

Econ 802

Lecture Notes on Chapter 5

Greg Dow

October 20, 2020

In These notes I will deal with four main topics:

- (a) The shapes of the average and marginal cost curves
- (b) The properties of the cost function in terms of prices
- (c) Shephard's Lemma
- (d) What returns to scale implies about long run average cost curves.

Average and marginal cost curves

If we know the cost function $c(w, y)$ we can write profit as $py - c(w, y)$. Suppose the input prices w are fixed and output price p is fixed. We want to choose the scalar output $y \geq 0$ to maximize profit.

The FOC is $p = \frac{\partial c(w, y)}{\partial y}$ or simply price = marginal cost.

The necessary SOC is $\frac{\partial^2 c(w, y)}{\partial y^2} \geq 0$ so MC is non-decreasing.

This indicates that the properties of the marginal cost curve are quite important in thinking about how a firm chooses output.

We will see later that the firm's output supply function $y(p)$ depends on the shape of the MC curve.

For the moment, I postpone this issue and focus on the features of the average and marginal costs.

In the next few pages, I hold the input price vector w constant and think about how $c(w, y)$ varies with y .

Go back to the definition of a cost function:

$$\begin{aligned} c(w, y) &\equiv \min_{x \geq 0} w \cdot x \text{ subject to } y = f(x) \\ &= w \cdot x(w, y) \end{aligned}$$

where $x(w, y)$ solves the minimization problem. This is the long run version of the cost function.

In the short run, we have some fixed inputs and some variable inputs. The fixed inputs might be buildings or equipment while the variable inputs might be labor and raw materials (the definitions can change depending on the real-world application).

Now we write the input vector as $x = (x_v, x_f)$ where x_v is the subvector of variable inputs and x_f is the subvector of fixed inputs. The corresponding prices are w_v and w_f .

The short run cost function is

$$\begin{aligned} c(w, y, x_f) &\equiv \min_{x_v \geq 0} w_v x_v + w_f x_f \\ &\quad \text{subject to } y = f(x_v, x_f) \\ &= \underbrace{w_v \cdot x_v(w_v, y, x_f)}_{\text{variable cost}} + \underbrace{w_f \cdot x_f}_{\text{fixed cost}} \end{aligned}$$

(3)

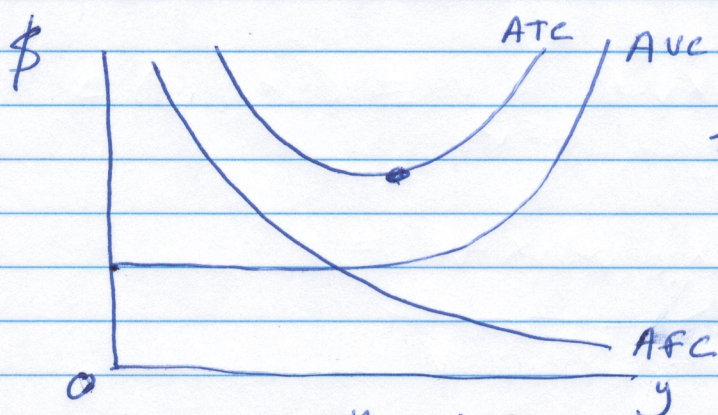
Notice that the optimal choices for the variable inputs $x_v(w_v, y, x_f)$ generally depend on the prices of these inputs, the output level and the levels of the fixed inputs (through the constraint). However they do not depend on the prices of the fixed inputs w_f because $w_f \cdot x_f$ is a constant and has no influence on the solution of the minimization problem.

In the long run, all inputs are variable, so the firm chooses both x_f and x_v to minimize cost. There is no fixed cost in the long run.

To get the short run average cost, we divide the equation at the bottom of p. 2 by the output y :

$$\underbrace{\frac{C(w, y, x_f)}{y}}_{\text{average total cost (or ATC)}} = \underbrace{\frac{w_v \cdot x_v(w_v, y, x_f)}{y}}_{\text{average variable cost (or AVC)}} + \underbrace{\frac{w_f \cdot x_f}{y}}_{\text{average fixed cost (or AFC)}}$$

We sometimes also refer to ATC as "short run average cost" (SAC). This often has a U shape:



Note: This means that ATC is close to AVC at large outputs.

The idea is that over some range AVC may be constant, but begins to rise as the firm approaches "capacity." Clearly AFC must always fall and approaches zero as $y \rightarrow \infty$.

4

Here is a basic fact about all cost curves (both short run and long run):

when AC is rising, MC is above it, and when AC is falling, MC is below it.

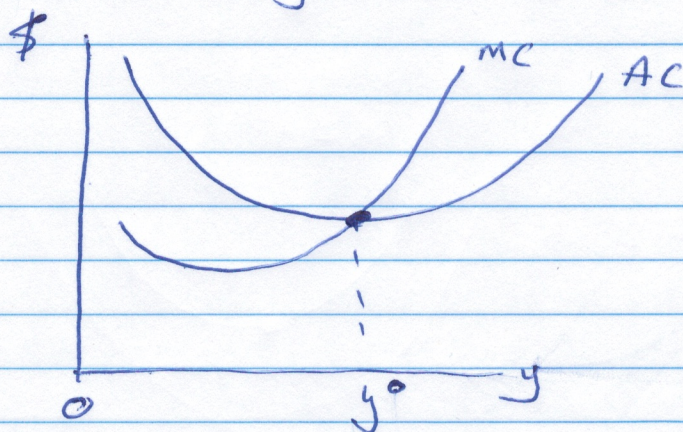
This is easy to prove with calculus.

Let $AC(y) = \frac{c(w, y)}{y}$ and differentiate with respect to y :

$$\begin{aligned}\frac{\partial AC(y)}{\partial y} &= \frac{\partial \left[\frac{c(w, y)}{y} \right]}{\partial y} = \frac{1}{y} \frac{\partial c(w, y)}{\partial y} - \frac{c(w, y)}{y^2} \\ &= \frac{1}{y} [MC(y) - AC(y)]\end{aligned}$$

This is positive ($AC(y)$ is rising at y) iff $MC(y) > AC(y)$ so the MC curve is above the AC curve at the output y ; and conversely when $AC(y)$ is falling at y .

Notice that if the output y^0 minimizes $AC(y)$ then $MC(y^0) = AC(y^0)$ so the MC curve passes through the AC curve at y^0 . Thus we get curves like this:



Also note that when AC is falling, MC can be either rising or falling. The only requirement is that MC must be below AC in this range.

(5)

The next issue is the relationship between SR average cost and LR average cost. You should think of cost min in the SR as the same as cost min in the LR except that there are more constraints in the SR due to the fixed inputs. Or, to put it another way, anything the firm does in the SR could also be done in the LR (the firm always has the option of choosing the same levels for the fixed inputs in the LR as it has in the SR).

Suppose the fixed inputs in the SR are $\bar{x}_f$. Then

$$c(w, y, \bar{x}_f) \geq c[w, y, x_f(w, y)] \equiv c(w, y)$$

where $x_f(w, y)$ are the optimal choices of the x_f vector. This just says that at (w, y) $\bar{x}_f$ might be the optimal inputs that would be chosen in the Long run, but they might not be. The firm never does worse by choosing the optimal levels $x_f(w, y)$ and it might do better, so short run cost is always at least as high as long run cost for the given (w, y) .

Dividing through by y , we get

$$\underbrace{\frac{c(w, y, \bar{x}_f)}{y}}_{\text{SAC}} \geq \underbrace{\frac{c[w, y, x_f(w, y)]}{y}}_{\text{LAC}} \equiv \frac{c(w, y)}{y}$$

Now consider a specific output $\bar{y}$ and let $\bar{x}_f = x_f(w, \bar{y})$ so $\bar{x}_f$ is what the firm would choose in the long run if it wanted to produce $\bar{y}$.

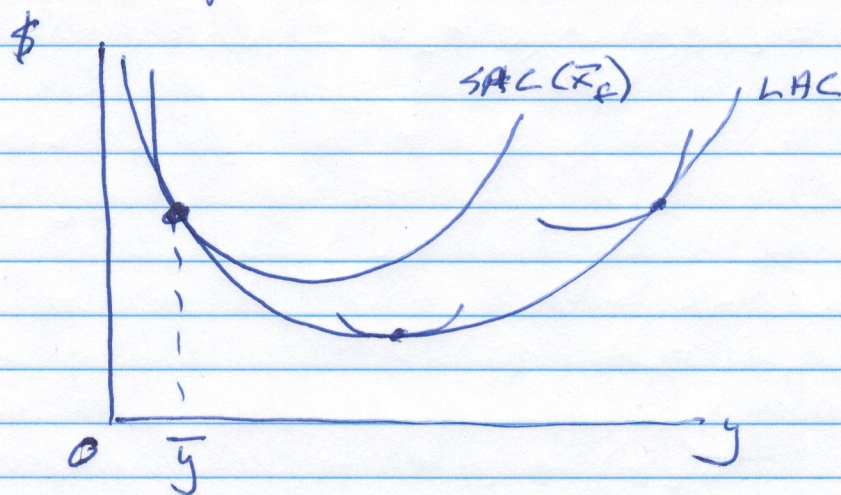
6

If the firm just happens to have $\bar{x}_F$ on hand in the SR, and it wants to produce $\bar{y}$, then it will have

$$\frac{c(w, \bar{y}, \bar{x}_F)}{\bar{y}} = \frac{c(w, \bar{y}, x_F(w, \bar{y}))}{\bar{y}} = \frac{c(w, \bar{y})}{\bar{y}}$$

In this special case, $SAC = LAC$ (long run average cost) at $\bar{y}$. However, for $y \neq \bar{y}$, the inputs $\bar{x}_F$ will not generally be optimal, and we typically get $SAC(y) > LAC(y)$.

Putting these results together, we have a tangency relationship between SAC and LAC :



If the firm has $\bar{x}_F$ in the SR, and it wants to produce $\bar{y}$, then $SAC(\bar{x}_F) = LAC$ at $\bar{y}$. But if the firm wants to produce some $y' \neq \bar{y}$, it will have $SAC(\bar{x}_F) > LAC$.

For different levels of the fixed inputs x_F , we generally get a tangency point at some other output level y (actually this is only true if x_F is a scalar; if it is a vector, there may not be any level of y for which x_F would be optimal in the LR).

Also observe that LAC does not pass through the minimum points of the SAC curves. If this happened when LAC was falling or rising we would contradict the fact that $LAC \leq SAC$ must always hold (try drawing some graphs and you will see the problem). It is only when LAC hits a minimum that LAC and the corresponding minimum of SAC coincide.

Properties of the Cost Function Involving Prices

So far we have held w fixed and considered what happens when we change y . Here we fix y and see what happens when we change w .

There are four properties to know about. These are closely related to the parallel properties for the profit function. I will state each one and discuss the proof.

① Non-decreasing in w : if $w' \geq w$ then $c(w', y) \geq c(w, y)$.

Proof: Let x' be optimal for w' . Then

$$c(w', y) = w'x' \geq wx' \geq c(w, y)$$

definition of the cost function using $w' \geq w$ the minimum cost at prices w is no greater than wx' (it may be smaller).

② Linear homogeneity in w : $c(\tau w, y) = \tau c(w, y)$ for all $\tau > 0$.

Fix the output y here.

Proof: Choose some arbitrary w^* and let x^* be cost min for prices w^* . This implies $w^*x^* \leq w^*x$ for all x such that $f(x) = y$. Therefore $\tau(w^*x^*) \leq (\tau w^*)x$ for all x such that $f(x) = y$.

(8)

This gives $(tw^*)x^* \leq (tw^*)x$ for all x such that $f(x) = y$.
 Therefore x^* minimizes cost at the prices (tw^*) .
 It follows that $c(tw, y) = (tw^*)x^* = tw^*x^* = tc(w, y)$.

Note: The key idea here is that if you multiply all the prices by the same positive constant, this has no effect on the optimal quantity choices. It is like deciding to measure the prices in euros instead of dollars. The feasible set is unchanged, the optimal solution is unchanged, and all you are doing is scaling the cost function up or down in the same proportion as the prices.

(3) Concavity in w :
$$c[tw + (1-t)w', y] \geq tc(w, y) + (1-t)c(w', y)$$
 for $0 \leq t \leq 1$.

This is just the definition of a concave function.

Proof: Let $w'' = tw + (1-t)w'$ and let x'' be the optimal inputs for the price vector w'' . Then

$$\begin{aligned} c[tw + (1-t)w', y] &= c(w'', y) = w''x'' \\ &= [tw + (1-t)w']x'' \end{aligned}$$

This follows from the fact that $w x'' \geq c(w, y)$ because $c(w, y)$ is the lowest possible cost at prices w and

$w' x'' \geq c(w', y)$ because $c(w', y)$ is the lowest possible cost at prices w' .

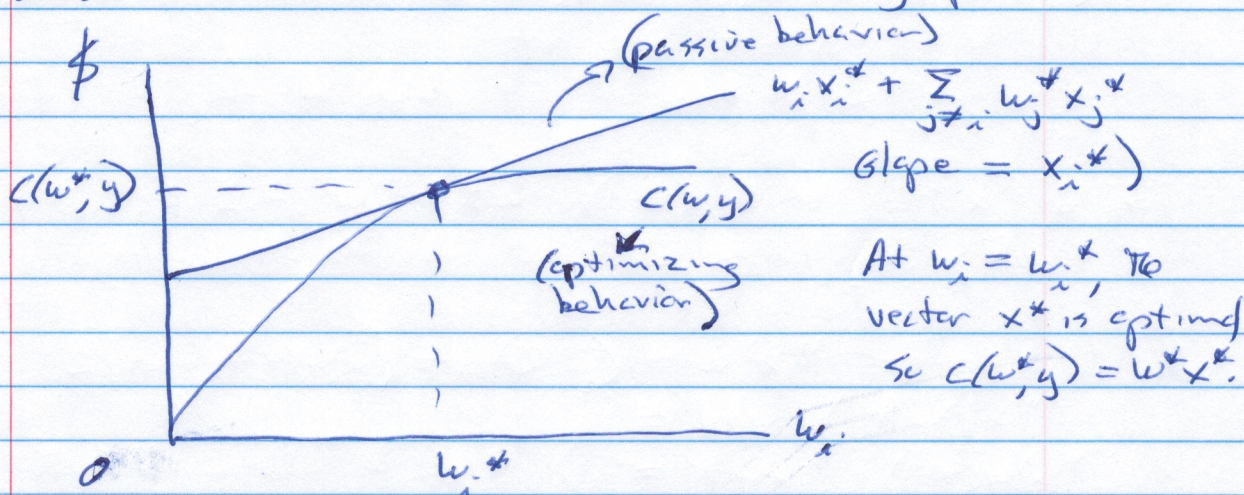
④ Continuity in w_i , for $w > 0$.

This follows from the Theorem of the Maximum
(see p. 506 in Varian)

Aside: here is some graphical intuition about the concavity result in (3). Let the minimum cost at w^* be $c(w^*, y)$. Suppose some price w_i changes, while all other prices $j \neq i$ remain at w_j^* . One thing the firm can always do in response to a price change is keep using its existing input vector x^* .

If it behaves in this passive way its expenditure on inputs will be $w_i x_i^* + \sum_{j \neq i} w_j^* x_j^*$

which is linear in w_i as shown in the graph:



However, when the price of input i changes (suppose it increases), the firm will generally want to change its input decision by substituting away from x_i as $w_i \uparrow$. Thus the minimum cost when it optimizes is typically less than the cost when it behaves passively, so the true cost function $c(w, y)$ is below the straight line. The same is true if w_i decreases. This gives the concave curvature for the true cost function shown in the graph.

Shephard's Lemma

You will recall that we differentiated the profit function to get the unconditional input demand functions using Hotelling's Lemma.

Here we do something very similar - we differentiate the cost function to get the conditional input demand functions using Shephard's Lemma.

The Lemma says: Let $x_i(w, y)$ be the conditional input demand for $i=1 \dots n$, given some fixed output y .

$$\text{Then } x_i(w, y) = \frac{\partial c(w, y)}{\partial w_i} \text{ for } i=1 \dots n$$

assuming the derivatives exist and $w > 0$.

Proof: Choose arbitrary prices $w^* > 0$ and let x^* be cost minimizing at the prices w^* . Define the function

$$g(w) \equiv c(w, y) - w x^* \leq 0$$

This is non-positive because for any given prices $w > 0$, $c(w, y)$ is the lowest possible cost of achieving y .

The vector x^* may or may not be optimal for w (in general it is not). But when the prices are w^* ,

we have $g(w^*) = c(w^*, y) - w^* x^* = 0$ because x^* is optimal at w^* by assumption. Therefore $g(w)$

reaches a maximum at w^* , and the FOC for a max must hold at this point. This implies

$$\text{for all } i=1 \dots n \mid \frac{\partial g(w^*)}{\partial w_i} = \frac{\partial c(w^*, y)}{\partial w_i} - x_i^* = 0. \text{ But } w^* \text{ and } x^* \text{ were arbitrary, and thus}$$

$$\frac{\partial c(w, y)}{\partial w_i} = x_i(w, y) \text{ for all } i=1 \dots n. \quad \text{Q.E.D.}$$

Shepherd's Lemma is helpful in proving comparative static results. It gives

$$\frac{\partial c(w, y)}{\partial w} = \left[\frac{\partial c(w, y)}{\partial w_1} \dots \frac{\partial c(w, y)}{\partial w_n} \right] = [x_1(w, y) \dots x_n(w, y)] \text{ so}$$

$$\frac{\partial^2 c(w, y)}{\partial w^2} = \frac{\partial x(w, y)}{\partial w} = \begin{bmatrix} \frac{\partial x_1}{\partial w_1} & \dots & \frac{\partial x_1}{\partial w_n} \\ \vdots & & \vdots \\ \frac{\partial x_n}{\partial w_1} & \dots & \frac{\partial x_n}{\partial w_n} \end{bmatrix}$$

We know from an earlier result that $c(w, y)$ is a concave function of its prices w . Therefore the Hessian of $c(w, y)$ is symmetric and negative semi-definite, which means that the substitution matrix is symmetric and negative semi-definite. Thus the cross-price effects are symmetric for the conditional input demands:

$$\frac{\partial x_i(w, y)}{\partial w_j} = \frac{\partial x_j(w, y)}{\partial w_i} \quad \text{for all } i, j$$

Also, the diagonal terms are non-positive, so

$$\frac{\partial x_i(w, y)}{\partial w_i} \leq 0 \quad (\text{conditional input demand curves cannot slope up})$$

Notice two things about this:

- ① you don't need to invert a matrix or use ~~Cramer's~~ Cramer's rule to get comparative static results.
- ② in some situations you may already know the cost function (maybe you estimated it from data on prices, output, and total expenditures) but you don't directly observe the input quantities. In a case like this, Shepherd's Lemma tells you everything you need to know about firm behavior.

Returns to Scale and The Long Run Average Cost Curve

This is the last topic in the Theory of the firm. I think it is useful in tying together a number of different ideas we have previously discussed.

The starting point is the following fact:

For any linearly homogeneous production function,
 $c(w, y) = y c(w, 1)$ where $c(w, 1)$ is the cost of one unit of output.

The proof of this fact is in Varian on p. 67.

What this says is that whenever a firm has constant returns (the production function is homogeneous of degree one), $AC = MC = c(w, 1)$ are constants that do not depend on output.

This should be intuitive: with CRS, it is always possible to scale the firm's activities up or down without changing cost per unit.

Another implication is that the conditional input demands $x_i(w, y)$ are linear in output y . To show this, use Shepherd's Lemma and differentiate $c(w, y) = y c(w, 1)$ with respect to the price w_i .

Again this should be intuitive: when the firm increases or decreases output, it scales the inputs up or down in the same proportion.

The first step in generalizing this result is to think about a Cobb-Douglas function where we could have CRS, IRS, or DRS.

$$\text{Let } c(w, y) = \min_{x_1, x_2} w_1 x_1 + w_2 x_2 \quad \text{subject to } A x_1^a x_2^b = y$$

Setting up a Lagrangean and using the FOC gives

$$c(w, y) = (a+b) \underbrace{\left[\frac{w_1^a w_2^b}{A a^a b^b} \right]^{\frac{1}{a+b}}}_{c(w, 1) \Rightarrow \text{cost when } y=1} y^{\frac{1}{a+b}}$$

You should check that you can carry out the algebraic steps leading to this result. Because the isoquants never touch the axes for a Cobb-Douglas function, you don't have to worry about Kuhn-Tucker multipliers. Also, because the production function is strictly quasi-concave (the input requirement sets are strictly convex) you don't have to worry about SOC; any point that satisfies the FOC solves the problem. In fact, the solution is unique, so the input demand functions are well-defined.

Now let's hold the price vector w fixed and stop writing it down for notational simplicity. We have

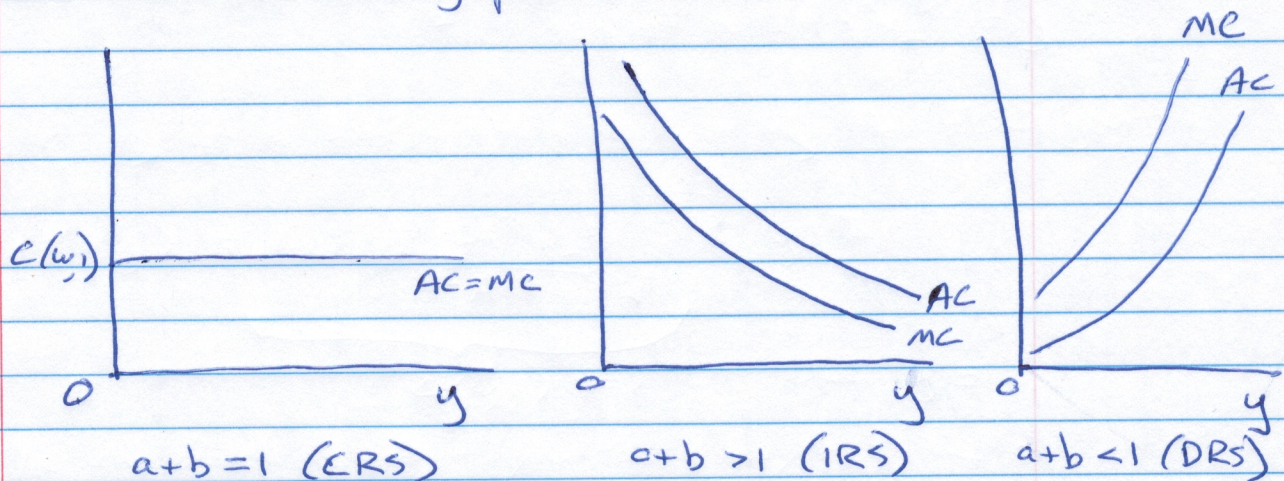
$$AC(y) = c(w, 1) y^{\frac{1}{a+b} - 1} = c(w, 1) y^{\frac{1-a-b}{a+b}}$$

$$MC(y) = \frac{c(w, 1)}{a+b} y^{\frac{1-a-b}{a+b}}$$

We already know that when $a+b=1$, we have CRS
 when $a+b > 1$ we have IRS
 when $a+b < 1$, we have DRS

(you should be able to prove that each of these statements is true.)

We can now draw graphs of the AC and MC curves:



Things to notice:

- ① These graphs are consistent with our earlier result that when AC is falling, MC is below it, and when AC is rising, MC is above it.
- ② There is a solution to the profit max problem in the CRS case iff $p \leq AC$. When $p < AC$, the unique solution is $y=0$. When $p = AC$, the firm has zero profit for every $y \geq 0$ and any output level is optimal. When $p > AC$, the firm can have unlimited profit by letting $y \rightarrow \infty$.
- ③ In the IRS case it is never possible to satisfy the necessary SOC for profit max because MC is always falling. Any solution to the FOC gives minimum profit, not maximum profit.
- ④ In the DRS case there is no difficulty with the profit max problem. A solution to the FOC will always satisfy the sufficient SOC because MC is always rising. There will be a unique solution for the optimal output and the resulting level of profit will be positive ($p = MC > AC$).

The next question is whether we can generalize these ideas further, beyond the Cobb-Douglas case.

In particular, could we get a U-shaped AC curve which is clearly impossible with a Cobb-Douglas function?

To do this, we will go back to the local elasticity with respect to scale, which was defined on p. 17 of Varian. I will use the following definition, which is equivalent to what Varian does:

$$e(x) = \sum_{i=1}^n \frac{\partial f(x)}{\partial x_i} \cdot x_i = \frac{\partial f(x)}{\partial x} \cdot x$$

Here is the result we want to prove:

Choose any output y^* and let x^* be the cost-minimizing way to produce y^* . Then

$$\begin{aligned} \frac{\partial AC}{\partial y} \Big|_{y^*} &> 0 \quad \text{if } e(x^*) < 1 \\ &= 0 \quad \text{if } e(x^*) = 1 \\ &< 0 \quad \text{if } e(x^*) > 1 \end{aligned}$$

This says that the long run average cost is rising at y^* if the elasticity with respect to scale has local DRS (at the point x^* the firm would use to produce y^*) and so on. There is a fixed price vector w in the background and x^* is cost-minimizing at these prices, but I will not write w down in what follows.

The proof involves several steps.

① we need to prove that $MC(y^*) = d(y^*)$, which says that marginal cost at an output y^* is always equal to the Lagrange multiplier in the cost min problem.

(16)

To show this consider the cost min problem for some arbitrary output y :

$$c(w, y) = \min w \cdot x \\ \text{subject to } f(x) = y$$

which gives the Lagrangian $L(x, d) = w \cdot x - d[f(x) - y]$

Using the FOC, we obtain $x(y)$ and $d(y)$. Since $x(y)$ is the cost-minimizing way to produce y and the constraint holds with equality, it must be true that

$$c(w, y) \equiv w \cdot x(y) - d(y)[f(x(y)) - y]$$

Now differentiate with respect to y :

$$\begin{aligned} \frac{\partial c(w, y)}{\partial y} &= w \cdot \frac{\partial x}{\partial y} - d(y) \frac{\partial f}{\partial x} \frac{\partial x}{\partial y} - \frac{\partial d}{\partial y} [f(x(y)) - y] + d(y) \\ &= \underbrace{\left[w - d \frac{\partial f}{\partial x} \right] \frac{\partial x}{\partial y}}_{= 0 \text{ from FOC}} - \underbrace{\frac{\partial d}{\partial y} [f(x(y)) - y]}_{= 0 \text{ from FOC}} + d(y) \end{aligned}$$

Therefore $\frac{\partial c(w, y)}{\partial y} = MC(y) = d(y)$ as we claimed.

This is an application of the envelope Theorem (see Chapter 3). Because we have already minimized cost, the FOC are satisfied, so we don't have to think about the indirect effects operating through $x(y)$ and $d(y)$ - the coefficients of $\frac{\partial x}{\partial y}$ and $\frac{\partial d}{\partial y}$ are automatically zero. All that matters is the direct effect where we differentiate with respect to y in the constraint, which gives $d(y)$.

This leads to a nice interpretation of the multiplier λ . If we had a maximization problem λ would give the rate at which the objective function would increase if we could relax the constraint (this is sometimes called a "shadow price"). Here we have a minimization problem so the multiplier gives the rate at which total cost could be reduced if we didn't have to produce as much output. This is equal to MC.

(2) The second step is just to write out the elasticity of output with respect to scale at the cost minimizing input vector x^* for the given output y^* :

$$e(x^*) = \frac{\sum_{i=1}^n \frac{\partial f(x^*)}{\partial x_i} \cdot x_i^*}{f(x^*)} = \frac{\frac{\partial f(x^*)}{\partial x} \cdot x^*}{f(x^*)}$$

(3) The third step is to carry out the differentiation of LR average cost with respect to output:

$$\frac{\partial AC(y)}{\partial y} = \frac{\partial \left[\frac{c(w, y)}{y} \right]}{\partial y} = \frac{1}{y} \left[\frac{\partial c(w, y)}{\partial y} - \frac{c(w, y)}{y} \right]$$

Note that we get the usual result: $AC \uparrow$ iff $MC > AC$
 $AC \downarrow$ iff $MC < AC$
 if AC is at a min, $AC = MC$.

(4) Finally substitute into the right hand side of (3) and evaluate everything at x^* and y^* :

$$\frac{\partial AC}{\partial y} \Big|_{y^*} = \frac{1}{y^*} \left[d(y^*) - \frac{w \cdot x^*}{y^*} \right]$$

where we used $\frac{\partial c(w, y^*)}{\partial y} = d(y^*)$ and $c(w, y^*) = w \cdot x^*$.

But from the FOC for cost min subject to $f(x) = y^*$, we have $w = d(y^*) \frac{\partial f(x^*)}{\partial x}$. Therefore:

$$\begin{aligned} \frac{\partial AC}{\partial y} \Big|_{y^*} &= \frac{d(y^*)}{y^*} \left[1 - \frac{\frac{\partial f(x^*)}{\partial x} \cdot x^*}{y^*} \right] \\ &= \frac{d(y^*)}{y^*} [1 - e(x^*)] \quad \text{where } d(y^*) > 0. \end{aligned}$$

This shows that $\frac{\partial AC}{\partial y} \Big|_{y^*} > 0$ if $e(x^*) < 1$ (local DRS)
 $= 0$ if $e(x^*) = 1$ (local CRS)
 < 0 if $e(x^*) > 1$ (local IRS)

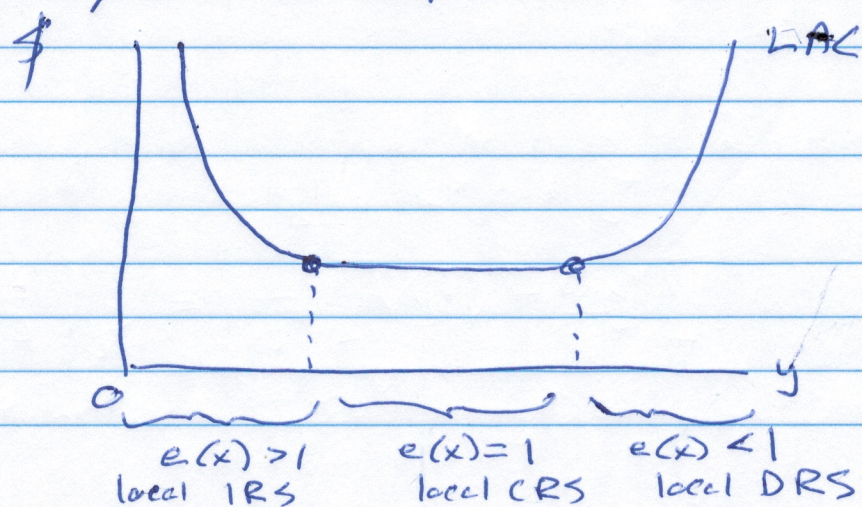
You can see that this is consistent with the previous results for the Cobb-Douglas case:

if $e(x) = 1$ globally, this holds at any cost min point, so AC is a constant.

if $e(x) < 1$ globally, we have rising AC at all outputs

if $e(x) > 1$ globally, we have falling AC at all outputs.

However now we have a generalization that allows the possibility of a U-shaped LAC curve like this:



People sometimes justify the assumption of local IRS at low output levels by using technological ideas (indivisibilities, engineering relationships between volumes and surface areas, etc) or organizational ideas (gains from specialization and division of labor).

Similarly people justify the assumption of local DRS at high output levels either technologically (there are some non-marketed inputs that are fixed even in the LR, like sun and rain) or organizationally (at large scales there are coordination or incentive issues, etc).

For the Cobb-Douglas function, you should check that $e(x) = a + b$ for all x , so the elasticity with respect to scale is a constant $\Rightarrow$ can't get U-shaped LAC.

We can extend this idea to any homogeneous production function of degree k , where $t^k f(x) = f(tx)$. If you differentiate with respect to t , this gives

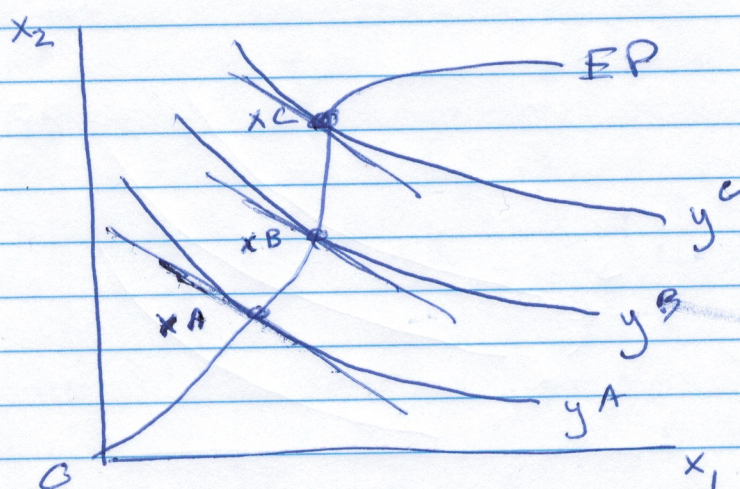
$$k t^{k-1} f(x) = \sum_{i=1}^n \frac{\partial f(tx)}{\partial x_i} \cdot x_i$$

Now set $t=1$ which gives $k f(x) = \sum_{i=1}^n \frac{\partial f(x)}{\partial x_i} \cdot x_i$
 or $k = e(x) = \frac{\sum_{i=1}^n \frac{\partial f(x)}{\partial x_i} x_i}{f(x)}$

So again the elasticity with respect to scale is a constant and we cannot have a U-shaped LAC.

One more concept: The expansion path (EP) for a fixed vector of input prices w is the set of all x vectors that are cost minimizing for some output $y > 0$.

Graphically the expansion path looks like this:



where y^A, y^B ,
and y^C are
isoquants,
 x^A, x^B, x^C
are cost-min
points, and the
slopes of the
isocost lines
are parallel
and determined by w .

For a fixed w , the only points that are relevant in determining whether LAC is rising or falling are the points along EP, because only these points will ever be cost-minimizing. No other x vectors will be used.

If the vector of input prices w changes then the slopes of the isocost lines change, and the entire EP will shift in the graph.

In general we don't know much about the shape of EP. However, there are important special cases.

For example in the Cobb-Douglas case, from FOC we have

$$\left. \begin{aligned} w_1 &= df_1 = a x_1^{a-1} x_2^b \\ w_2 &= df_2 = b x_1^a x_2^{b-1} \end{aligned} \right\} \Rightarrow \frac{w_1}{w_2} = \frac{a}{b} \frac{x_2}{x_1}$$

so the ratio $\frac{x_2}{x_1}$ is constant or $\frac{x_2}{x_1} = \frac{b}{a} \frac{w_1}{w_2}$
and the expansion path is a ray from the origin.

This can be generalized to any homogeneous function of degree k (which must have partial derivatives of degree $k-1$).

In this case the FOC for cost min gives

$$\frac{w_1}{w_2} = \frac{\frac{\partial f(x)}{\partial x_1}}{\frac{\partial f}{\partial x_2}} = \frac{t^{k-1} \frac{\partial f(x)}{\partial x_1}}{t^{k-1} \frac{\partial f(x)}{\partial x_2}} = \frac{\frac{\partial f(tx)}{\partial x_1}}{\frac{\partial f(tx)}{\partial x_2}}$$

This implies that if x is an expansion path, then so is the point tx for $t > 0$.

Again, this implies that EP is a ray from the origin.

These results can be extended slightly further: The EP is also a ray from the origin if the production function is homothetic (see Varian's discussion on pp. 18-19). I won't prove this, but you should be able to see what is going on from the book.

That's all for Chapter 5!