with all the properties of a cost function, can we know there exists some production function that would have generated the function we constructed?

- Given the assumptions that we have made about production functions, in particular, that they are strictly increasing and strictly quasiconcave, the answer is yes.
- It follows from the mathematical "duality" between the constrained optimization problems used to describe firm behavior. These duality relationships are important in both theoretical and empirical analysis.
- They allow us to choose on the basis of convenience or tractability whether to begin with a specification of the production function or with a specification of the cost function. If we choose the latter, duality allows us to be confident we are carrying along the same primitive assumptions about production technology with which we otherwise would have had to begin.

- Applied researchers need no longer begin their study of the firm with detailed knowledge of the technology and with access to relatively obscure engineering data. Instead, they can estimate the firm's cost function by using observable market input prices, costs of production and levels of output. They can then "recover" the underlying production function from the estimated cost function.
- Here is the main theorem about duality in production.
- (The corresponding theorem about duality in consumption shows how to construct a utility function from an expenditure function).

- **THEOREM (Recovering a Production Function from a Cost Function)**: Let $c : \mathbb{R}^n_{++} \times \mathbb{R}_+ \to \mathbb{R}_+$ satisfy the 7 previously-stated properties of a cost function. Then the function $f : \mathbb{R}^n_+ \to \mathbb{R}_+$ defined by

$$f(\mathbf{x}) \equiv \max\{y \geq 0 \,|\, \mathbf{w} \cdot \mathbf{x} \geq c(\mathbf{w}, y), \text{ for all } \mathbf{w} \gg \mathbf{0}\}$$

  is increasing, unbounded above, and quasiconcave. Moreover, the cost function generated by f is c.

- We illustrate this theorem about duality in production using an example. From before, we have shown that given the CES production function
$$y = f(\mathbf{x}) \equiv (x_1^\rho + x_2^\rho)^{1/\rho},$$
the corresponding CES cost function is
$$c(\mathbf{w}, y) = y \left(w_1^{\rho/(\rho-1)} + w_2^{\rho/(\rho-1)} \right)^{(\rho-1)/\rho}.$$
- Now, we start with the cost function c and "work backwards" to discover the underlying production function f.

- $$f(\mathbf{x}) \equiv \max\left\{y \geq 0 \,|\, \mathbf{w}\cdot\mathbf{x} \geq c(\mathbf{w}, y), \text{ for all } \mathbf{w} \gg \mathbf{0}\right\}$$

$$f(\mathbf{x}) \equiv \max\{y \geq 0 \,|\, w_1 x_1 + w_2 x_2$$
$$\geq y\left(w_1^{\rho/(\rho-1)} + w_2^{\rho/(\rho-1)}\right)^{(\rho-1)/\rho} \text{ for all } \mathbf{w} \gg \mathbf{0}\}$$

$$\mathcal{L} \equiv y - \lambda\left[y\left(w_1^{\rho/(\rho-1)} + w_2^{\rho/(\rho-1)}\right)^{(\rho-1)/\rho} - (w_1 x_1 + w_2 x_2)\right]$$

$$\frac{\partial \mathcal{L}}{\partial y} = 1 - \lambda\left(w_1^{\rho/(\rho-1)} + w_2^{\rho/(\rho-1)}\right)^{(\rho-1)/\rho} = 0$$

$$\frac{\partial \mathcal{L}}{\partial w_i} = \lambda\left[x_i - y\left(w_1^{\rho/(\rho-1)} + w_2^{\rho/(\rho-1)}\right)^{-1/\rho} w_i^{1/(\rho-1)}\right] = 0$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = w_1 x_1 + w_2 x_2 - y\left(w_1^{\rho/(\rho-1)} + w_2^{\rho/(\rho-1)}\right)^{(\rho-1)/\rho} = 0$$

-
$$x_1 = y\frac{\rho-1}{\rho}\left(w_1^{\rho/(\rho-1)} + w_2^{\rho/(\rho-1)}\right)^{-1/\rho}\frac{\rho}{\rho-1}w_1^{1/(\rho-1)}$$
$$x_2 = y\frac{\rho-1}{\rho}\left(w_1^{\rho/(\rho-1)} + w_2^{\rho/(\rho-1)}\right)^{-1/\rho}\frac{\rho}{\rho-1}w_2^{1/(\rho-1)}$$
implies that
$$\frac{x_1}{x_2} = \left(\frac{w_1}{w_2}\right)^{1/(\rho-1)}$$
- Since only relative prices matter, we now set $w_2 = 1$, which implies that
$$\left(\frac{x_1}{x_2}\right)^{\rho-1} = w_1$$

- $$w_1 x_1 + w_2 x_2 = y\left(w_1^{\rho/(\rho-1)} + w_2^{\rho/(\rho-1)}\right)^{(\rho-1)/\rho}$$

$$\left(\frac{x_1}{x_2}\right)^{\rho-1} x_1 + x_2 = y\left(\left(\frac{x_1}{x_2}\right)^{(\rho-1)\rho/(\rho-1)} + 1\right)^{(\rho-1)/\rho}$$

$$\left(\frac{x_1}{x_2}\right)^{\rho} x_2 + x_2 = y\left(\left(\frac{x_1}{x_2}\right)^{\rho} + 1\right)^{(\rho-1)/\rho}$$

$$x_2\left(\left(\frac{x_1}{x_2}\right)^{\rho} + 1\right) = y\left(\left(\frac{x_1}{x_2}\right)^{\rho} + 1\right)^{(\rho-1)/\rho}$$

$$(x_2^{\rho})^{1/\rho}\left(\left(\frac{x_1}{x_2}\right)^{\rho} + 1\right)^{1/\rho} = (x_1^{\rho} + x_2^{\rho})^{1/\rho} = y$$

- We have recovered the CES production function!

The Competitive Firm

- Assume that the firm is both a perfect competitor on input markets and a perfect competitor on its output market.
- It is a perfect competitor on output markets if it believes the amount it produces and sells will have no effect on prevailing market prices.
- The competitive firm sees the market price for its product, assumes it will remain the same regardless of how much or how little it sells, and makes its plans accordingly.
- Such a firm is thus a **price taker** on both output and input markets.

Profit Maximization

- Profit is the difference between revenue from selling output and the cost of acquiring the factors necessary to produce it.
- The competitive firm can sell each unit of output at the market price p. Its revenues are therefore a simple function of output $R(y) = py$.
- Suppose the firm is considering output y^0. If $\mathbf{x}^0$ is a feasible vector of inputs to produce y^0 and $\mathbf{w}$ is the vector of factor prices, the cost of using $\mathbf{x}^0$ to produce y^0 is simply $\mathbf{w} \cdot \mathbf{x}^0$.
- This plan would therefore yield the firm profits of $py^0 - \mathbf{w} \cdot \mathbf{x}^0$.

- We suppose the overriding objective is to maximize profits. The firm therefore will choose the level of output and that combination of factors that solve the following problem:

$$\max_{(\mathbf{x},y)\geq\mathbf{0}} \; py - \mathbf{w}\cdot\mathbf{x} \quad \text{s.t.} \quad f(\mathbf{x}) \geq y,$$

  where $f(\mathbf{x})$ is a production function satisfying the previously-stated assumptions.
- Because the production function is strictly increasing, the constraint will hold with equality, $y = f(\mathbf{x})$, so we can rewrite the maximization problem as

$$\max_{\mathbf{x}\in\mathbb{R}^n_+} \; p\,f(\mathbf{x}) - \mathbf{w}\cdot\mathbf{x}.$$

- $$\max_{\mathbf{x}\in\mathbb{R}^n_+} \; p\,f(\mathbf{x}) - \mathbf{w}\cdot\mathbf{x}.$$

  Assume that this profit-maximization problem has an interior solution at the input vector $\mathbf{x}^* \gg \mathbf{0}$ and that $y^* = f(\mathbf{x}^*)$. Then the first-order conditions are

  $$p\frac{\partial f(\mathbf{x}^*)}{\partial x_i} - w_i = 0 \quad \text{for all } i = 1, \ldots, n,$$

  and rearranging yields

  $$p\frac{\partial f(\mathbf{x}^*)}{\partial x_i} = w_i \quad \text{for all } i = 1, \ldots, n.$$

- $$p\frac{\partial f(\mathbf{x}^*)}{\partial x_i} = w_i \quad \text{for all } i = 1, \ldots, n.$$

  The product of the output price with the marginal product of input i is the **marginal revenue product** of input i. It gives the rate at which revenue increases per additional unit of input i. At the optimum, this must equal the cost per unit of input i, namely, w_i.
- Assuming further that all the w_i are positive

  $$\frac{\partial f(\mathbf{x}^*)/\partial x_i}{\partial f(\mathbf{x}^*)/\partial x_j} = \frac{w_i}{w_j} \quad \text{for all } i, j.$$

  This is the same as the condition for cost-minimizing input choice, so profit maximization requires cost minimization in production.

- $$\max_{y \geq 0} \; p\,y - c(\mathbf{w}, y)$$

  If $y^* > 0$ is the optimal output, it therefore satisfies the first-order condition
  $$p - \frac{\partial c(\mathbf{w}, y^*)}{\partial y} = 0,$$
  so output is chosen so that price equals marginal cost. The second order conditions require that
  $$-\frac{\partial^2 c(\mathbf{w}, y^*)}{\partial y^2} \leq 0 \quad \text{or} \quad \frac{\partial^2 c(\mathbf{w}, y^*)}{\partial y^2} \geq 0 \quad \text{or} \quad \frac{d\,mc(y)}{dy} \geq 0$$
  so marginal cost must be increasing in y at the optimum.

- When the profit-maximization problem has a unique solution for each price vector $(p, \mathbf{w})$, the optimal choice of output $y^* = y(p, \mathbf{w})$ is called the firm's **output supply function** and the optimal choice of inputs $\mathbf{x}^* = \mathbf{x}(p, \mathbf{w})$ gives the vector of firm **input demand functions**.
- **DEFINITION**: The firm's **profit function** depends only on input and output prices and is defined as the maximum-value function

$$\pi(p, \mathbf{w}) \equiv \max_{(\mathbf{x}, y) \geq \mathbf{0}} py - \mathbf{w} \cdot \mathbf{x} \quad \text{s.t.} \quad f(\mathbf{x}) \geq y.$$

Example: The profit function for Cobb-Douglas technology

- Consider the problem of maximizing profits for the production function of the form $f(x) = x^a$ where $a > 0$. Then $\pi \equiv px^a - wx$. The first-order condition is

$$pax^{a-1} - w = 0$$

  and the second-order condition reduces to

$$pa(a-1)x^{a-2} \leq 0.$$

- The second-order condition can only be satisfied when $a \leq 1$, which means that the production function must have constant or decreasing returns to scale for competitive profit maximization to be meaningful.

- If $a = 1$, the first-order condition reduces to $p = w$. Hence, when $p = w$ and $\pi \equiv px^a - wx = px - px = 0$, any value of x is a profit-maximizing choice.
- When $a < 1$, $pax^{a-1} - w = 0$ implies that the input demand and output supply functions are

$$x(p, w) = \left(\frac{w}{ap}\right)^{1/(a-1)} \quad y(p, w) = f(x(p, w)) = \left(\frac{w}{ap}\right)^{a/(a-1)}$$

and the profit function is

$$\begin{aligned}
\pi(p, w) &= p\left(\frac{w}{ap}\right)^{a/(a-1)} - w\left(\frac{w}{ap}\right)^{1/(a-1)} \\
&= \left(\frac{w}{ap}\right)^{1/(a-1)}\left[p\left(\frac{w}{ap}\right)^{(a-1)/(a-1)} - w\right] \\
&= w\left(\frac{1-a}{a}\right)\left(\frac{w}{ap}\right)^{1/(a-1)}
\end{aligned}$$

- **THEOREM (Properties of the Profit Function)**: If f satisfies the previously-stated assumptions (continuous, strictly increasing, strictly quasiconcave), then for $p \geq 0$ and $\mathbf{w} \geq \mathbf{0}$, the profit function $\pi(p, \mathbf{w})$, where well-defined, is continuous and

  1. Increasing in p.

  2. Decreasing in $\mathbf{w}$.

  3. Homogeneous of degree one in $(p, \mathbf{w})$.

  4. Convex in $(p, \mathbf{w})$.

  5. Differentiable in $(p, \mathbf{w}) \gg \mathbf{0}$. Moreover, under the additional assumption that f is strictly concave (Hotelling's lemma)

$$\frac{\partial \pi(p, \mathbf{w})}{\partial p} = y(p, \mathbf{w}) \text{ and } -\frac{\partial \pi(p, \mathbf{w})}{\partial w_i} = x_i(p, \mathbf{w}), \quad i = 1, \ldots, n.$$

- **Proof**: Proofs of properties 1-3 are left as exercises.
- To prove property 4 that π is convex in $(p, \mathbf{w})$, let y and $\mathbf{x}$ maximize profits at p and $\mathbf{w}$, and let y' and $\mathbf{x}'$ maximize profits at p' and $\mathbf{w}'$.
- Define $p^t \equiv tp + (1-t)p'$ and $\mathbf{w}^t \equiv t\mathbf{w} + (1-t)\mathbf{w}'$ for $0 \leq t \leq 1$, and let y^* and $\mathbf{x}^*$ maximize profits at p^t and $\mathbf{w}^t$.
- Then

$$\pi(p, \mathbf{w}) = py - \mathbf{w} \cdot \mathbf{x} \geq py^* - \mathbf{w} \cdot \mathbf{x}^*$$

$$\pi(p', \mathbf{w}') = p'y' - \mathbf{w}' \cdot \mathbf{x}' \geq p'y^* - \mathbf{w}' \cdot \mathbf{x}^*$$

- So, for $0 \leq t \leq 1$,

$$\begin{aligned} t\pi(p, \mathbf{w}) + (1-t)\pi(p', \mathbf{w}') &\geq (tp + (1-t)p')y^* - (t\mathbf{w} + (1-t)\mathbf{w}')\mathbf{x}^* \\ &= p^t y^* - \mathbf{w}^t \mathbf{x}^* = \pi(p^t, \mathbf{w}^t). \end{aligned}$$

- To prove property 5 (Hotelling's lemma), we use the Envelope Theorem:

$$\pi(p, \mathbf{w}) \equiv \max_{(\mathbf{x}, y) \geq \mathbf{0}} py - \mathbf{w} \cdot \mathbf{x} \quad \text{s.t.} \quad f(\mathbf{x}) \geq y$$

$$\mathcal{L}(\mathbf{x}, y, \lambda) \equiv py - \mathbf{w} \cdot \mathbf{x} - \lambda \left[y - f(\mathbf{x}) \right],$$

$$\frac{\partial V(\mathbf{a})}{\partial a_j} = \frac{\partial \pi(p, \mathbf{w})}{\partial p} = \left. \frac{\partial \mathcal{L}}{\partial a_j} \right|_{\mathbf{x}(\mathbf{a}), \lambda(\mathbf{a})} = \frac{\partial \mathcal{L}(\mathbf{x}^*, y^*, \lambda^*)}{\partial p} = y(p, \mathbf{w})$$

$$\frac{\partial V(\mathbf{a})}{\partial a_j} = \frac{\partial \pi(p, \mathbf{w})}{\partial w_i} = \left. \frac{\partial \mathcal{L}}{\partial a_j} \right|_{\mathbf{x}(\mathbf{a}), \lambda(\mathbf{a})} = \frac{\partial \mathcal{L}(\mathbf{x}^*, y^*, \lambda^*)}{\partial w_i} = -x_i(p, \mathbf{w})$$

□

- Question: In the proof of Hotelling's lemma, where did we use the strict concavity of f assumption?

- Answer: For the Envelope Theorem to apply, $\mathbf{x}(\mathbf{a})$ must uniquely solve the constrained maximization problem.

- **THEOREM (Properties of Output Supply and Input Demand Functions)**: Suppose that f is a strictly concave production function satisfying the standard assumptions and that its associated profit function $\pi(p, \mathbf{w})$ is twice continuously differentiable. Then, for all $p > 0$ and $\mathbf{w} \gg \mathbf{0}$ where it is well defined:

1. Homogeneous of degree zero:

$$y(tp, t\mathbf{w}) = y(p, \mathbf{w}) \text{ for all } t > 0.$$

$$x_i(tp, t\mathbf{w}) = x_i(p, \mathbf{w}) \text{ for all } t > 0 \text{ and } i = 1, \ldots, n.$$

2. Own-price effects:

$$\frac{\partial y(p, \mathbf{w})}{\partial p} \geq 0,$$

$$\frac{\partial x_i(p, \mathbf{w})}{\partial w_i} \leq 0 \text{ for all } i = 1, \ldots, n.$$

- 3. The substitution matrix

$$\begin{pmatrix} \frac{\partial y(p,\mathbf{w})}{\partial p} & \frac{\partial y(p,\mathbf{w})}{\partial w_1} & \cdots & \frac{\partial y(p,\mathbf{w})}{\partial w_n} \\ -\frac{\partial x_1(p,\mathbf{w})}{\partial p} & -\frac{\partial x_1(p,\mathbf{w})}{\partial w_1} & \cdots & -\frac{\partial x_1(p,\mathbf{w})}{\partial w_n} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{\partial x_n(p,\mathbf{w})}{\partial p} & -\frac{\partial x_n(p,\mathbf{w})}{\partial w_1} & \cdots & -\frac{\partial x_n(p,\mathbf{w})}{\partial w_n} \end{pmatrix}$$

is symmetric and positive semidefinite.

- **Proof**: To prove property 1 (homogeneous), first note that the profit function $\pi(p, \mathbf{w})$ is homogeneous of degree one in $(p, \mathbf{w})$.
- Then we use the theorem: if $f(\mathbf{x})$ is homogeneous of degree k, its partial derivatives are homogeneous of degree $k-1$.
- It then follows from Hotelling's lemma

$$\frac{\partial \pi(p, \mathbf{w})}{\partial p} = y(p, \mathbf{w}) \text{ and } -\frac{\partial \pi(p, \mathbf{w})}{\partial w_i} = x_i(p, \mathbf{w}), \quad i = 1, \dots, n$$

  that both output supply $y(p, \mathbf{w})$ and input demand $x_i(p, \mathbf{w})$ are homogeneous of degree zero in $(p, \mathbf{w})$.

- To prove property 2 (own-price effects), we begin with Hotelling's lemma rearranged

$$y(p, \mathbf{w}) = \frac{\partial \pi(p, \mathbf{w})}{\partial p},$$

$$x_i(p, \mathbf{w}) = -\frac{\partial \pi(p, \mathbf{w})}{\partial w_i}, \quad i = 1, \ldots, n.$$

Because these hold for all p and $\mathbf{w}$, differentiate both sides to obtain

$$\frac{\partial y(p, \mathbf{w})}{\partial p} = \frac{\partial^2 \pi(p, \mathbf{w})}{\partial p^2},$$

$$\frac{\partial x_i(p, \mathbf{w})}{\partial w_i} = -\frac{\partial^2 \pi(p, \mathbf{w})}{\partial w_i^2}, \quad i = 1, \ldots, n.$$

- Because $\pi(p, \mathbf{w})$ is convex in $(p, \mathbf{w})$, its second-order own partials are all nonnegative, so

$$\frac{\partial y(p, \mathbf{w})}{\partial p} = \frac{\partial^2 \pi(p, \mathbf{w})}{\partial p^2} \geq 0,$$

$$\frac{\partial x_i(p, \mathbf{w})}{\partial w_i} = -\frac{\partial^2 \pi(p, \mathbf{w})}{\partial w_i^2} \leq 0, \quad i = 1, \ldots, n.$$

- To prove property 3 (substitution matrix), we need to show that the substitution matrix is equal to the Hessian matrix of second-order partials of the profit function.

- $$H = \begin{pmatrix} \frac{\partial^2 \pi(p,\mathbf{w})}{\partial p \partial p} & \frac{\partial^2 \pi(p,\mathbf{w})}{\partial p \partial w_1} & \cdots & \frac{\partial^2 \pi(p,\mathbf{w})}{\partial p \partial w_n} \\ \frac{\partial^2 \pi(p,\mathbf{w})}{\partial w_1 \partial p} & \frac{\partial^2 \pi(p,\mathbf{w})}{\partial w_1 \partial w_1} & \cdots & \frac{\partial^2 \pi(p,\mathbf{w})}{\partial w_1 \partial w_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 \pi(p,\mathbf{w})}{\partial w_n \partial p} & \frac{\partial^2 \pi(p,\mathbf{w})}{\partial w_n \partial w_1} & \cdots & \frac{\partial^2 \pi(p,\mathbf{w})}{\partial w_n \partial w_n} \end{pmatrix}$$

  $$= \begin{pmatrix} \frac{\partial y(p,\mathbf{w})}{\partial p} & \frac{\partial y(p,\mathbf{w})}{\partial w_1} & \cdots & \frac{\partial y(p,\mathbf{w})}{\partial w_n} \\ -\frac{\partial x_1(p,\mathbf{w})}{\partial p} & -\frac{\partial x_1(p,\mathbf{w})}{\partial w_1} & \cdots & -\frac{\partial x_1(p,\mathbf{w})}{\partial w_n} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{\partial x_n(p,\mathbf{w})}{\partial p} & -\frac{\partial x_n(p,\mathbf{w})}{\partial w_1} & \cdots & -\frac{\partial x_n(p,\mathbf{w})}{\partial w_n} \end{pmatrix}$$

  This must be symmetric by Young's theorem and positive semidefinite by the convexity of the profit function.

$\square$

Example: The CES profit function

- Let the production function be of the CES form

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

where $\beta < 1$, $-\infty < \rho < 1$ and $\rho \neq 0$. We solve for the corresponding profit function.

- Form the Lagrangian for the profit-maximization problem

$$\pi(p, \mathbf{w}) \equiv \max_{(\mathbf{x}, y) \geq \mathbf{0}} py - \mathbf{w} \cdot \mathbf{x} \quad \text{s.t.} \quad f(\mathbf{x}) \geq y$$

$$\mathcal{L}(x_1, x_2, y, \lambda) \equiv py - (w_1 x_1 + w_2 x_2) - \lambda \left[y - (x_1^\rho + x_2^\rho)^{\beta/\rho}\right]$$

- $$\mathcal{L}(x_1, x_2, y, \lambda) \equiv py - (w_1 x_1 + w_2 x_2) - \lambda \left[y - (x_1^\rho + x_2^\rho)^{\beta/\rho} \right]$$

  Assuming an interior solution, the first-order conditions are

  $$\frac{\partial \mathcal{L}}{\partial x_i} = -w_i + \lambda \frac{\beta}{\rho} (x_1^\rho + x_2^\rho)^{(\beta/\rho)-1} \rho x_i^{\rho-1} = 0, \quad i = 1, 2$$

  $$\frac{\partial \mathcal{L}}{\partial y} = p - \lambda = 0$$

  $$\frac{\partial \mathcal{L}}{\partial \lambda} = (x_1^\rho + x_2^\rho)^{\beta/\rho} - y = 0$$

  and simplifying yields

  $$p\beta (x_1^\rho + x_2^\rho)^{(\beta/\rho)-1} x_i^{\rho-1} = w_i, \quad i = 1, 2.$$

- It follows from

$$p\beta(x_1^\rho + x_2^\rho)^{(\beta/\rho)-1} x_i^{\rho-1} = w_i, \quad i = 1,2$$

that

$$\left(\frac{x_1}{x_2}\right)^{\rho-1} = \frac{w_1}{w_2} \quad \text{or} \quad \frac{x_1}{x_2} = \left(\frac{w_1}{w_2}\right)^{1/(\rho-1)}.$$

- Substituting into the constraint yields

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

$$y^{1/\beta} = \left(x_2^\rho \left(\frac{w_1}{w_2}\right)^{\rho/(\rho-1)} + x_2^\rho\right)^{1/\rho} \left[\left(\frac{w_2}{w_2}\right)^{\rho/(\rho-1)}\right]^{1/\rho}$$

$$y^{1/\beta} = \frac{x_2}{w_2^{1/(\rho-1)}} \left(w_1^{\rho/(\rho-1)} + w_2^{\rho/(\rho-1)}\right)^{1/\rho}$$

- So rearranging, using $r \equiv \rho/(\rho - 1)$ gives the conditional input demands

$$x_2 = y^{1/\beta} w_2^{1/(\rho-1)} \left(w_1^r + w_2^r\right)^{-1/\rho}$$

$$x_1 = x_2 \left(\frac{w_1}{w_2}\right)^{1/(\rho-1)} = y^{1/\beta} w_1^{1/(\rho-1)} \left(w_1^r + w_2^r\right)^{-1/\rho}$$

- Substituting into the first-order condition

$$p\beta(x_1^\rho + x_2^\rho)^{(\beta/\rho)-1} x_1^{\rho-1} = w_1$$

using $y = (x_1^\rho + x_2^\rho)^{\beta/\rho}$ yields

$$p\beta \left(y^{\rho/\beta}\right)^{(\beta/\rho)-1} y^{(\rho-1)/\beta} w_1 \left(w_1^r + w_2^r\right)^{-(\rho-1)/\rho} = w_1$$

- $$p\beta \left(y^{\rho/\beta}\right)^{(\beta/\rho)-1} y^{(\rho-1)/\beta} w_1 \left(w_1^r + w_2^r\right)^{-(\rho-1)/\rho} = w_1$$
$$p\beta y^{(\beta-1)/\beta} \left(w_1^r + w_2^r\right)^{-(\rho-1)/\rho} = 1$$
$$y = (p\beta)^{-\beta/(\beta-1)} \left(w_1^r + w_2^r\right)^{\beta(\rho-1)/\rho(\beta-1)}$$

  This yields the output supply function.

- $$y = (p\beta)^{-\beta/(\beta-1)} \left(w_1^r + w_2^r\right)^{\beta(\rho-1)/\rho(\beta-1)}$$

  Now substituting for y using the output supply function back into the conditional input demands:

  $$x_i = y^{1/\beta} w_i^{1/(\rho-1)} \left(w_1^r + w_2^r\right)^{-1/\rho}$$

  $$x_i = (p\beta)^{-1/(\beta-1)} \left(w_1^r + w_2^r\right)^{(\rho-1)/\rho(\beta-1)} \cdot w_i^{1/(\rho-1)} \left(w_1^r + w_2^r\right)^{-1(\beta-1)/\rho(\beta-1)}$$

  $$x_i = w_i^{1/(\rho-1)} (p\beta)^{-1/(\beta-1)} \left(w_1^r + w_2^r\right)^{(\rho-\beta)/\rho(\beta-1)}$$

  This yields the input demand functions.

- Next note that

$$w_i x_i = w_i^{(\rho-1+1)/(\rho-1)} (p\beta)^{-1/(\beta-1)} (w_1^r + w_2^r)^{(\rho-\beta)/\rho(\beta-1)}$$

$$= w_i^r (p\beta)^{-1/(\beta-1)} (w_1^r + w_2^r)^{(\rho-\beta)/\rho(\beta-1)}$$

so

$$w_1 x_1 + w_2 x_2 = (w_1^r + w_2^r) (p\beta)^{-1/(\beta-1)} (w_1^r + w_2^r)^{(\rho-\beta)/\rho(\beta-1)}$$

$$= (p\beta)^{-1/(\beta-1)} (w_1^r + w_2^r)^{(\rho\beta-\rho+\rho-\beta)/\rho(\beta-1)}$$

$$= (p\beta)^{-1/(\beta-1)} (w_1^r + w_2^r)^{\beta(\rho-1)/\rho(\beta-1)}$$

- Finally, we use this information to form the profit function

$$\pi(p, \mathbf{w}) = py - (w_1 x_1 + w_2 x_2)$$

$$= \left[p\,(p\beta)^{-\beta/(\beta-1)} - (p\beta)^{-1/(\beta-1)}\right](w_1^r + w_2^r)^{\beta(\rho-1)/\rho(\beta-1)}$$

$$= \left[p^{(\beta-1)/(\beta-1)}\,(p\beta)^{-\beta/(\beta-1)} - (p\beta)^{-1/(\beta-1)}\right](w_1^r + w_2^r)^{\beta/r(\beta-1)}$$

$$= p^{-1/(\beta-1)}\left[\beta^{-\beta/(\beta-1)} - \beta^{-1/(\beta-1)}\right](w_1^r + w_2^r)^{\beta/r(\beta-1)}$$

$$= p^{-1/(\beta-1)}\beta^{-\beta/(\beta-1)}(1-\beta)\,(w_1^r + w_2^r)^{\beta/r(\beta-1)} > 0$$

The profit function is well-defined for $\beta < 1$ (decreasing returns to scale) but not defined for $\beta = 1$ (constant returns to scale).

History of Economic Ideas

- The elasticity of substitution is due to John Hicks (1932, *Theory of Wages*).
- The properties of the cost function were developed by several authors, but the most systematic treatment is in Ronald Shephard (1953, *Cost and Production Functions*). Shephard's lemma was first presented in this book, as was the basic duality between cost and production functions.
- The properties of the profit function were developed by Harold Hotelling (1932), John Hicks (1946, *Value and Capital*) and Paul Samuelson (1947, *Foundations of Economic Analysis*).