Mathematics Review

Revised: January 5, 2012

Mathematics is a precise and efficient language for expressing quantitative ideas, including many that come up in business. What follows is an executive summary of everything you'll need in this course: functions, exponents and logarithms, derivatives, and spreadsheets, each illustrated with examples.

Functions

In economics and business, we often talk about relations between variables: demand depends on price, cost depends on quantity produced, price depends on yield, output depends on input, and so on. We call these relations *functions*. More formally, a function f assigns a (single) value y to each possible value of a variable x. We write it this way: $y = f(x)$. Perhaps the easiest way to think about a function is to draw it: put x on the horizontal axis and plot the values of y associated with each x on the vertical axis. In a spreadsheet program, you might imagine setting up a table with a grid of values for x. The function would then be a formula that computes a value y for each value of x.

Example (demand functions). We may be interested in the sensitivity of demand for our product to its price. If demand is q and price p, an example of a demand function relating the two is

$$q \;=\; a + bp,$$

where a and b are "parameters" (think of them as fixed numbers whose values we haven't bothered to write down). Sensitivity of demand to price is summarized by b, which we'd expect to be negative (demand falls as price rises).

Example (bond yields). The price p and yield y for a one-year zero-coupon bond might be related by

$$p \;=\; 100/(1+y),$$

where 100 is the face value of the bond. Note the characteristic inverse relation: high yield, low price.

Example (production functions). In this class we'll relate output Y to inputs of capital K and labor L. (In macroeconomics, capital refers to plant and equipment.) It'll look a little strange the first time you see it, but a convenient example of such a function is

$$Y = K^{\alpha} L^{1-\alpha},$$

where α is a number between zero and one (typically we set $\alpha = 1/3$). This is a modest extension of our definition of a function — Y depends on two variables, not one — but the idea is the same.

Exponents and logarithms

Exponents and logarithms are useful in lots of situations: elasticities, compound interest, growth rates, and so on. Here's a quick summary.

Exponents. Exponents are an extension of multiplication. If we multiply x by itself, we could write either $x \times x$ or x^2, where 2 is an exponent (or power). In general, we can write x^a to mean (roughly) "x multiplied by itself a times," although this language may seem a little strange if a isn't a positive whole number like 2 or 3. We can nevertheless compute such quantities for any value of a we like as long as x is positive. (Think about how you'd do this in a spreadsheet.)

The most useful properties of exponents are

$$\begin{aligned} x^a x^b &= x^{a+b} \\ x^a y^a &= (xy)^a \\ (x^a)^b &= x^{ab}. \end{aligned}$$

You can work these out for yourself using our multiplication analogy.

Logarithms. By logarithm we mean the function "LN" in Microsoft Excel, OpenOffice Calc, or Google spreadsheets, sometimes called the natural logarithm. This will seem strange, perhaps worse, but the natural logarithm of a number x comes from the power of a number e, a mathematical constant that is approximately 2.718. If $x = e^y$, then y is the logarithm of x, expressed $y = \log x$. (There are other logarithms based on powers of other numbers, but we'll stick with e.) Suppose, instead, you know that y is the logarithm of x. How do you find x? From the definition, apparently $x = e^y$. In Excel, this is written "exp(y)". As a check, you might verify that $\log 6 = 1.792$ and $\exp(1.792) = 6.00$.

The most useful properties of logarithms are

$$\begin{aligned} \log(xy) &= \log x + \log y \\ \log(x/y) &= \log x - \log y \\ \log(x^a) &= a \log x \\ \log(e^x) &= x \\ e^{\log x} &= x. \end{aligned}$$

In short, logarithms convert multiplication into addition, division into subtraction, and "exponentiation" into multiplication. In each case, an operation is converted into a simpler one: addition, for example, is simpler than multiplication.

Example (demand functions). A more useful demand function is $q = ap^b$, which is linear in logarithms:

$$\log q \ = \ \log a + b \log p.$$

This follows from the first and third properties of logarithms. Here b is the price elasticity.

Example (compound interest). Our earlier relation between price and yield was based on a compounding interval of one year, the same as the maturity of the bond. In practice, people use lots of different compounding intervals, creating no end of confusion. US treasuries, for example, are based on semi-annual compounding, which implies

$$p \ = \ 100/(1 + y/2)^2,$$

where y is the "semi-annually compounded" yield. If we compound n times a year, the relation is

$$p \ = \ 100/(1 + y/n)^n.$$

If $n = 2$ we compound twice a year (semi-annually), if $n = 12$ we compound twelve times a year (monthly), and so on. For n large, this becomes (trust us)

$$p \ = \ 100\, e^{-y} \ = \ 100\ \exp(-y),$$

where y is referred to as the "continuously-compounded" yield. With continuous compounding, how do we find the yield if we know the price? The answer: use $y = \log 100 - \log p$. (This follows from the first and fourth properties of logarithms.)

Example (long bonds). The choice of compounding interval is arbitrary — we can choose any interval we like. For a zero-coupon bond with a maturity of m years, three versions of the relation between the price p and yield y are

$$\begin{aligned} \text{Annual compounding:} &\quad p \ = \ 100/(1 + y)^m \\ \text{Semi-annual compounding:} &\quad p \ = \ 100/(1 + y/2)^{2m} \\ \text{Continuous compounding:} &\quad p \ = \ 100\, e^{-my} \ = \ 100\ \exp(-my). \end{aligned}$$

The choice is a matter of convenience and tradition: each definition of the yield contains the same information. What's nice about continuous compounding is that, once you take logs, you simply multiply the yield times the number of periods, rather than the more complicated compounding we usually get.

Example (growth in the US). In the US, real GDP was \$10,074.8b in 2002, \$10,381.3b in 2003. What was the growth rate? Approach 1 (annual compounding): solve

$$1 + g \ = \ 10381.3/10074.8$$

for the (simple) growth rate g. The answer: $g = (10381.3/10074.8) - 1 = 0.0304 = 3.04\%$. Approach 2 (continuous compounding): solve

$$e^{\gamma} \;=\; 10381.3/10074.8$$

for γ. The answer: $\gamma = \log(10381.3/10074.8) = 0.0300 = 3.00\%$. The answers are very similar, which will be true as long as the growth rates are small. If you plot both over time for the US, you'll have a hard time telling the difference. We'll typically use the second approach.

Example (growth in Korea). GDP per capita in Korea was \$770 in 1950, \$14,343 in 2000, measured in 1990 US dollars. What was the average annual growth rate? Approach 1 (annual compounding): find the number g satisfying

$$14343 \;=\; (1+g)^{50} 770.$$

How do we find g? Using logarithms, of course! Note that

$$\log(14343/770) = 50 \log(1+g).$$

Since $\log(14343/770) = 2.925$, $\log(1+g) = 2.925/50 = 0.0585$, and $1+g = \exp(0.0585) = 1.0602$. Thus the growth rate was 6.02% a year, which is extraordinarily high. Approach 2 (continuous compounding): solve

$$14343 \;=\; e^{50\gamma}\, 770.$$

The answer: $50\gamma = \log(14343/770) = 2.925$, so $\gamma = 2.925/50 = 0.0585$ or 5.85%.

Slopes and derivatives

The slope of a function is a measure of how steep it is: the ratio of the change in y to the change in x. For a straight line, we can find the slope by choosing two points and computing the ratio of the change in y to the change in x. For some functions, though, the slope (meaning the slope of a straight line tangent to the function) is different at every point.

The *derivative* of a function $f(x)$ is a second function $f'(x)$ that gives us its slope at each point x if the function is continuous (no jumps) and smooth (no kinks). Formally, we say that the derivative is

$$\frac{\Delta y}{\Delta x} \;=\; \frac{f(x+\Delta x) - f(x)}{\Delta x}$$

for a "really small" Δx. (You can imagine doing this on a calculator or computer using a particular small number, and if the number is small enough your answer will be pretty close.) We express the derivative as $f'(x)$ or dy/dx and refer to it as "the derivative of y with respect to x." The d's are intended to be suggestive of small changes, analogous to Δ but with the understanding that we are talking about very small changes.

So the derivative is a function $f'(x)$ that gives us the slope of a function $f(x)$ at every possible value of x. What makes this useful is that there are some relatively simple mechanical rules for finding f' for common functions f. See Exhibit 1. If these rules are new to you, take them as facts to be memorized and put to work.

Example (marginal cost). Suppose total cost c is related to the quantity produced q by

$$c \;=\; 100 + 10q + 2q^2.$$

Marginal cost is the derivative of c with respect to q. How does it vary with q? The derivative of c with respect to q is

$$dc/dq \;=\; 10 + 4q,$$

so marginal cost increases with q.

Example (bond duration). Fixed income analysts know that prices of bonds with long maturities are more sensitive to changes in their yields than those with short maturities. They quantify sensitivity with duration D, defined as

$$D \;=\; -\frac{d \log p}{dy}.$$

In words, duration is the ratio of the percent decline in price (the change in the log) over the increase in yield for a small increase. Two versions follow from different compounding conventions. With annual compounding, the price of an m-year zero-coupon bond is related to the yield by $p = 100/(1+y)^m$. Therefore

$$\log p \;=\; \log 100 - m \log(1+y)$$

and duration is $D = m/(1+y)$. With continuous compounding, $p = 100 \exp(-my)$, $\log p = \log 100 - my$, and $D = m$. In both cases, it's clear that duration is higher for long maturity bonds (those with large m).

Example (marginal product of capital). Suppose output Y is related to inputs of capital K and labor L by

$$Y \;=\; K^{\alpha} L^{1-\alpha}$$

for α between zero and one. If we increase K holding L fixed, what happens to output? We call the changes in output resulting from small increases in K the marginal product of capital. We compute it as the derivative of Y with respect to K holding L constant. Since we're holding L constant, we call this a *partial derivative* and write it

$$\frac{\partial Y}{\partial K} \;=\; \alpha K^{\alpha-1} L^{1-\alpha} \;=\; \alpha (L/K)^{1-\alpha}.$$

Despite the change in notation, we find the derivative in the usual way, treating L like any other constant.

Finding the maximum of a function

An important use of derivatives is to find the maximum (or minimum) of a function. Suppose we'd like to know the value of x that leads to the highest value of a function $f(x)$, for values of x between two numbers a and b. We can find the answer by setting the derivative $f'(x)$ equal to zero and solving for x. Why does this work? Because a function is "flat" (has zero slope) at a maximum. (That's true, anyway, as long as the function has no jumps or kinks in it.) We simply put this insight to work.

Fine points (feel free to skip). Does this always work? If we set the derivative equal to zero, do we always get a maximum? The answer, in a word, is no. Here are some of the things that could go wrong: (i) The point could be a minimum, rather than a maximum. (ii) The maximum could be at one of the endpoints, a or b. There's no way to tell without comparing your answer to $f(a)$ and $f(b)$. (iii) There may be more than one "local maximum" (picture a wavy line). (iv) The slope might be zero without being either a maximum or a minimum: for example, the function might increase for a while, flatten out (with slope of zero), then start increasing again. An example is the function $f(x) = x^3$ at the point $x = 0$. [You might draw functions for each of these problems to illustrate how they work.] If you want to be extra careful, there are ways to check for each of these problems. One is the co-called second-order condition: a point is a maximum if the second derivative [the derivative of $f'(x)$] is negative. All of these things can happen in principle, but one of our jobs is to make sure they do not happen in this class. And they won't.

Example (maximizing profit). Here's an example from Firms & Markets. Suppose a firm faces a demand for its product of $q = 10 - 2p$ (q and p being quantity and price, respectively). The cost of production is 2 per unit. What is the firm's profit function? What level of output produces the greatest profit?

Answer. Profit is revenue (pq) minus cost ($2q$). The trick (and this isn't calculus) is to express it in terms of quantity. Apparently we need to use the demand curve to eliminate price from the expression for revenue: $p = (10 - q)/2$ so $pq = [(10 - q)/2]q$. Profit (expressed as a function of q) is therefore

$$\text{Profit}(q) \;=\; [(10-q)/2]q - 2q \;=\; 5q - q^2/2 - 2q.$$

To find the quantity associated with maximum profit, we set the derivative equal to zero:

$$\frac{d\text{Profit}}{dq} \;=\; 3 - q \;=\; 0,$$

so $q = 3$. What's the price? Look at the demand curve: if $q = 3$, then p satisfies $3 = 10 - 2p$ and $p = 7/2$.

Example (demand for labor). A firm produces output Y with labor L and a fixed amount of capital K, determined by past investment decisions, subject to the production function $Y = K^\alpha L^{1-\alpha}$. If each unit of output is worth p dollars and each unit of labor costs w dollars, then profit is

$$\text{Profit} \;=\; pK^\alpha L^{1-\alpha} - wL.$$

The optimal choice of L is the value that sets the derivative equal to zero:

$$\frac{\partial \text{Profit}}{\partial L} \;=\; p(1-\alpha)(K/L)^\alpha - w \;=\; 0.$$

(We use a partial derivative here to remind ourselves that K is being held constant.) The condition implies

$$L \;=\; K\left[\frac{p(1-\alpha)}{w}\right]^{1/\alpha}.$$

You can think of this as the demand for labor: given values of K, p, and w, it tells us how much labor the firm would like to hire. As you might expect, at higher wages w, labor demand L is lower.

Spreadsheets

Spreadsheets are the software of choice in many environments. If you're not familiar with the basics, here's a short overview. The structure is similar in Microsoft Excel, OpenOffice Calc, and Google documents.

The first step is to make sure you have access to one of these programs. If you have one of them on your computer, you're all set. If not, you can download OpenOffice at www.openoffice.org or open a Google spreadsheet at docs.google.com. Both are free.

In each of these programs, data (numbers and words) are stored in tables with the rows labeled with numbers and the columns labeled with letters. Here's an example:

	A	B	C
1	x1	x2	
2	3	25	
3	8	13	
4	5	21	
5			

The idea is that we have two (short) columns of data, with variable x1 in column A and variable x2 in column B.

Here are some things we might want to do with this data, and how to do it:

- Basic operations. Suppose you want to compute the natural logarithm of element B2 and store it in C2. Then in C2 you would type: `=LN(B2)`. (Don't type the period, it's part of the punctuation of the sentence.) The answer should appear almost immediately. If you want to add the second observation (row 3) of x1 and x2 and put in in C3, then in C3 you type: `=A3+B3`.
- Statistics. Suppose you want to compute the sample mean and standard deviation of x1 and place them at the bottom of column A. Then in A5 type: `=AVERAGE(A2:A4)`. That takes the numbers in column A from A2 to A4 and computes the sample mean or average. The standard deviation is similar: in A6 you type `=STDEV(A2:A4)`. Finally, to compute the correlation between x1 and x2, you type (in any cell you like): `=CORREL(A2:A4,B2:B4)`.

If you're not sure what these functions refer to, see the links to the Kahn Academy videos at the end.

Review questions

If you're not sure you followed all this, give these a try:

1. Growth rates. Per capita income in China was 439 in 1950, 874 in 1975, and 3425 in 2000, measured in 1990 US dollars. What were the annual growth rates in the two subperiods?

   Answer. The average continuously-compounded growth rates were 2.75% and 5.46%. The simple growth rates (these are harder) are 2.79% and 5.62%, so there's not much difference between them.

2. Find the derivative of each of these functions:
   (a) $2x + 27$ [2]
   (b) $2x^2 + 3x + 27$ $[4x + 3]$
   (c) $2x^2 + 3x - 14$ $[4x + 3]$
   (d) $(x - 2)(2x + 7)$ $[4x + 3]$
   (e) $\log(2x^2 + 3x - 14)$ $[(4x + 3))/(2x2 + 3x - 14)]$
   (f) $3x^8 + 13$ $[24x^7]$

(g) $3x^{2/3}$ [$2x^{-1/3} = 2/x^{1/3}$]

(h) $2e^{5x}$ [$10e^{5x}$]

Answers in brackets [ ].

3. Suppose output is related to the amount of capital used by

$$Y \;=\; 27K^{1/3}.$$

Compute the marginal product of capital (the derivative of Y with respect to K) and describe how it varies with K.

Answer. The marginal product of capital is $\text{MPK} = 9K^{-2/3} = 9/K^{2/3}$ is positive and falls as we increase K. We call this diminishing returns: the more capital we add, the less it increases output.

4. Find the value of x that maximizes each of these functions:

(a) $x^2 - 2x$ [$f'(x) = 2x - 2 = 0$, $x = 1$]

(b) $2\log x - x$ [$f'(x) = 2/x - 1 = 0$, $x = 2$]

(c) $5x^2 - 2x + 11$ [$f'(x) = 10x - 2 = 0$, $x = 1/5$]

5. You have the following data: 4, 6, 3, 4, 5, 8, 5, 3, 6. What is the mean? (Use a spreadsheet program to do the calculation.)

Answer. 4.89.

If you're looking for more

If these notes seem mysterious to you, we recommend the Kahn Academy. He has wonderful short videos on similar topics, including logarithms (look for "Proof: $\log a$ ..."), calculus (look for "Calculus: Derivatives ..."), and statistics (start at the top). For spreadsheets, the Google doc tutorial is quite good.

Exhibit 1. Rules for Computing Derivatives

Function $f(x)$	Derivative $f'(x)$	Comments
	Rules for Specific Functions	
a	0	a is a number
$ax + b$	a	a and b are numbers
ax^b	bax^{b-1}	a and b are numbers
ae^{bx}	bae^{bx}	a and b are numbers
$a \log x$	a/x	a is a number
	Rules for Combinations of Functions	
$g(x) + h(x)$	$g'(x) + h'(x)$	
$ag(x) + bh(x)$	$ag'(x) + bh'(x)$	a and b are numbers
$g(x)h(x)$	$g(x)h'(x) + g'(x)h(x)$	
$g(x)/h(x)$	$[g'(x)h(x) - g(x)h'(x)]/[h(x)]^2$	$h(x) \neq 0$
$g[h(x)]$	$g'[h(x)]h'(x)$	"chain rule"