Lecture Note 5: Production

Ting Liu

Econ 500 – Microeconomic Theory I

State University of New York– Stony Brook

1 Introduction

In this note we introduce the production side of the economy. We will start our analysis from a very simplified setting in which the firm takes prices as given and makes input and output decisions with the goal of maximizing profit. This is meant to capture situations in which

- there are many firms
- homogeneous product
- each firm is a price taker

. We will focus attention on single-product firms, that is firms producing a single output with (possibly) many inputs.

2 Technology

- Let y denote the output of the firm, and $(x_1, x_2 \ldots, x_n)$ the inputs. We will adopt the convention that both y and the x_i's are nonnegative.

- We will be interested in the **frontier** of the production set, that is in vectors which generate the maximum output for a given level of inputs. The frontier can be described by a *production function*:

$$y = f(x_1, x_2 \ldots, x_n).$$

Given the current technology and the firm's organization, y is the maximum outputs the firm can produce.

- The production function plays a role similar to the one played by the utility function in consumer theory; notice however that production is a physical quantity, so that the exact amount produced is important. *Monotonic transformations are not allowed* because it is not ordinal.

- When we have only two inputs, we can use a graphical representation similar to the indifference curves we have seen in consumer theory. An *isoquant* is the set of input levels such that a certain production level $\overline{y}$ is achieved. Formally, if $f(x_1, x_2)$ is the production function then the isoquant is the set of points:

$$I(\overline{y}) = \{(x_1, ..., x_n)| \, f(x_1, ..., x_n) = \overline{y}\}.$$

Higher levels of $\overline{y}$ are associated to higher isoquants.Figure representation...

- Perfectly substitutes, Leontiff production funcation, Cobb Doglas production function. Examples: two men and a truck moving company has the leontiff production function and Potatoes from different states are perfect substitutes for McDonald's french fry.

- Isoquant has similar properties as indifference curve i) an isoquant farther away from the origin is associated with a higher output ii) isoquants do not cross iii) isoquants are downward sloping

- The *marginal product* of input x_i is defined (much in the same way as marginal utility) as the unit increase Δy in production which is obtained when input x_i only is increased by Δx_i, that is $\frac{f(x_1,..x_i+\Delta x_i,...x_n)-f(x_1,..x_i,...x_n)}{\Delta x_i}$. When the increment Δx_i goes to zero, the marginal product can be expressed as the partial derivative $\frac{\partial f}{\partial x_i}$.

- It is customary to distinguish between a '*short-run*,' and a '*long-run*' problem. The difference is that in the short run some factors are hard to change (e.g. size of the plant) and they are therefore considered fixed. The firm therefore maximizes over a limited number of variables (e.g. raw materials). Thus, the difference between the two problems lies only in the number of variables that we can control. Which inputs are fixed in the short run depend on the nature of the industry. Usually, the capitla input is fixed in manufacture industry. By contrast, labor input is more likely to be fixed in some service industry. Take the university as another example. It is difficult to find qualified faculties in a short period of time but is realtively easier to get more facilites or rent office spaces.

- In the short run, some inputs are fixed. For example, capital input in manufacture industry is fixed in the short run. That is, $f\left(L,\overline{K}\right)$. In this case, **marginal product** of labor first increases due to specialization but eventually decreases due to *law of diminishing returns.* Thomas Malthus predicted mass starvation on the basis of *law of diminishing returns.*

- In the long run, firms can adjust all the inputs. A firm's willingess to substitute one input for another while holding output constant is captured by *Marginal Rate of Technical Substitution.* The *technical rate of substitution* between factor x_i and factor x_j is the increase Δx_j necessary to compensate a decrease Δx_i, if we want to keep production constant. From the graphical point of view, it is the slope of the isoquant. To obtain an analytical expression, consider infinitesimal changes dx_i and dx_j such that production is unchanged, that is $dy = 0$. Since:

$$dy = \frac{\partial f}{\partial x_i} dx_i + \frac{\partial f}{\partial x_j} dx_j$$

when we set $dy = 0$ we obtain:

$$\frac{dx_i}{dx_j} = -\frac{\frac{\partial f}{\partial x_j}}{\frac{\partial f}{\partial x_i}}$$

so that the technical rate of substitution is given by the ratio of marginal products.

Example. Consider the Cobb-Douglas technology $y = x_1^{0.4} x_2^{0.5}$. Then the equation of the isoquant curve of level $\overline{y}$ is $x_2 = \frac{\overline{y}^2}{x_1^{0.8}}$. The marginal product of the first input is $\frac{\partial f}{\partial x_1} = 0.4 x_1^{-0.6} x_2^{0.5}$, and the marginal product of the second input is $\frac{\partial f}{\partial x_2} = 0.5 x_1^{0.4} x_2^{-0.5}$. The technical rate of substitution is given by $\frac{dx_2}{dx_1} = -\frac{0.4 x_1^{-0.6} x_2^{0.5}}{0.5 x_1^{0.4} x_2^{-0.5}} = -0.8 \frac{x_2}{x_1}$.

- Isoquant is convex because of the *law of diminishing returns.* Take the labor & capitpal input as an example. As a firm has more employees and fewer capital input, it needs more labor input to compensate for a unit loss in capital input because labor comes less productive compared with capital.

- In the long run, firms can adjust all the inputs. An important notion in production theory is that of *returns to scale.* Roughly, returns to scale indicate how much production is increased when all inputs are increased by the same percentage. In particular, let $t > 1$ be the scale by which inputs are increased. Then:

- A production function has *decreasing* returns to scale if $f(tx_1, \ldots, tx_n) < tf(x_1, \ldots, x_n)$.

- A production function has *constant* returns to scale if $f(tx_1, \dots, tx_n) = tf(x_1, \dots, x_n)$.
- A production function has *increasing* returns to scale if $f(tx_1, \dots, tx_n) > tf(x_1, \dots, x_n)$.

A production function may display different returns to scale at different points. Usually a firm has increasing return to scale when it is small and constant return to scale and decreasing return to scale when it grows bigger and bigger. This is because that when the firm is small, it benefit from specilization when it has more of every inputs. For example, an auto maker initially produce all the cars in one factory. When it has two factories, it can specialize in SUV in one factory and Sedan in another factory. Both factories will become more productive due to specialization.However, when the firm become bigger, it is harder for the manager to monitor each worker and it becomes less and less efficient.

Example. For a Cobb-Douglas technology $y = x_1^\alpha x_2^\beta$ we have

$$f(tx_1, tx_2) = (tx_1)^\alpha (tx_2)^\beta = t^{\alpha+\beta} x_1^\alpha x_2^\beta = t^{\alpha+\beta} f(x_1, x_2).$$

Therefore, if $\alpha + \beta < 1$ returns to scale are decreasing, if $\alpha + \beta = 1$ returns to scale are constant and if $\alpha + \beta > 1$ returns to scale are increasing.

3 Profit Maximization

- The firm's problem is to choose the variables y, $x_1, \dots, x_n$ to maximize profit subject to technological constraints. Let p be the price of output, and $w_1, \dots, w_n$ the prices of inputs. Profit is given by:

$$\underbrace{\pi(y, x_1, \dots, x_n)}_{\text{Profit}} \underset{=}{=} \underbrace{py}_{\text{Revenue}} \underset{-}{-} \underbrace{\sum w_i x_i}_{\text{Costs}}.$$

The problem is therefore:

$$\max_{x_1,\ldots,x_n} \quad pf\left(x_1,\ldots,x_n\right) - w_1x_1 - \ldots - w_nx_n.$$

S.t.

$$x_i \geq 0$$

where f is the production function.

- The first order condition which has to be satisfied for each variable x_i

$$p\frac{\partial f\left(x_1,\ldots,x_n\right)}{\partial x_i} \leq w_i, \text{ with equality if } x_i > 0.$$

Let's focus on the interior solution. Remember that $\frac{\partial f}{\partial x_i}$ is the extra quantity that we obtain when we increase x_i. This extra quantity yields an extra value of $p\frac{\partial f}{\partial x_i}$, which is the marginal revenue obtained. The extra cost is w_i, the unit cost of factor x_i. Therefore, the condition says that x_i should be increased up to the point at which the extra cost of one unit of input equals the extra revenue generated by that unit of input.

- Another way to interpret the optimal solution is "last dollar rule"

$$\frac{\frac{\partial f}{\partial x_i}}{\frac{\partial f}{\partial x_j}} = \frac{w_i}{w_j}$$

$$\frac{\frac{\partial f}{\partial x_i}}{w_i} = \frac{\frac{\partial f}{\partial x_j}}{w_j}$$

- If the function $f\left(x_1,\ldots,x_n\right)$ is concave then the first order conditions are also sufficient. If not, we will have to be careful about second order conditions. At any rate, the solution to the maximization problem yields *factor demand curves*:

$$x_1^*\left(p,w_1,\ldots,w_n\right),\ldots,x_n^*\left(p,w_1,\ldots,w_n\right)$$

which are similar to the demand curves that we have seen for consumer theory. Thus, $x_i^*(p, w_1, \ldots, w_n)$ describes the optimal quantity of input x_i chosen by the firm in order to maximize profits when the price of output is p and the input prices are $w_1, \ldots, w_n$.

- The profit function $\pi^*(p, w_1, \ldots, w_n)$ is defined by:

$$\pi^*(p, w_1, \ldots, w_n) = pf(x_1^*, \ldots, x_n^*) - w_1 x_1^* - \ldots - w_n x_n^*$$

where the x_i^* are the demand functions. The demand functions and the profit functions have the following properties.

- **The factor demand functions** $x_1^*(p, w_1, \ldots, w_n)$ **are homogeneous of degree zero.**To see this, suppose we have $(tp, tw_1, \ldots, tw_n)$. The firm's maximization problem becomes

$$\max_{x_1,\ldots,x_n} \quad t[pf(x_1, \ldots, x_n) - w_1 x_1 - \ldots - w_n x_n],$$

which will yeild the same factor demand fucntion because t is just a scalar.

- **The profit function** $\pi^*(p, w_1, \ldots, w_n)$ **is homogeneous of degree 1**.So, $\pi^*(tp, tw_1, \ldots, tw_n) = t\pi^*(p, w_1, \ldots, w_n)$.

- **The profit function** $\pi^*(p, w_1, \ldots, w_n)$ **is convex**. To see this, consider two vectors (p^A, w^A) and (p^B, w^B), and their convex combination

$$(p^\alpha, w^\alpha) = \left(\alpha p^A + (1-\alpha) p^B, \alpha w^A + (1-\alpha) w^B\right).$$

Let x^A, x^B and x^α be the input demands corresponding to the three price vectors and y^A, y^B and y^α be the corresponding output. Then

$$\begin{aligned}
\pi\left(p^\alpha, w^\alpha\right) &= p^\alpha y^\alpha - w^\alpha x^\alpha \\
&= (\alpha p^A + (1-\alpha)\, p^B) y^\alpha - (\alpha w^A + (1-\alpha)\, w^B) x^\alpha \\
&= \alpha(p^A y^\alpha - w^A x^\alpha) + (1-\alpha)\left(p^B y^\alpha - w^B x^\alpha\right) \\
&\leq \alpha(p^A y^A - w^A x^A) + (1-\alpha)\left(p^B y^B - w^B x^B\right)
\end{aligned}$$

where the inequality follows from the fact that (y^α, x^α) is not necessarily the optimal demand vector at $\left(p^A, w^A\right)$ or $\left(p^B, w^B\right)$. **Convexity of π implies that firms like uncertainty in price**.

- When the production function f is concave then the factor demand correspondence is convex. Strict quasi-concavity of f corresponds to the case in which x^* is single-valued.

Proof: Suppose $x_1^*, x_2^* = \arg\max_x pf\left(x\right) - w.x$.So, $\pi\left(x_1^*\right) = \pi\left(x_2^*\right) = \pi^*$. Consider $x^\alpha = \alpha x_1^* + (1-\alpha) x_2^*$, with $\alpha \in (0,1)$. The profit associated with x^α is

$$\begin{aligned}
& pf(x^\alpha) - wx^\alpha \\
\geq\ & p[\alpha f(x_1^*) + (1-\alpha) f(x_2^*)] - w[\alpha x_1^* + (1-\alpha) x_2^*] \\
=\ & \alpha(pf(x_1^*) - wx_1^*) + (1-\alpha)(pf(x_2^*) - wx_2^*) \\
=\ & \alpha\pi^* + (1-\alpha)\pi^* \\
=\ & \pi^*.
\end{aligned}$$

Hence, $x^\alpha = \arg\max_x pf\left(x\right) - w.x.$. When $f(.)$ is strictly concave, then

$$pf(x^\alpha) - wx^\alpha > \pi^*,$$

which contradicts the assumption that $x_1^*, x_2^* = \arg\max_x pf\left(x\right) - w.x$. Q.E.D.

- When x^* is single valued then by the envelope theorem $\frac{\partial \pi}{\partial w_i} = -x_i^*$ and $\frac{\partial \pi}{\partial p} = f\left(x_1^*, \ldots, x_n^*\right)$.To see this,

$$\begin{aligned}
\frac{\partial \pi}{\partial w_i} &= p \sum_{k=1}^{n} \frac{\partial f\left(x_1^*, \ldots, x_n^*\right)}{x_k} \frac{\partial x_k^*(p, w_1, ..., w_n)}{\partial w_i} \\
&\quad - \sum_{k=1}^{n} w_k \frac{\partial x_k^*(p, w_1, ..., w_n)}{\partial w_i} - x_i^*.
\end{aligned}$$

By the F.O.C., we have

$$\begin{aligned}
\frac{\partial \pi}{\partial w_i} &= \sum_{k=1}^{n} w_k \frac{\partial x_k^*(p, w_1, ..., w_n)}{\partial w_i} - \sum_{k=1}^{n} w_k \frac{\partial x_k^*(p, w_1, ..., w_n)}{\partial w_i} - x_i^* \\
&= -x_i^*.
\end{aligned}$$

$$p \frac{\partial f}{\partial x_i} = w_i.$$

Another way to see this is

$$\begin{aligned}
\frac{\partial \pi}{\partial w_i} &= \sum_{k=1}^{n} \frac{\partial \pi}{\partial x_k} \frac{\partial x_k}{\partial w_i} - x_i \\
&= -x_i.
\end{aligned}$$

The second equality follows from the Envelope Theorem.

- Similarly,

$$\begin{aligned}
\frac{\partial \pi}{\partial p} &= f\left(x_1^*, \ldots, x_n^*\right) + p \sum_{k=1}^{n} \frac{\partial f\left(x_1^*, \ldots, x_n^*\right)}{x_k} \frac{\partial x_k^*(p, w_1, ..., w_n)}{\partial p} \\
&\quad - \sum_{k=1}^{n} w_k \frac{\partial x_k^*(p, w_1, ..., w_n)}{\partial p} \\
&= f\left(x_1^*, \ldots, x_n^*\right) + \sum_{k=1}^{n} w_k \frac{\partial x_k^*(p, w_1, ..., w_n)}{\partial p} \\
&\quad - \sum_{k=1}^{n} w_k \frac{\partial x_k^*(p, w_1, ..., w_n)}{\partial p} \\
&= f\left(x_1^*, \ldots, x_n^*\right).
\end{aligned}$$

- We can establish the following *law of supply*. Let (p^a, w^a), $\left(p^b, w^b\right)$ be two price vectors, and let $(y^a, -x^a)$, $\left(y^b, -x^b\right)$ the corresponding supplies. Then

$$\left((p^a, w^a) - \left(p^b, w^b\right)\right)\left((y^a, -x^a) - \left(y^b, -x^b\right)\right) \geq 0.$$

- This follows from

$$\begin{aligned} & \left((p^a, w^a) - \left(p^b, w^b\right)\right)\left((y^a, -x^a) - \left(y^b, -x^b\right)\right) \\ = \ & (p^a, w^a)\left((y^a, -x^a) - \left(y^b, -x^b\right)\right) + \\ & \left(p^b, w^b\right)\left(\left(y^b, -x^b\right) - (y^a, -x^a)\right) \\ \geq \ & 0 \end{aligned}$$

  because each vector is profit maximizing at the corresponding prices. This in particular implies that **when it is only the price of the output to increase then the supply of the output increases**, and that **when it is only the price of an input to increase then the demand of that input decreases.**

- A factor demand curve is *always* downward sloping with respect to the price of the factor. From this point of view, it is better behaved that the consumer's demand function, the reason being that no wealth effects appear.

4 Cost Functions

One key assumption in the preceding section was that the output price p was given, and that the firm could sell any desired quantity at that price. In other words, the firm was operating in a competitive market. Notice also that the firm was simultaneously deciding the quantity y and the inputs $x_1, \ldots, x_n$.

- An alternative approach is the following.

1. First, for any given level of production y find out the cost-minimizing combination of factors producing y. This will depend on the factor prices and on the quantity to be produced, so we will have functions $x_1^*(w_1, \ldots, w_n, y), \ldots, x_n^*(w_1, \ldots, w_n, y)$. Figure representation.

- These are called *conditional (or derived) factor demands*. They differ from the factor demands that we have seen in the previous section because the quantity to be produced y is taken as given.

- Define the cost function of the firm as:

$$c(w_1, \ldots, w_n, y) = \sum_{i=1}^{n} w_i x_i^*(w_1, \ldots, w_n, y).$$

The cost function $c(w_1, \ldots, w_n, y)$ tells us what is the lowest cost, at given factor prices, to produce y

2 Then chooses y which maixmizes

$$\pi = py - c(w_1, \ldots, w_n, y).$$

- This approach has the advantage of being useful also in non-competitive markets for the product. In a competitive market, revenue is simply py. In a non-competitive market the price may depend on y, so that revenue is a more complicated function $p(y)y$. At any rate, it will remain true that for any level of y chosen the firm should select inputs so as to minimize the cost of production. Figure illustration of $MR(y)$ under competitive market and monopoly market, respectively.

- Let us now take a more formal look at the problem of cost minimization. The problem can be formally written as:

$$\min_{x_1,\ldots,x_n} \qquad w_1x_1 + w_2x_2 + \ldots + w_nx_n$$

s.t.

$$f(x_1,\ldots,x_n) \geq y$$

where y is the required level of production. The constraint must be binding at the optimum. The Lagrangian is:

$$L = w_1x_1 + w_2x_2 + \ldots + w_nx_n - \lambda\left(f(x_1,\ldots,x_n) - y\right)$$

so that the first order condition for x_i is:

$$w_i = \lambda \frac{\partial f}{\partial x_i}$$

Dividing side by side the FOCs for x_i and x_j we obtain

$$\begin{aligned} \frac{w_i}{w_j} &= \frac{\frac{\partial f}{\partial x_i}}{\frac{\partial f}{\partial x_j}} \\ \frac{\frac{\partial f}{\partial x_j}}{w_j} &= \frac{\frac{\partial f}{\partial x_i}}{w_i} = \frac{1}{\lambda}, \quad \text{"last dollar rule!"} \end{aligned}$$

Figure illustration of two inputs. This says that at the cost-minimizing point, for any pair of inputs the ratio of the factor prices has to equal the (negative of) the marginal rate of transformation (this is very similar to the equality between MRS and relative prices in consumer theory).

- The first order condition is sufficient if $f(x)$ is concave.

- Firms' cost minimization problem is similar to expenditure minimization problem in consumer theory.

Example. Suppose that the production function is $y = (x_1^\alpha + x_2^\alpha)^\rho$ with $\rho \in (0,1)$, $\alpha \in (0,1)$. The Lagrangian is:

$$L = w_1 x_1 + w_2 x_2 - \lambda \left[(x_1^\alpha + x_2^\alpha)^\rho - y \right]$$

and the FOCs are:

$$\begin{aligned} w_1 &= \lambda \rho \left(x_1^\alpha + x_2^\alpha \right)^{\rho - 1} \alpha x_1^{\alpha - 1} \\ w_2 &= \lambda \rho \left(x_1^\alpha + x_2^\alpha \right)^{\rho - 1} \alpha x_2^{\alpha - 1} \end{aligned}$$

Hence:

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

together with

$$y = (x_1^\alpha + x_2^\alpha)^\rho$$

we can find x_1

$$\left(x_1^\alpha + \left(\frac{w_1}{w_2} \right)^{\frac{\alpha}{1-\alpha}} x_1^\alpha \right)^\rho = y \qquad \rightarrow \qquad x_1 = \frac{y^{\frac{1}{\alpha \rho}}}{\left(1 + \left(\frac{w_1}{w_2} \right)^{\frac{\alpha}{1-\alpha}} \right)^{\frac{1}{\alpha}}}$$

Therefore, the conditional demands for factors are:

$$\begin{aligned} x_1^* (w_1, w_2, y) &= \left(1 + \left(\frac{w_1}{w_2} \right)^{\frac{\alpha}{1-\alpha}} \right)^{-\frac{1}{\alpha}} y^{\frac{1}{\alpha \rho}} \\ x_2^* (w_1, w_2, y) &= \left(\frac{w_1}{w_2} \right)^{\frac{1}{1-\alpha}} \left(1 + \left(\frac{w_1}{w_2} \right)^{\frac{\alpha}{1-\alpha}} \right)^{-\frac{1}{\alpha}} y^{\frac{1}{\alpha \rho}} \end{aligned}$$

The cost function is:

$$\begin{aligned} c (w_1, w_2, y) &= \left(w_1 + w_2 \left(\frac{w_1}{w_2} \right)^{\frac{1}{1-\alpha}} \right) \left(1 + \left(\frac{w_1}{w_2} \right)^{\frac{\alpha}{1-\alpha}} \right)^{-\frac{1}{\alpha}} y^{\frac{1}{\alpha \rho}} \\ &= w_1 \left(1 + \left(\frac{w_1}{w_2} \right)^{\frac{\alpha}{1-\alpha}} \right)^{1 - \frac{1}{\alpha}} y^{\frac{1}{\alpha \rho}} \end{aligned}$$

The cost function $c(w_1, \ldots, w_n, y)$ and the conditional demand $x_i(w_1, \ldots, w_n, y)$ satisfy the following properties:

- $x(w_1, \ldots, w_n, y)$ **is homongenuous degree zero in factor prices** $(w_1, \ldots, w_n)$.((Use a figure with two inputs to show it)

- $c(w_1, \ldots, w_n, y)$ **is homogeneous of degree 1 in factor prices** $(w_1, \ldots, w_n)$ **and nondecreasing in** y (proof by contradiction). The first statement follows from the fact that $x(w_1, \ldots, w_n, y)$ is homongenuous degree zero in factor prices

- $c(w_1, \ldots, w_n, y)$ **is a concave function of** $(w_1, \ldots, w_n)$. The intutition is the same as the concavity of the expenditure function $e(p, u)$. Figure illustration.

Proof: consider w^1(factor price vector) and w^2. The corresponding conditional factor demand is $x(w^1, y)$ and $x(w^2, y)$. Consider a new price vector $\alpha w^1 + (1-\alpha)w^2$. Its corresponding conditinal factor demand is $x(\alpha w^1 + (1-\alpha)w^2, y)$.We have

$$\begin{aligned} c(\alpha w^1 + (1-\alpha)w^2, y) &= [\alpha w^1 + (1-\alpha)w^2].x(\alpha w^1 + (1-\alpha)w^2, y) \\ &= \alpha w^1.x(\alpha w^1 + (1-\alpha)w^2, y) \\ &\quad +(1-\alpha)w^2.x(\alpha w^1 + (1-\alpha)w^2, y) \\ &\geq \alpha w^1 x(w^1, y) + (1-\alpha)w^2 x(w^2, y) \\ &= \alpha c(w^1, y) + (1-\alpha)c(w^2, y) \end{aligned}$$

- **When c is differentiable,** $\dfrac{\partial c}{\partial w_i} = x_i\left(w_1, \ldots, w_n, y\right)$ (this is called Shepard's lemma). To see this

$$
\begin{aligned}
\frac{\partial c}{\partial w_i} &= x_i + \sum_{k=1}^{n} w_k \frac{\partial x_k(w,y)}{\partial w_i} \\
&= x_i + \lambda \sum_{k=1}^{n} \frac{\partial f}{\partial x_k} \frac{\partial x_k(w,y)}{\partial w_i}, \text{ by the F.O.C.} \\
&= x_i.
\end{aligned}
$$

The last equality follows by total differentiate the production constraint with respect to w_i

$$
\begin{aligned}
\frac{\partial f(x_1(w,y), ..., x_n(w,y))}{\partial w_i} &= 0 \\
\sum_{k=1}^{n} \frac{\partial f}{\partial x_k} \frac{\partial x_k(w,y)}{\partial w_i} &= 0.
\end{aligned}
$$

- Additional properties of the cost functions are obtained by making assumptions on the technology. Suppose that $f(x)$ is homogeneous of degree 1 (constant return to scale). Then, the conditional demand function $x(w,y)$ are also homogeneous of degree 1 in y. To see this, observe that when the technology has constant returns to scale, if x is such that $y = f(x)$ then αx is such that $\alpha y = f(\alpha x)$. But now observe that if x minimizes at w the cost of producing y, it must be the case that αx minimizes the cost of producing αy at the same prices. Consider the cost minimization problem

$$\min_x w.x$$

s.t.

$$f(x) = y.$$

The solution of this minimization program is identical to solution of

$$\min_x w.\alpha x$$

s.t.

$$\begin{aligned} \alpha f(x) &= \alpha y \\ &\Leftrightarrow \\ f(\alpha x) &= \alpha y \end{aligned}$$

Hence, $c(\alpha y) = w.\alpha x = \alpha w.x = \alpha c(y)$.

- If $f(x)$ is strictly concave (convex) then $c(w, y)$ is strictly convex (concave) in y. [HW problem]

Proof: Consider $y = f(x), y' = f(x')$ such that $c(y) = w.x$ and $c(y') = w.x'$. For any $\alpha \in (0, 1)$, consider the cost minimization problem

$$\min_x wx$$

s.t.

$$f(x) = \alpha y + (1-\alpha)y'.$$

First, consider strictly concave $f(x)$.

$$\begin{aligned} f(\alpha x + (1-\alpha)x') &> \alpha f(x) + (1-\alpha)f(x') \\ &= \alpha y + (1-\alpha)y'. \end{aligned}$$

Hence,

$$\begin{aligned} c(\alpha y + (1-\alpha)y') &< w.(\alpha x + (1-\alpha)x') \\ &= \alpha w.x + (1-\alpha)w.x' \\ &= \alpha c(y) + (1-\alpha)c(y'). \end{aligned}$$

Now, consider strictly convex $f(x)$. Suppose

$$\begin{aligned}
c(\alpha y + (1-\alpha)y') &\leq \alpha c(y) + (1-\alpha)c(y') \\
&= \alpha w.x + (1-\alpha)w.x' \\
&= w.(\alpha x + (1-\alpha)x') \Leftrightarrow \\
w.x(w, \alpha y + (1-\alpha)y') &\leq w.(\alpha x + (1-\alpha)x') \Leftrightarrow \\
x(w, \alpha y + (1-\alpha)y') &\leq \alpha x + (1-\alpha)x'.
\end{aligned}$$

Since $f()$ is nondecreasing,

$$\begin{aligned}
f(x(w, \alpha y + (1-\alpha)y')) &\leq f(\alpha x + (1-\alpha)x') \\
&< \alpha f(x) + (1-\alpha)f(x') \text{ by convexity of } f() \\
&= \alpha y + (1-\alpha)y'.
\end{aligned}$$

This contradicts that $x(w, \alpha y + (1-\alpha)y')$ is the conditional factor demand for $\alpha y + (1-\alpha)y'$.

5 The Cost Function and the Supply of the Firm

When factor prices are taken as given, we can write the cost function as $c(y)$, that is simply as a function of the level of production. From the cost function we can derive two other curves which play an important role in the analysis of profit maximization:

- The *marginal cost* function gives the cost, at any given level of y of slightly increasing production. It is obtained as the derivative of the cost function, that is $MC(y) = c'(y)$.
- The *average cost* function gives the mean cost of producing y units. It is given by $AC(y) = \frac{c(y)}{y}$.

Example continued. If we define $k = w_1 \left(1 + \left(\frac{w_1}{w_2}\right)^{\frac{\alpha}{1-\alpha}}\right)^{1-\frac{1}{\alpha}}$ then the cost function previously obtained can be written as $c(y) = ky^{\frac{1}{\alpha\rho}}$. The marginal cost function is $MC(y) = \frac{k}{\alpha\rho} y^{\frac{1}{\alpha\rho}-1}$. The average cost function is $AC(y) = ky^{\frac{1}{\alpha\rho}-1}$. Suppose $\alpha\rho < 1$. Figure of AC and MC.

Suppose now that the firm requires a fixed set up cost F. The optimal choices of x_1 and x_2 do not change. The cost function become $c(y) = cy^{\frac{1}{\alpha\rho}} + F$. The marginal cost function remains the same as before, since the derivative does not change. The average cost however changes, it is now $AC(y) = cy^{\frac{1}{\alpha\rho}-1} + \frac{F}{y}$, and it has an U-shaped curve. It is customary to call average variable cost the first term and average fixed cost the second term.Figure of AC and MC.

- Notice the following:

$$\begin{aligned} AC'(y) &= d(\frac{C(y)}{y})/dy \\ &= \frac{C'(y)y - C(y)}{y^2} \\ &= \frac{C'(y) - AC(y)}{y}. \end{aligned}$$

Hence, $AC'(y) = 0$ iff $C'(y) = AC(y)$; $AC'(y) > 0$ iff $C'(y) > AC(y)$; $AC'(y) > 0$ iff $C'(y) < AC(y)$. Figure

The form of the cost functions is related to the returns to scale in the following way:

- If returns to scale are *decreasing* then the average cost function $AC(y)$ must be increasing. This is called *diseconomies of scale.* This follows from the fact that an increase in the scale of production requires a more than proportional increase in inputs. So, $MC(y) > AC(y)$.

- If returns to scale are *constant* then average cost is also constant (*no economies of scale*), that is $AC(y) = c$, for some fixed value c. This in turn implies that the cost function is linear, $c(y) = cy$, and that $MC(y) = AC(y)$. Proof:

- If returns to scale are *increasing* then the average cost function $AC(y)$ must be decreasing (*economies of scale*), and in this case $MC(y) < AC(y)$.

- The standard textbook exposition involves average cost curves which are U-shaped. One case in which such a form arises is when there are fixed set up costs and a convex variable cost function.

- If we take the factor prices as given, then the maximization problem of the competitive firm can be written as:

$$\max_{y \geq 0} \qquad py - c(y).$$

If the solution involves $y > 0$ then the first order condition is

$$p = c'(y) \tag{1}$$

that is, the price should equal marginal cost.

- Notice that the second order condition requires $c''(y) \geq 0$, which means that at the optimal point the marginal cost has to be increasing .

- **Example.** If $y = (x_1^\alpha + x_2^\alpha)^\rho$ then $((tx_1)^\alpha + (tx_2)^\alpha)^\rho = t^{\alpha\rho}(x_1^\alpha + x_2^\alpha)^\rho$, so that we have decreasing returns if $\alpha\rho < 1$. The average cost function $AC(y) = cy^{\frac{1}{\alpha\rho}-1}$ is increasing if $\alpha\rho < 1$, and in that case $MC(y) = \frac{c}{\alpha\rho}y^{\frac{1}{\alpha\rho}-1} > AC(y)$. If $\alpha\rho = 1$ then returns to scale are constant, and in this case we have $AC(y) = MC(y) = c$. (The case of $\alpha\rho > 1$ increasing return to scale, the second order conditions are not satisfied.)

At last, consider the case in which $c(y)$ is differentiable and $AC(y)$ reaches a minimum, say at y^*. Then it must be the case that $MC(y^*) = AC(y^*)$. This comes from the condition that $\frac{dAC}{dy} = 0$.

- The really well-behaved case is the one with decreasing returns. In this case, $AC(y)$ is increasing and hence $MC(y)$ is increasing. So, the second order condition is always satisfied.

- Suppose there are increasing returns to scale. In this case, the average cost is decreasing. In this case the optimal point is either zero production (when the price is low enough) or production at full capacity. When $y > 0$ we can write profit as $\pi(y) = y[p - AC(y)]$. Now suppose that profit is maximized at some level $y^* > 0$. Consider now slightly increasing the quantity produced at $y^* + \varepsilon$. Then:

$$\pi(y^* + \varepsilon) = (y^* + \varepsilon)[p - AC(y^* + \varepsilon)] > y^*[p - AC(y^*)] = \pi(y^*)$$

where the inequality follows from the fact that AC is decreasing. Thus, if a positive production is optimal it must be the highest possible production. Example: health insuranc market.

- When there are constant returns to scale then $\pi(y) = (p - c)y$. In this case, optimal production is zero if $p < c$; it is indeterminate if $p = c$ and it is equal to the highest possible production if $p > c$.

Supply

- Supply curve is the portion of the $MC(y)$ above the minimum of the $AC(y)$ curve.

  Firms should shut down if price is below the minimum of $AC(y)$.

Figure, MWG 5.D.1 & 5.D.4

A fixed set up cost will result in a U shaped $AC(y)$

- long run cost is the lower envelope of the short run cost.

Figure illustration