Adv. PSE I: Public Policy in Open Economies
Math Primer

Severin Weingarten

August 28, 2014

Contents

0 About This Document

This primer is designed as an introduction to the mathematical techniques that will be used throughout the MA course ‘Adv. PSE I’ and other courses offered by the chair of public finance. The selection of techniques is not meant to be complete, it is meant to fit into roughly two hours of teaching or careful reading. A more comprehensive introduction to mathematics for economists is taught by Oliver Kirchkamp as the first part of ‘MW26.1 Approaches to Economic Science’.

Chiang (1984) is a very good (English) book to look up specific math problems. It is available in the library: `WIR:GC:100:C532:(4):2005`.

1 Functions with a Single Variable and Their Derivatives

1.1 Functions

A function describes how measures of one or more inputs translate into a single measure of output. Throughout this primer we will use the example of a production function.

Imagine us running a factory that produces fresh water from sea water.[1] Since plenty of sea water is available, we ignore it for the moment and assume that our output of fresh water Y only depends on the number of machines (i.e. the size of the factory or capital) K available to perform the transformation.

$$Y = Y(K) = f(K) = 5 \cdot K \tag{1}$$

Equation (1) is an example of a production function that could apply in this context. Y, $Y(K)$, and $f(K)$ are different names that are often used for this sort of function, $5 \cdot K$ is the function's specific functional form. It contains information about our technology of production: For every unit of capital (e.g. for every machine) we are able to produce 5 units of fresh water Y (per day). For example, $K = 8$ units of capital would allow us to produce $Y = 5 \cdot 8 = 40$ units of fresh water.

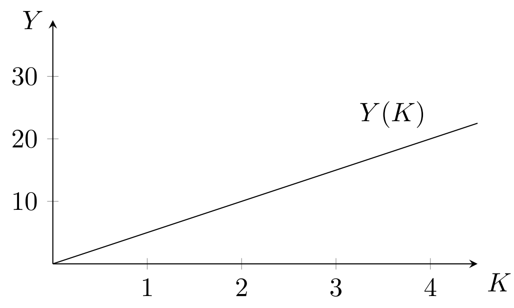

Figure 1: A linear production function with a single input K

We plot the function in figure 1 to get a better understanding of how it works. Admittedly, this is a pretty trivial function, but we will make things more complicated soon.

1.2 Derivatives

Our production function tells us something about the relationship between K and Y, but sometimes we are more interested in the relationship between

[1] The example may seem slightly unconventional at first, but we will get back to the choice of the example later.

a change in K and a change in Y. For example when we are considering an increase in capital, we care about the additional output that this increase will generate. The first-order derivative of a function provides us with this information.

If we know the functional form of our production function, and in our case we do, we can use the rules of differentiation to calculate the derivative. Tables 2 and 3 in appendix A contain those rules. Appendix B provides you with some exercises that you can use to practice the application of the rules if you want to brush up on your skills.

$$\frac{\partial Y}{\partial K} = Y_K = Y'(K) = f'(K) = 5 \tag{2}$$

Equation (2) is the first-order derivative of our production function with respect to K. Again $\frac{\partial Y}{\partial K}$, Y_K, $Y'(K)$, and $f'(K)$ are different names for the derivative (we will mostly be using the first and the second one), while 5 is the specific functional form of our derivative.

Here, the first-order derivative does not depend on K, i.e. there is no K in 5. Regardless of how much K we are already using, adding one additional unit of capital K to our inputs will result in 5 additional units of fresh water Y being produced.

The second derivative is nothing more than the first-order derivative of the first-order derivative. It tells us how the relationship between a change in K and a change in Y changes with K. This may seem slightly confusing at first, but continue to read. In the next example, things will become clearer.

$$\frac{\partial^2 Y}{\partial K^2} = Y_{KK} = Y''(K) = f''(K) = 0 \tag{3}$$

Equation (3) is the second-order derivative of our production function. Again, $\frac{\partial^2 Y}{\partial K^2}$, Y_{KK}, $Y''(K)$, and $f''(K)$ are different ways of referring to the second derivative of Y with respect to K and 0 is the specific functional form of the derivative in the case of our production function.

The value 0 of the second derivative implies that, in our case, the level of K does not have an effect on the relationship between changes in K and Y. Actually, we already discovered this characteristic of our production above: an increase in K by one unit will always translate into an increase in Y by five units. The word 'always' precisely refers to the fact that this assertion does not depend on the level of K.

We will now look at a production function with a different functional form. This will also allow us to revisit the concept of derivatives.

$$Y(K) = K^{0.5} \tag{4}$$

Above, we used a linear production function. Instead, we will now use the function given by equation (4). From now on, we will also stick to the following names for our function and its derivatives: $Y(K)$ is our production function, Y_K is its first-order derivative (with respect to K) and Y_{KK} is its second derivative (with respect to K).

Figure 2 plots the new production function. It rises quickly for low levels of K and more slowly for higher levels of K. This assumption seems more realistic if the plant is run by a fixed number of workers.[2] As the number of machines increases, technicians will have less time to attend to each individual machine. This will decrease the productivity of each individual machine, because it will take longer, until somebody notices machine failures and has time to fix them. We will still be able to produce more fresh water if we add additional machines, but the additional output Y after adding a machine will be lower if we already

[2] The assumption of a fixed number of workers is not explicitly visible in the production function. We will take a closer look at it in section 2.

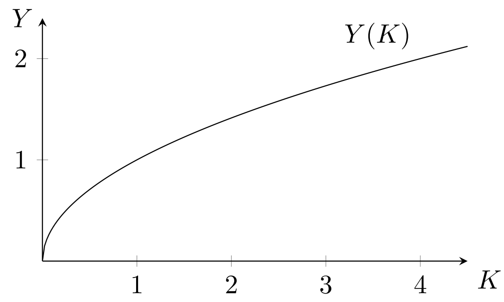

Figure 2: A concave production function with a single input K

have a lot of machines. These characteristics of our technology are reflected in the derivatives of our production function.

$$Y_K = 0.5\,K^{-0.5} \tag{5}$$

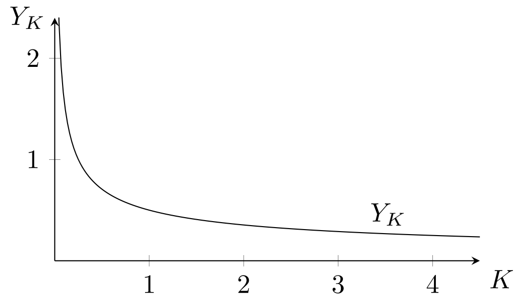

Figure 3: The first-order derivative of a concave production function

Equation (5) gives the first-order derivative of the new production function and figure 3 plots it. The derivative is positive but declines as more capital is used. The output contribution of additional machines decreases with the number of machines in use.

$$Y_{KK} = -0.5\,K^{-1.5} \tag{6}$$

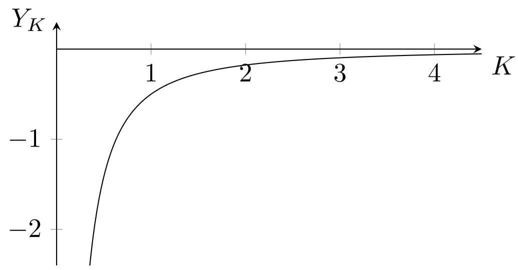

Figure 4: The second derivative of a concave production function

Equation (6) and figure 4 repeat the same exercise for the second derivative.

Above, we interpreted the value of the second derivative as the effect of changes in K on the relationship between changes in K and Y. We can now understand this explanation by looking at our example.

Our second derivative is negative for all values of K, i.e. an increase in K will reduce the amount of Y that we can gain by adding even more K. However, this negative effect approaches zero for large values of K. As the benefit of adding K (measured by the first-order derivative) becomes smaller, so does the negative effect of adding K on that benefit (measured by the second derivative).

In case you are confused, table 1 summarizes the interpretation of the value of a function and its derivatives using the example of a production function.

Value of ...	Interpretation
$Y(K)$	Amount of output produced with inputs K
Y_K	Benefit of adding one more unit of K
Y_{KK}	Change of benefit of adding one more unit of K

Table 1: Interpretation of a production function and its derivatives

1.3 General Assumptions vs Specific Functional Forms

We have already seen two different functional forms of production functions: $Y(K) = 5 \cdot K$ and $Y(K) = K^{0.5}$. Specific functional forms are nice, because they allow us to easily plot functions and quickly get an impression of their shape. They also allow us to calculate derivatives if we are interested in the behavior of the function. However, most of the time, we will not use specific functional forms, because they introduce more (implicit) assumptions about how *exactly* a function looks than we are willing to make.

Instead of something like $Y(K) = K^{0.5}$ we will simply state that there is a function $Y(K)$ with $Y_K > 0$ and $Y_{KK} < 0$. These assumptions are much less 'heavy', but carry roughly the same meaning: Output increases in K, but as K gets larger the same increase in K leads to smaller and smaller increases in Y. Economists use this particular set of assumptions quite often and call it 'positive and decreasing returns to scale'.[3]

2 Assumptions

In economic models we often make assumptions which seem overly simplistic or extremely unbalanced. That can be a bad thing, but often it is not.

Not all assumptions are made consciously. Some mathematical properties can be interpreted in different ways.

Let's take a look at some of the assumptions that we have made above:

1. Sea water doesn't appear in the production function. ⇒ Fits our picture of the real world. There is so much of it, that we probably don't have to care.
2. There is only one type of machine. All machines work in the same way. ⇒ Doesn't fit, but not too far from the truth. It is probably ok, to assume some average technology across all machines.
3. Machines are the only input (factor of production). ⇒ Definitely not true. Somebody has to operate those machines.
    - Can be interpreted as there being a single operator—or any fixed number of operators. ⇒ Probably not so bad a fit after all.
4. Machines can't break. ⇒ Probably fine in the short-run. Completely unrealistic in the long-run.
    - K can be interpreted as a flow measure, as operating costs instead of a stock. ⇒ Probably fine, but does not really fit into the picture of the real world that we have drawn above.
5. No external events: floods, sabotages, strikes ... ⇒ Completely unrealistic in the long-run.

[3] Another set of assumptions that is often used are the Inada conditions (Inada 1963). Intuitively, they carry the same meaning as 'positive and decreasing returns to scale' but include additional assumptions that ensure that certain problems can be solved in a way that satisfies mathematicians and economists at the same time.

- These things happen. However, they don't happen very often and here we are only interested in modeling how things work 'on regular days'.

Take away:

1. Some assumptions are completely unrealistic, but it doesn't matter, because we are focusing on something else. (As long at there is no interaction between this subject and another.)
2. Some assumptions are made implicitly and can be interpreted in different ways.
3. Some assumptions are unrealistic and have to be replaced by better ones.

3 Functions with Many Variables

We will now extend our previous example of a production function by explicitly adding labor L as an input in equation (7). We end up with a function that has two inputs.

$$Y(K,L) = K^{0.5}L^{0.4} \tag{7}$$

Since this is a three-dimensional function (two inputs plus one output) we will not attempt to plot it. However, we can plot the relationship between each input and the output under the assumption that the other input remains fixed at some level $\bar{K}$ or $\bar{L}$. This assumption is often called the ceteris-paribus (c.p.) assumption. Figure 5 shows the two resulting diagrams.

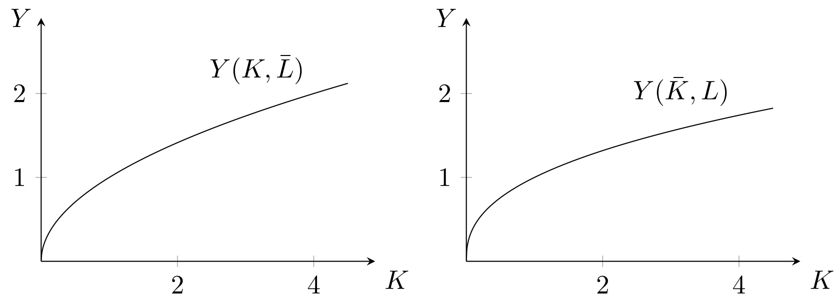

Figure 5: Plots of a production function with two inputs, one input fixed

Derivatives of functions with more than one input are also calculated under the ceteris-paribus assumption.[4] Our new production function has two first-order derivatives: each of them tells us something about the relationship between output and one of the inputs, given that the other input remains fixed.

$$Y_K = 0.5K^{-0.5}L^{0.4} \tag{8}$$
$$Y_L = 0.4K^{0.5}L^{-0.6} \tag{9}$$

The derivatives are given by equations (8) and (9). Both derivatives are positive: adding more of one production factor, while keeping the other fixed, increases output.

$$Y_{KK} = -0.25K^{-1.5}L^{0.4} \tag{10}$$
$$Y_{LL} = -0.16K^{0.5}L^{-1.6} \tag{11}$$

We can also calculate two second-order derivatives (sometimes called the 'pure' second-order derivatives) of our production function; we simply have to take the derivative of each first-order derivative with respect to the same input as before.

[4] If you don't like this assumption, wait for section 6 where we will learn about the total differential. It will allow us to look at the change of multiple variables at the same time.

The second-order derivatives are given by equations (10) and (11). Both are negative: increasing the amount of an input reduces the output gains that are derived from a one-unit increase in the same input.

However, a third second-order derivative (the 'cross' derivative) is still missing. We can calculate it by picking either first-order derivative and differentiating it with respect to the input with respect to which it has not been differentiated before.[5]

$$Y_{KL} = Y_{LK} = 0.2K^{-0.5}L^{-0.6} \tag{12}$$

Equation (12) gives the cross derivative of our production function. It is positive: increasing the amount of one input increases the input gains that can be obtained from adding an additional unit of the other input. This is an important consequence of the functional form that we have picked for our production function and it is also an assumption that we will often make when we do not use specific functional forms: our inputs complement each other. It is more productive to have a balanced mix of the two than a lot of one, but very little of the other.

If you think about it, this assumption makes a lot of sense in our example. In the long run, machines will fail without engineers, but engineers will have a hard time transforming sea water into fresh water without any machines. A mix of both will result in more output.

4 Homogeneity and the Euler Equation

Up to this point, we have not talked about the cost of our inputs K and L and actually, we will delay most of our thoughts about that until section 5. However, assume for a second that capital is costly and that we have to pay for it. Now take another look at figure 2, which plots our earlier production function where capital K was the only input. You can immediately tell that there must be an output level that we cannot reach. The more we produce, the more expensive becomes the next increase in outputs. At some point, producing an additional unit of output will be so expensive, in terms of capital, that we will not be interested in producing it.

This is basically the same result that we got from the interpretation of the second-order derivative of the production function in equation (6).

It would be good to know, whether our new production function, which has two inputs K and L, suffers from a similar problem. Do we get less bang for the buck, as we produce more and more? To make things a little more concrete: "Does adding 10% more capital and 10% more labor yield more than, less than, or exactly 10% more output?" Unfortunately, things are a little complicated with our new production function, because there is no single second-order derivative for changes in K and L that would provide us with this kind of information.

Actually, not all functions allow us to answer this question unambiguously. Those who do are called homogeneous functions and, fortunately, our new production function is homogeneous.

$$\begin{aligned} Y(\delta K, \delta L) &= (\delta K)^{0.5} \cdot (\delta L)^{0.4} \\ &= \delta^{0.5} \cdot K^{0.5} \cdot \delta^{0.4} \cdot L^{0.4} \\ &= \delta^{0.9} \cdot K^{0.5} \cdot L^{0.4} \\ &= \delta^{0.9} \cdot Y(K, L) \end{aligned} \tag{13}$$

Mathematically, a function is homogeneous if we can multiply each input with some constant, say δ, and rearrange the function until we end up with the

[5] It does not matter which of the two first-order derivatives we pick. We will always end up with the same result. This mathematical law is so famous that it has two different names: depending on the context it is either called the 'Schwarz integrability condition' or 'Young's theorem'.

product of the constant and the original function. In (13) we perform this exercise for our production function. We are able to extract δ, which proves that our production function is indeed homogeneous.

But what about our original question? We can answer it, by looking at the exponent of δ in the last line of (13). It is called the degree of homogeneity. Sometimes it is also referred to as the λ of a homogeneous function. If $\lambda < 1$, the output contribution of an increase in all inputs declines as we produce more and more output. This was the case for our old production function $Y(K) = K^{0.5}$ and it is also the case for our new production function $Y(K, L) = K^{0.5}L^{0.4}$, which has a degree of homogeneity of 0.9. If $\lambda > 1$ the opposite would be true: the output contribution of an increase in all inputs would increase as we produce more. Finally, a function with $\lambda = 1$ is called linearly homogeneous and a proportional increase in inputs always yields an output gain of the same proportion, independently of how much is already being produced.

As an aside: The degree of homogeneity tells us something important about the nature of our technology. $\lambda < 1$ implies decreasing, $\lambda = 1$ constant, and $\lambda > 1$ increasing returns to scale. Non-increasing returns to scale are considered a precondition for the survival of a competitive market. Increasing returns to scale lead to one of the traditional forms of market failure: a natural monopoly. One big company can squeeze all others out of the market, because it can produce goods more cheaply than any smaller competitor. One solution is for the government to step in and produce the good on its own. Classic examples of technologies with increasing returns to scale are railroads and phone lines.

To understand the economic importance of homogeneous production functions, let's briefly consider the economic surroundings of our little fresh water factory. We are using capital and employing labor, but maybe we do not own all of the capital ourselves and even if we did, why shouldn't we use it for something else than fresh water production. We certainly do not own the workers who work in our factory. Why don't they go and work somewhere else? The answer is simple: we earn a return on our capital and we pay the workers, and if we did not, we would have to close shop.

One of the results that you might remember from a course on microeconomics is that, under perfect competition, all factors of production are payed according to their marginal productivity.

The problem is, that non-homogeneous production functions do not easily yield this fundamental result. The Euler theorem is a mathematical law, that can help us to understand this problem. It tells us that, as long as the degree of homogeneity is smaller than or equal to one, we will always produce enough output to pay our inputs their marginal productivities.

$$\lambda Y(K, L) = Y_K \cdot K + Y_L \cdot L \tag{14}$$

Equation (14) states the Euler theorem for a production function with two inputs K and L. The right-hand side sums up the marginal productivities of all inputs, each multiplied with the employed quantity of the input. This is the total payment that we will have to make to meet the demand of capital owners and workers. According to the theorem this payment equals $\lambda Y(K, L)$. Since we produce $Y(K, L)$ and the payment is only the share λ of that, we are fine as long as $\lambda \leq 1$.

The calculation in (15) demonstrates that the Euler theorem holds for the production function $Y(K, L) = K^{0.5}L^{0.4}$ that we have used in the previous sections.

$$\begin{aligned} Y_K \cdot K + Y_L \cdot L &= \\ \left(0.5 \cdot K^{-0.5}L^{0.4}\right) \cdot K + \left(0.4 \cdot K^{0.5} \cdot L^{-0.6}\right) \cdot L &= \\ 0.5 \cdot K^{0.5}L^{0.4} + 0.4 \cdot K^{0.5}L^{0.4} &= \\ 0.9 \cdot K^{0.5}L^{0.4} &= 0.9 \cdot Y(K, L) \end{aligned} \tag{15}$$

Because the Euler theorem gives us this nice result and because it only applies to homogeneous functions, we will often assume that our functions are homogeneous.

In fact, we will often assume linearly homogeneous production functions, i.e. functions with $\lambda = 1$. First, because they cause economic profits (i.e. what is left to us as the owner of the factory after we have payed off our inputs) to be zero right after profit maximization. If we include price adjustments on the product market into our model, the same is true for functions that have $\lambda < 1$, but sometimes we do not want to model the product market at all and then it is convenient to arrive at zero profits right away. Second, linearly homogeneous functions have a number of other nice mathematical properties that make it easy to work with them. Chiang (1984, p. 411–414) has many more details on this topic.

5 Optimization

By now, we have learned quite a bit about derivatives, but we haven't really used them for anything. Derivatives are so important to economists, because they are the basis of optimization calculus.

Optimization uses the most basic interpretation of a derivative: the value of the derivative is the steepness of the original function.[6] Since an optimum is either a minimum or a maximum of the original function, it is also a place, where the steepness equals zero. As an analogy, imagine yourself climbing a mountain (the function): a point from where you can climb up further (first-order derivative different from zero) cannot be the peak.

$$\begin{aligned} Y_K &\overset{!}{=} 0 \\ Y_L &\overset{!}{=} 0 \end{aligned} \tag{16}$$

To find the optimum of a function, we simply ask: what is a combination of inputs, for which all of the function's first-order derivatives are zero? The conditions in (16) summarize this question as mathematical expressions for the case of two inputs K and L.

Unfortunately, this method only gives us candidates for optima. In order to make sure, that such a candidate is really an optimum and an optimum of the desired kind (maximum or minimum) we also have to check that the second-order derivatives conform to the conditions in (17).

$$\begin{aligned} Y_{KK} > 0 \quad &\text{and} \quad Y_{LL} > 0 \quad \text{(Minimum)} \\ Y_{KK} < 0 \quad &\text{and} \quad Y_{LL} < 0 \quad \text{(Maximum)} \\ Y_{KL} > 0 & \end{aligned} \tag{17}$$

These conditions have intuitive explanations as well.[7] However, since we will rarely use second-order conditions in the course, we will skip them here and continue with an example instead.

We will now apply optimization to our example of producing sea water from fresh water. Of course, there is no point in optimizing output, if inputs are costless. We could simply produce as much as we wanted or needed. There would not be any limits. To make things more realistic, we will assume that

[6] Or in the case of more than one input, the steepness 'when looking into the direction of that input'.

[7] Suffice it to say that we have to exclude the possibility of having reached a plateau instead of the peak of our imaginary mountain. First-order derivatives are short-sighted in the sense that they will not be able to see that we can climb up further if the function is flat over too long a distance.

inputs have to be paid for: each unit of capital costs r (**r**eal interest rate[8]) and each unit of labor costs w (**w**age). We will also assume that Y is already measured in terms of its market value (or, to the same effect, that we can sell each unit of Y at price 1).

$$\max_{K,L} \quad \pi(K,L) = Y(K,L) - C(K,L) \tag{18}$$

$$\max_{K,L} \quad \pi(K,L) = Y(K,L) - r \cdot K - w \cdot L \tag{19}$$

Expressions (18) and (19) state our optimization problem in a formal way. $\max_{K,L}$ indicates that we are looking for a maximum and that our choice variables are K and L. We can adjust only K and L to reach the optimum. All other variables are called parameters. We want to find the optimal values of our choice variables given our parameters r and w.

As required by (16), we take derivatives and set them to zero. This yields equations (20) and (21). We call these equations first-order conditions (FOC). We will use this term all the time during the course.

$$\frac{\partial \pi}{\partial K} = Y_K - r \overset{!}{=} 0 \tag{20}$$

$$\frac{\partial \pi}{\partial L} = Y_L - w \overset{!}{=} 0 \tag{21}$$

Mathematically there can only be an optimum, where all FOCs are fulfilled. Often, we will combine two or more FOCs to arrive at yet another equation that will hold in the optimum.

$$\frac{Y_K}{r} = \frac{Y_L}{w} \tag{22}$$

Since there are only two choice variables in our example, we only get two FOCs. Combining them is quite straightforward. We get equation (22), which has an important economic interpretation that you may remember from courses in microeconomics or economic policy: in the optimum, all factors have the same marginal productivity per unit of currency.

This gives us a rough guideline for optimizing our production: We ask ourselves how production would change if we added more capital or labor. If we divide our estimates of these changes by the respective input prices, we get the terms on both sides of equation (22). If one of the terms is higher, we will use a little more of that factor and pay for it with a reduction in the use of the other factor. We know from the second-order derivatives that such adjustments will eventually equalize the two terms.[9]

If this rough guideline isn't enough for us, we can also calculate the actual optimal ratio of our inputs. This does, however, require that we agree on a specific functional form of $Y(K,L)$ and substitute its derivatives into equation (22). If we stick to the production function $Y(K,L) = K^{0.5}L^{0.4}$ from the previous section—the derivatives of which we have already calculated above—we get equation (23).

$$\begin{aligned}
\frac{0.6K^{-0.4}L^{0.4}}{r} &= \frac{0.4K^{0.6}L^{-0.6}}{w} \\
\frac{0.6K^{-0.4}L^{0.4}}{0.4K^{0.6}L^{-0.6}} &= \frac{r}{w} \\
\frac{3}{2}\frac{L}{K} &= \frac{r}{w} \\
\frac{L}{K} &= \frac{2}{3}\frac{r}{w} \qquad (23)
\end{aligned}$$

[8] The term 'real' bears no specific meaning in this context. In models that include inflation, economists often differentiate between a real interest rate r and a nominal interest rate i. Since there is no inflation in our model, we have to pick one of the letters and it is common practice to use r.

[9] If you do not understand this last bit, ask yourself what the second-order derivatives tell us about changes in Y_K and Y_L as we adjust inputs.

Using a specific functional form, the output ratio can be expressed in terms of the price ratio $\frac{r}{w}$ and the $\frac{2}{3}$ which accounts for the fact that the marginal productivities of the two factors are not the same. If the prices of the two inputs were the same, we would still use more capital than labor, because it has the higher marginal productivity in our production function.

Summarizing the results of this section, we have seen that fairly simple optimization techniques can be used to determine the input ratio that leads to maximum output. We have plugged in a specific functional form for our production function to illustrate things a bit. But for most of the course, we will stick to fairly abstract production functions with very few additional assumptions such as linear homogeneity.

6 The Total Derivative and What It Is Used For

In this section, we will look at a very simple example of a mathematical model that will be used in the course. The model can be analyzed using our final mathematical technique: the total differential.

We will stick to the example of fresh water production, but we will now take the perspective of a government that wants to generate tax revenues by levying a tax on capital. Our goal for this section is to find a rule that tells us, how firms will change their decision about how much capital and labor to use (the main result of the previous section) if we introduce a tax (or increase its rate).

First, we have to introduce the capital tax into the maximization calculus (19) of the firm. That is easily done: taxes represent additional costs. If, for example, a firm has to pay a return of $r = 10\%$ to the owners of the capital it uses and an additional $\tau = 2\%$ in taxes, we would simply add up the rate of return r and the tax rate τ and multiply them with the amount of capital used. Our new maximization calculus is then given by (24).

$$\max_{K,L} \quad \pi(K,L) = Y(K,L) - (r+\tau)\cdot K - w\cdot L \tag{24}$$

Repeating the calculation from the previous section we get the first-order conditions (25) and (26). Combining the two conditions gives us equation (27).

$$\frac{\partial \pi}{\partial K} = Y_K - r - \tau \overset{!}{=} 0 \tag{25}$$

$$\frac{\partial \pi}{\partial L} = Y_L - w \overset{!}{=} 0 \tag{26}$$

$$\frac{Y_K}{r+\tau} = \frac{Y_L}{w}$$

$$\frac{Y_K}{Y_L} = \frac{r+\tau}{w} \tag{27}$$

This result tells us that increasing the tax rate τ will cause firms to use less capital. How can we see that? An increase in τ will cause the right-hand side of equation (27) to increase. Firms cannot change anything about the right-hand side, because it only contains parameters. The left-hand side, however, contains Y_K and Y_L, both of which are derivatives of $Y(K,L)$ and themselves functions of K (and L). If only we could tell, how K had to change, to increase the term on the right. Well, we can: Above, we noted that we are willing to assume $Y_{KK} < 0$ and $Y_{KL} > 0$. We will use these assumptions now. According to the former an increase in capital will decrease Y_K and according to the latter it will increase Y_L. Since that is the opposite of what we want, we know that firms will use a decrease in capital to adjust the right-hand side of equation (27) if the government increases the tax rate τ. Voila.

The total differential is a technique that allows us to translate (the admittedly fairly exhausting) reasoning above into a very short calculation. It will also tell us more precisely how strong the effect of an increase in the tax rate on capital usage is.

First, some theory: A single equation can have many total differentials; one for each combination of variables that are allowed to change. Before we calculate a total differential, we have to choose variables that will be allowed to change. The total differential will tell us, how changes in the chosen variables fit together so that the equation does still hold after the changes.

For example, we have seen that if we change τ in equation (27), some other variable has to be adjusted to compensate for that. We decided to look at how an adjustment in K would be able to compensate for an adjustment in τ. Our chosen variables are K and τ.

Calculating a total differential can sometimes be tedious, but the individual steps are fairly easy. On both sides of the equation, take the derivatives of each side of the equation, once with respect to each variable that is allowed to change, add to each derivative an abbreviation for the change in that variable (e.g. K or τ) and sum up the resulting terms. To make things slightly easier we will first remove the fractions by rearranging from (27) to (28).

$$Y_K \cdot w = Y_L \cdot (r + \tau) \tag{28}$$

$$\underbrace{Y_{KK} \cdot w \,\mathrm{d}K}_{\frac{\partial \mathrm{LHS}}{\partial K}} + \underbrace{0 \,\mathrm{d}\tau}_{\frac{\partial \mathrm{LHS}}{\partial \tau}} = \underbrace{Y_{LK} \cdot (r + \tau) \,\mathrm{d}K}_{\frac{\partial \mathrm{RHS}}{\partial K}} + \underbrace{Y_L \,\mathrm{d}\tau}_{\frac{\partial \mathrm{RHS}}{\partial \tau}} \tag{29}$$

$$\big(Y_{KK} \cdot w - Y_{LK} \cdot (r + \tau)\big) \,\mathrm{d}K = Y_L \,\mathrm{d}\tau$$

$$\tfrac{\mathrm{d}K}{\mathrm{d}\tau} = \frac{\overbrace{Y_L}^{>0}}{\underbrace{Y_{KK}}_{<0} \cdot w - \underbrace{Y_{LK}}_{>0} \cdot (r + \tau)} \tag{30}$$

The total differential is then given by equation (29). Let's go through it from left to right: The first term on the left-hand side (LHS) is the derivative of $Y_K \cdot w$ with respect to K. w is a constant, which remains in place, the derivative of Y_K is given by Y_{KK}. The second term on the LHS is the derivative of $Y_K \cdot w$ with respect to τ. However, neither Y_K nor w depend on τ.

The first term on the right-hand side (RHS) is the derivative of $Y_L \cdot (r + \tau)$ with regard to K. We treat $(r + \tau)$ as a constant and get Y_{LK} as the derivative of Y_L with regard to K. The second term on the RHS is the derivative of $Y_L \cdot (r + \tau)$ with regard to τ. We treat Y_L and r as constants.

The last step contains an important assumption: By treating r as a constant, we assume that it is not influenced by a change in τ. We will see in one of the first lectures that this is often called the small-country assumption, because investors can avoid an impact of the tax on their return by moving capital to countries which do not raise their taxes.

Finally, we rearrange and arrive at equation (30). The LHS of this equation tells us that we have found the answer to our question. We are looking at an expression for the relationship between changes in K and changes in τ. The term on the RHS is always negative: an increase in τ will lead to a decrease in K. Additionally, if we were interested in determining the strength of the effect, we could further analyze the magnitude of the parameters.

Since we are taking the perspective of the government, we can now include this reaction of the firms into our own optimization calculus and ask: Given that firms will always react as in equation (30), what is our optimal choice for the tax rate? This, of course, depends on our reasons for levying the tax and it is one of the questions that we will be asking many times throughout the remainder of the course.

In this section we have learned to use the total differential. It allows us to treat an equation as a condition and ask: Given that this condition must hold, how will a change in one variable translate into a change in another variable? Or more generally: How are the changes in different variables related to each other?

References

Chiang, Alpha (1984). *Fundamental Methods of Mathematical Economics*. 3rd ed. Singapore: McGraw-Hill, 1984.

Inada, Ken-Ichi (1963). “On a Two-Sector Model of Economic Growth: Comments and a Generalization”. In: *The Review of Economic Studies* 30.2 (June 1963), pp. 119–127.

A Rules of differentiation

f	f'
a	0
au	au'
$u+v$	$u'+v'$
uv	$u'v+uv'$
$u(v)$	$u'(v)v'$
$\frac{u}{v}$	$\frac{u'v-v'u}{v^2}$

Table 2: Rules of differentiation

f	f'
x^n	nx^{n-1}
e^x	e^x
$\ln x$	$1/x$
$\sin x$	$\cos x$
$\cos x$	$-\sin x$

Table 3: Some derivatives

Table 2 lists the most important rules of differentiation; a is a constant, u and v are functions. Table 3 gives the derivatives of frequently used functions; n is a constant and x is the input.

When taking the partial derivative of a function with more than one input, simply treat all terms as constants that do not contain the variable with regard to which you are differentiating.

B Exercises: Differentiation

1. Calculate all first-order partial derivatives of each function, determine whether the function is homogeneous and, if so, calculate its degree of homogeneity.
   (a) $Y(K,L) = K^{0.3}L^{0.2}$
   (b) $Y(K,L,H) = K^{0.3}L^{0.7}H^{0.2}$
   (c) $Y(K,L) = K^{0.5} + L^{0.5}$
   (d) $Y(K,L) = K^{0.3} + L^{0.7}$
   (e) $Y(K,L) = \ln(KL)$
   (f) $Y(K,L) = \sqrt{KL}$
2. Calculate the optimal input ratio for the following combinations of production functions and cost functions. (The production function in (d) has three inputs and, as a consequence, three ratios that characterize the optimal mix of inputs. Calculate all three.) Use the total derivative to determine how firms' usage of capital K will react to an exogenous change in their usage of labor L.
   (a) $Y(K,L)$ and $C(K,L) = rK^2 + wL^2$
   (b) $Y(K,L) = K^{\alpha}L^{\beta}$ and $C(K,L) = rK + wL$
   (c) $Y(K,L) = ln(KL)$ and $C(K,L) = rK + wL$
   (d) $Y(K,L,H) = K^{\alpha}L^{\beta}H^{\gamma}$ and $C(K,L) = rK + w_L L + w_H H$