Advanced Monetary Policy

Technical note

Math tools

rhys.m.bidder@kcl.ac.uk

Spring 2024

Abstract

This note lays out various mathematical tools that we will use in the course. As the course progresses I will add anything that has caused confusion, so please LMK. Many of these tools you won't need to memorize or use directly in assessment, but you should be able to understand vaguely how they are being applied. In fact, you will use many of them in a concrete way when working with Matlab.

1 Notation

You will have already encountered much of the notation that economists frequently use. Here is a brief summary.

For scalar variables (i.e. a variable that is just 'a number'), we often use lower case letters. Perhaps we use c for consumption, i for the interest rate. We may use Greek lower case letters too, say π for inflation, though they are often used to denote scalar *parameters* too. For example, σ is often used to denote a standard deviation, or μ to denote a mean. Loosely speaking, *variables* are things, like consumption, that vary over time in our models, while *parameters* are fixed numbers describing some property of the economy. We might occasionally imagine would would happen if a parameter changes (say, a parameter that describes the preferences of a central bank) but you should regard that as the model changing, rather a change *within* a given model.

Lower case is also used for vectors, say consumption in multiple periods, such as (see here for vectors in LaTex):

$$c = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$$

though people sometimes like to use $\vec{c}$. Personally I don't bother with the latter unless it's very unclear from the context whether c is a scalar or a vector. Above, I have written a vector in column form, which can be a pain when writing things down, as it takes up space. If I *do* want to write out some elements explicitly (rather than using c or $\vec{c}$), I would likely express it as a row vector with a transposition operator - the symbol $'$. This indicates it should be understood as being transposed (or flipped so that columns are rows and rows are columns).

$$c = \begin{pmatrix} c_1 \, c_2 \, \cdots \, c_n \end{pmatrix}'$$

The two definitions of c above are equivalent. Note that some textbooks use ^{T} rather than $'$.

Upper case letters like A or B are often used to denote matrices (2-D arrays of numbers). We will come back to these later in the course. Like vectors, I will be trying to avoid any complex use of them but we *will* be using them in Matlab. In fact, they are much easier to understand in Matlab!

A two by two matrix might be written (see here for matrices in LaTex):

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

where it is common to use the subscripts such that the first index of a_{ij} denotes the row (i), and the second denotes the column (j). It is standard to denote the elements (here, a) of a matrix with the lower case of the letter that denotes the matrix itself (here, A). As another example, in

$$B = \begin{pmatrix} 7 & -1 \\ -2 & 4 \end{pmatrix}$$

we have that $b_{21} = -2$.

Upper case Greek letters often also are used to denote matrices (for example Σ for a covariance matrix) though sometimes, such as in the case of Π, they may also denote a gross rate. For example, π might be the *net* rate of inflation, say 0.02 (indicating 2%), while $\Pi = 1.002$ denotes the associated *gross* rate of inflation. This sort of upper/lower relation also often applies to standard letters - for example, R might denote the gross real interest rate, while r denotes the net real rate ($R = 1 + r$).[1]

To denote a time period, we often use t. If we're referring to 1 or τ periods in the future from the perspective of a period t, we might write $t+1$ or $t+\tau$. Often we will subscript a variable by a time period, as many of our variables - for example consumption, c - progress over time (consumption this year, next year, today, tomorrow...) - so we might write c_t to mean consumption in period t.

If we are indexing 'things', we might use an index i, j, or k (we already saw t being used to index 'time' in the previous paragraph). That is, when I want to refer to a generic person, or maybe a generic Treasury maturity, I might say 'person i' or 'bond maturity k'. For example, consumption of person i in period $t + \tau$ might be written $c_{i,t+\tau}$) or the yield in t on a k-maturity bond as $y_{k,t}$.

This isn't weird or magic - it's just solving the question of how to write down stuff about multiple things (people, time periods,...) without having to use numbers and pages and pages of paper. Suppose I am writing a model that applies now (t) and in the infinite future. I'm going to struggle to write down $c_t, c_{t+1}, c_{t+2}, c_{t+3}, \ldots$ (an infinite set of $c_{t+\tau}$) on a piece of paper, right? Much better instead to say $t + \tau$ and say τ implicitly goes from 0 to infinity (or write $\tau = 0, \ldots \infty$).

[1] Warning, people sometimes use R and r when talking about nominal interest rates. I try to avoid that and use i for the net nominal rate and R^N for the gross nominal. If I am talking about gross rates I actually will frequently write R^{Re} and R^{Nom} or something like that.

2 Variables, coefficients, and exponents

In your previous Monetary/Financial course you encountered (something like) the following equations:

$$\pi = \beta E[\pi] + \kappa(y - y^*) + \epsilon_{PC} \tag{1}$$

$$y - y^* = E[y - y^*] - \sigma(i - E[\pi]) + \epsilon_{IS} \tag{2}$$

$$i = a(\pi - \pi^*) + b(y - y^*) + \epsilon_{Pol} \tag{3}$$

with Equation 1 being the New Keynesian Phillips Curve, Equation 2 being an aggregate demand, or 'dynamic IS' curve (sometimes I might say 'schedule' instead of curve) and Equation 3 being a Taylor-type policy rule for the short rate.[2] We have in this system of equations:

- *Variables:* Inflation (π), Output (y), 'Trend' output (y^*) and random shocks (ϵ_{PC}, ϵ_{IS} and ϵ_{Pol})
  - Inflation, output and 'trend' output are *endogenous* variables (determined within the model)
  - The shocks are *exogenous* variables (introduced *deus ex machina* according to a random process)
  - See here for more on endogenous vs exogenous
- *Coefficients:*
  - β, κ, σ, a and b are all coefficients - they **multiply** the variables
- *Parameters:*
  - All the coefficients (except maybe κ as discussed below) are parameters
  - Target inflation (π^*) is also a parameter and is a constant (or you could think of it as a 'coefficient' multiplying the number '1')
  - Implicitly the shocks will have some distribution - maybe joint normal, with a mean vector $\vec{\mu}$ and a 3-by-3 covariance matrix Σ, which are also parameters

Let us briefly comment on κ and whether it is a parameter. If you take these three equations as a starting point, then it would be regarded as a parameter. However, as we will see later in the course, when we (partly) derive the NK Phillips Curve, it actually is a mix of more primitive parameters that describe a more elaborate underlying model. So, in a sense, κ is a parameter that is a function of other parameters.

Now, if a number is multiplied by itself b times then we might write, for, $b = 4$

$$2 \times 2 \times 2 \times 2$$

[2] The equations are simplified (note the absence of time subscripts) but gets across the points I want to make.

Obviously that's ridiculously inefficient and confusing notation, so we would in fact instead say that 2 is raised to the exponent, 4 and write 2^4. Generally, we will often encounter expressions where a variable is multiplied by a coefficient and also raised to an exponent. Speaking generally, we might encounter something like ax^b. Notice that we have

$$ax^b \times cx^d = acx^{b+d}$$

So the coefficient of the product is the product of the individual coefficients and the exponent is the sum of the individual exponents.[3]

Now, we also write $\frac{1}{x^b}$ as x^{-b}.[4] Note that dividing x by x^2 is like multiplying x by $\frac{1}{x^2}$. So we use the rule above and say the answer is

$$x \times x^{-2} = x^1 \times x^{-2} = x^{1-2} = x^{-1} = \frac{1}{x}$$

where we made explicit that x is itself to the power 1.

But what about x divided by x. We know that is 1 regardless of what x and we see that this means $x^0 \equiv 1$.[5] In fact, any power of x divided by itself is 1 since

$$x^b \div x^b = x^{b-b} = x^0 = 1$$

Suppose we want to take something with an exponent and raise it to another exponent? Suppose we want $(2^3)^2$? Well, remember first that $(2^3)^2 \equiv (2^3) \times (2^3)$ which is $(2 \times 2 \times 2) \times (2 \times 2 \times 2)$, which is 2^6. Generally we have

$$(ax^b)^c = a^c x^{bc}$$

Finally, though this is really just an reuse/re-expression of previous results, suppose we are looking for a square root of a number, x. That is a number s such that $s^2 = x$. For example, we know that 2 and -2 are both square roots of 4 as $2^2 = 4$ and $(-1)^2 = 4$. Using our results on multiplication of expressions with variables raised to exponents we see that s is naturally written $s = x^{\frac{1}{2}}$, as

$$x^{\frac{1}{2}} \times x^{\frac{1}{2}} = x^{\frac{1}{2}+\frac{1}{2}} = x^1 \equiv x$$

[3]Note that $ax^b \times cy^d = abx^b y^c$ you only add exponents that apply to the same variable. Here we have two variables x and y.

[4]Based on the discussion below, you can look at this as $x^0 \div x^b = x^{0-b} = x^{-b}$

[5]There are conventions for how to handle 0^0, see here, but frequently it is useful to define it to be 1. Similarly, the convention also applies to ∞^0.

Generally, we have that if $x^a = y$ then $x = y^{\frac{1}{a}}$, which defines the a^{th} root (if it exists).

3 Functions

A function is a rule that maps an input to a single output, over some domain of inputs. It may be that two inputs give the same output, but it is key that for any single input, there is only one output.[6] You have encountered many functions in your studies so far (and in life generally). For example, a Taylor rule for monetary policy is a function that maps from inputs such as the inflation rate, the output gap and possibly a random shock, to the short rate that the policymaker is 'supposed' to set. We could write

$$i_t = f(\pi_t, y_t - y_t^*, \epsilon_t)$$

with the function, f, defined as

$$f(x, y, z) = \pi^* + \alpha(x - \pi^*) + \gamma y + z$$

where π^* is a fixed number (the inflation target) and α and γ are fixed coefficients. π^*, α and γ are parameters and together with the structure of the equation (what gets added to, or multiplied by, what etc.) they define the function. Typically people do not use notation that makes the parameters explicit, though when programming up such things, that often is useful (we will discuss this in another technical note on Matlab). In that case, you might see something like

$$f(x, y, z; \pi^*, \alpha, \gamma) = \pi^* + \alpha(x - \pi^*) + \gamma y + z$$

with the ; separating the *arguments of the function* (here denoted x, y and z) from parameters used to define it.

Note that the function can be defined using any sort of dummy variables or symbols to represent the arguments. What do I mean by that? Well, above, I defined f using x, y and z but when I used it in the Taylor rule, I imagined inputting inflation, the output gap and the shock, in the three 'slots' of the function. It doesn't matter what symbols I use when *defining* the function and to emphasize this, when defining f I used y as the second argument but at that point, I was not claiming any relationship with the y in the model (output). When I actually come to use the function, I can stick in whatever I want

[6] $f(x) = x^2$ is an example of a function where two inputs may give the same output. For example $(-2)^2 = 2^2 = 4$ so both -2 and 2 give the same output.

and in that 'slot' (the second 'slot' of the function) I inserted the output gap, $y_t - y_t^*$. I put this slightly awkward situation in as an example to alert you to something that often confuses students. Bottom line, When a function is being defined, you can do pretty much whatever with the arguments in the function definition. I could have defined f instead as

$$f(\dagger, \ddagger, \frownie) = \pi^* + \ddagger(\dagger - \pi^*) + \frownie y + z$$

As another example, commonly when defining a utility function that will eventually be used with consumption, c, as an argument, I might still define it as $u(x) = \log x$, or whatever...

Sometimes when I'm referring to a function, I may just refer to a function f without even mentioning its arguments or parameters. Of course, we can call functions by pretty much any name, but f is commonly the first choice! If there is obvious interpretation, like for a 'utility function' in economics it's common to use u or U.

4 Basic calculus

Economics involves a lot of optimization. Maximizing profit or utility, minimizing cost and so forth. Calculus (and differentiation in particular) is hard to avoid. I will use it in deriving various things but I don't expect a deep understanding from you.

Happily, in the simple cases we work with, there are some very basic rules of thumb that will be enough to understand everything we do. You won't need to remember them, but you should be able to refer to them when understanding some of the equations we derive. I am not going to provide a deep dive into differentiation but a very nice (and short) guide is here.

Very briefly, if I have a function f, then differentiation gives me another function that gives the slope of f at any particular value of its input. Often it is denoted f'. Suppose the function takes a nice simple form like

$$f(x) \equiv ax^b$$

So, if I 'input' x the function outputs x raised to an exponent (or power) b and multiplied by a coefficient a. In that case the rule for differentiating it is simply that we multiply the expression by the exponent and then lower the exponent by 1 to get

$$f'(x) = abx^{b-1}$$

Suppose I have $a = 1$ and $b = 1$. Then f describes a straight line and I plot it in Figure 1 in the left panel. The slope of f is always 1 as $f(x) = x$. If I increase x by ϵ then $f(x)$ increases by ϵ. I plot f' in the right panel. The slope of f is constant (i.e. it's the same - equal to 1 - regardless of x), which is what f' reflects (it's flat at the value 1).

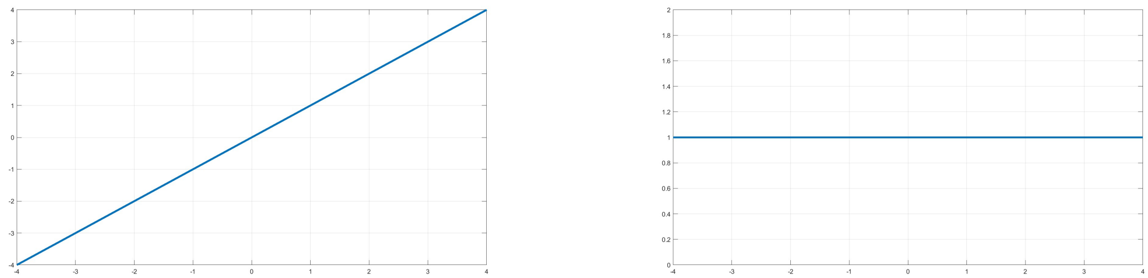

Figure 1: Left panel: $f(x) = x$. Right panel: $f'(x) = 1$

Suppose I have $a = 1$ and $b = 2$. Then f describes a 'parabola' (a U shape), as shown in Figure 2 in the left panel. The slope of f in this case depends on the input. You can see that the slope is negative when $x < 0$, equal to zero when $x = 0$ and positive when $z > 0$. Indeed you can see that the slope is gradually less negative as we approach $x = 0$ from below, and gets increasingly positive as x increases above 0. What is the slope? Well, using our rule we can obtain f' which implies $f'(x) = 2x$.

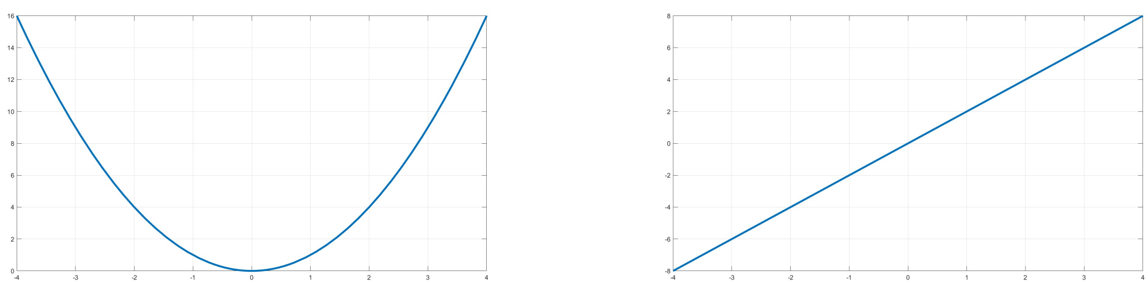

Figure 2: Left panel: $f(x) = x^2$. Right panel: $f'(x) = 2x$

Think about it: $f'(x) = 2x$ makes sense, qualitatively. It is negative for $x < 0$, zero at $x = 0$ and positive at $x > 0$. It is also increasing with x, so it must be gradually getting less negative as it approaches $x = 0$ from below and more positive as it continues moving upwards.

Differentiation gives you the slope of a function or, alternatively, the rate of change for an infinitesimal change in the input. Note that the point at which it is being evaluated matters. For some functions, f' is a constant function - i.e. whatever x is, it gives the same value (like in our first example above) but this is not a generic thing. In our second example, the derivative (the value that f' returns) depends on x. How weird would it be if there was one *number* for the slope of a given function? Obviously that can't be right - we know (see the second example) that the slope can be all sorts of values at different inputs, so naturally f' will have to be a function too - not a single number, but a rule that gives back a number if you plug in an input, x.

Will all our functions look like ax^b? Nope. But in this course most functions will be built up from such expressions. They may involve sums of them, or products (multiplied together) or perhaps ratios (divided by eachother), but there are simple rules for how to handle this.

First, consider a situation where we have

$$\begin{aligned} f(x) &= g(x) + h(x) \\ g(x) &= ax^b \\ h(x) &= cx^d \end{aligned}$$

It turns out that differentiation is linear. So the derivative of the sum is equal to the sum of derivatives. That is $f'(x) = g'(x) + h'(x)$. We know how to differentiate g and h so we're in business. So in this case $f'(x) = abx^{b-1} + cdx^{d-1}$.

Now consider a situation where we have

$$f(x) = g(x)h(x)$$

It turns out that there is a rule for what to do here, called the 'product rule' which tells us what f' is. Happily the expression is quite simple

$$f'(x) = g(x)h'(x) + g'(x)h(x)$$

The algebra to get the answer is a little annoying, but you know g and h, you know how to get g' and h', and then you just need to multiply and add things together.[7]

Finally, there is also a quotient rule that tells us what to do when we need to differentiate a ratio of functions:

$$f(x) = \frac{g(x)}{h(x)}$$

[7] The answer is $ac(b+d)x^{b+d-1}$ - try it. You will initially get something like $acdx^{b+d-1} + abcx^{b+d-1}$ and then you can simplify/factorize like I have, though that tidying up doesn't change the answer.

We have in this case that[8]

$$f'(x) = \frac{g'(x)h(x) - g(x)h'(x)}{h(x)^2}$$

Note that for this and the previous example we actually could have simplified the product and ratio *before* doing differentiation on $f(x) = acx^{b+d}$ and $f(x) = \frac{a}{c}x^{b-d}$. But I wanted to show how the rules worked. Usually though, we don't have such a simple form. For example, suppose in the second example we had $h(x) = 1 + cx^d$. Try it.

It's important to note that one can (and typically will) combine these rules. For example, suppose in the last case, the top part of the ratio, was itself the product of two functions. Then we would get g' using the product rule, and then use that within the quotient rule application.

There is another rule that you should know about: the 'chain rule'. This is used when you are differentiating something 'inside' another function. So if I had a function f such that $f(x) = g(h(x))$ where g and h are other functions, it turns out that $f'(x) = h'(x)g'(h(x))$. So to get the slope of f as x changes you ask first how quickly the 'inside' (h) function changes with x and then ask how quickly the 'outside' function (g) changes with its input, which is being 'fed' to it by h. Effectively it's a multiplication of slopes, which makes sense, right?[9]

Note also that will often be working with functions that take two inputs (perhaps consumption today and consumption tomorrow). For our basic purposes not much changes, except we often will differentiate with respect to one variable at a time, imagining the other is fixed. What does this look like? Imagine we have

$$f(x, y) = ax^b y^c + dx^f y$$

We will take a 'partial' derivative individually with respect to x and also with respect to y

$$\begin{aligned} f_x(x, y) &= abx^{b-1}y^c + dfx^{f-1}y \\ f_y(x, y) &= acx^b y^{c-1} + dx^f \end{aligned}$$

We see that f_x and f_y are obtained using our standard rules, treating the 'y bits' as fixed when differentiating with respect to x and the 'x bits' fixed when differentiating with respect to y.

Why does differentiation come up a lot in economics? Because we often want to find the minimum or maximum of a function and this often implies we are looking for a value of x when the *level* of the derivative, f', is zero, which indicates that the *slope* of f is zero.

[8] The answer, after some tidying, is $\frac{a(b-d)}{c}x^{b-d-1}$.

[9] We see an example of this in the consumption-savings note I upload, where we express utility simply in terms of first period consumption, c_1, and replace c_2 with a function (implied by the intertemporal) budget constraint of c_1.

5 Maximization

In our simple cases, we will be doing maximization and minimization (to minimize f just maximize $-f$) in such a way that the only thing we need to do is find values of the inputs where the first derivatives (perhaps partial derivatives if we have multiple inputs) are equal to zero.

Intuitively, this is like finding a flat spot in the dimension of whatever variable(s) we're choosing to maximize our function. This is a bit like being at the top of a hill - go left or right and you'll be going downwards. If there is a non-zero slope then you must be able to go higher in some direction - so you can't be at a maximum. The zero slope is (in our applications) a necessary condition and (again, in our applications) they will also be sufficient.

Let us consider $f(x) = -x^2$, which is plotted in the left panel of Figure 3. You should be able to see that the maximum of f is zero and it is located at $x = 0$. Note the difference between the max and the value of the input that induced the max. The latter is called the 'arg max'. We know how to find f'. In this case, $f'(x) = 2x$. When is $2x = 0$? When $x = 0$ so we get the right result. We will see examples of maximization with two arguments when we obtain the Euler equation in the technical note on the 2-period consumption-savings problem.

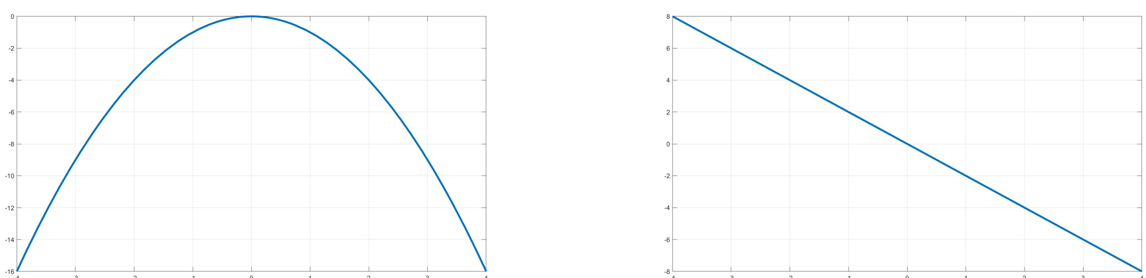

Figure 3: Left panel: $f(x) = -x^2$. Right panel: $f'(x) = -2x$

6 Logs and exponentials

A logarithm of y 'to base x' is the value to which x must be raised to make it equal y. That is, the log is defined by

$$x^{\log_x(y)} \equiv y$$

We will typically be working with the 'natural' logarithm which has the exponential constant, e, as its base. e is a number - a very particular number - but it's still 'just a number'. We have that

$$e = \sum_{n=0}^{\infty} \frac{1}{n!} \approx 2.71828$$

where the notation $n!$ doesn't mean we're amazed at the letter n, but $n! \equiv n \times (n-1) \times (n-2) \ldots 2 \times 1$, which is referred to as 'n factorial'. You should know that $\log_e(y)$ is sometimes written $\ln(y)$ but we will typically use $\log(y)$ in this course. For more on this stuff you could check out here.

Logarithms have some properties that can be useful when manipulating equations.

- Log of product = sum of logs

$$\log(xy) = \log(x) + \log(y)$$

- Exponents become coefficients

$$\log(x^y) = y \log(x)$$

- Log of ratio = difference in logs[10]

$$\log\left(\frac{x}{y}\right) = \log(x) - \log(y)$$

- Log of unity = 0 (anything raised to 0 equals unity)

$$\log(1) = 0$$

A few more properties of logs pop up a lot in economics (and in any scientific field):

- $\log(1+x) \approx x$ for small x
  - This is useful for manipulating gross and net interest rates.
- The difference in logs $\approx$ percentage difference (for small differences).
  - Log of today's value of a variable minus the log of yesterday's value $\approx$ the percent *growth rate*
  - $\log c_t - \log c_{t-1} \approx \frac{c_t - c_{t-1}}{c_{t-1}}$

[10] This is implied by log of product rule, combined with our earlier discussion of exponents