Intermediate Macroeconomics:
Math

Eric Sims
University of Notre Dame

Fall 2012

1 Introduction

Economics makes use of math. Much of what we will do in this course will use some relatively straightforward math of one variety or another. Before we get started, we will review some basic concepts.

2 Notation

A variable is something which can be represented by a number that can change, either for reasons "outside of the model" or as a consequence of a change in another variable. Parameters, in contrast, are taken to be constants which are fixed. I will typically (with a few exceptions) use Latin letters to denote variables and Greek letters to denote parameters.

There are two main kinds of variables in economics – endogenous and exogenous. Exogenous variables are taken "as given" when conducting a modeling exercise. Endogenous variables are variables whose values we want to determine given the values of the exogenous variables. For example, in a macro model, you might think of the level of productivity as exogenous, and you want to use that to determine values for variables like consumption, investment, etc., which are endogenous.

A lot of what we'll be doing in macro is dynamic – that is, looking at the behavior of variables across time. I'll often times write variables with time subscripts to denote that the variable is observed in units of time; e.g. x_t is the value of x observed at time t, and x_{t+1} is the value observed in period $t + 1$, etc.

I'll most often denote time by $t = 0, 1, 2, \ldots$. Whatever we want to call the time does not matter, so I usually just use integers in ascending order. The first period (usually either denoted 0 or 1) could be 1948, 1948 quarter 1, 1948 month 1, or 1948 week one. Years, quarters (three consecutive months), months, weeks, days, and hours are just different frequencies of time observations. In macro when we talk about time we usually are referring to years or quarters, the constraint being that the large macroeconomic data like gross domestic product are only collected at a quarterly

1

frequency. Williamson's textbook uses the notation that x is a value of the variable x observed in the current period, while x' is the value in the next period (be the periods days, months, quarters, or years). We'll often be working with two periods, which is a convenient abstraction – basically dividing up time into the present and the future.

Sometimes it is convenient to write sums of variables across time using shorthand notation. Suppose we want to look at the sum of the realizations of a variable x across time, where we take time to "begin" in period t:

$$S = x_t + x_{t+1} + x_{t+2} + x_{t+3} + \cdots + x_{t+T}$$

Written in summation notation, this would be:

$$S = \sum_{j=0}^{T} x_{t+j}$$

Above j is just a dummy indicator. What we do is to start at the "bottom" of the upper-case Greek sigma, Σ, and evaluate the inside at $j = 0$. Then we go to $j = 1$, which yields x_{t+1}, and add that to the first part. Then we keep going until we get to the top part of the upper-case sigma, $j = T$. Note that the sum of a constant times a variable is equal to the constant times the sum, since the constant would appear in every term in the sum:

$$S = \sum_{j=0}^{T} \alpha x_{t+j} = \alpha \sum_{j=0}^{T} x_{t+j}$$

The way I've written the sum above, time "begins" at t and moves forward from there (you can also move backward, and can use the summation operator notation analogously). That is arbitrary. Sometimes we'll also write it where time begins at 0 and then we can use t as the dummy indicator instead of j:

$$S = x_0 + x_1 + x_2 + x_3 + \ldots x_T$$
$$S = \sum_{t=0}^{T} x_t$$

These are equivalent ways of writing the same concept. The choice of when time "begins" is just arbitrary, and the integers that we use to keep track of time ($t + j$ or t) are also arbitrary – they could be years, months, whatever.

3   Exponents and Logs

We will be making frequent use of exponents and natural logs. The following are a sequence of rules for exponents. x denotes some random variable and α and γ are constants.

$$x^1 = x$$
$$x^0 = 1$$
$$x^{-1} = \frac{1}{x}$$
$$x^{-\alpha} = \frac{1}{x^\alpha}$$
$$x^\alpha x^\gamma = x^{\alpha+\gamma}$$
$$\frac{x^\alpha}{x^\gamma} = x^{\alpha-\gamma}$$
$$(x^\alpha)^\gamma = x^{\alpha\gamma}$$
$$x^\alpha y^\alpha = (xy)^\alpha$$

The natural log, which I'll denote by ln or log, is the inverse operator for the exponential function, which I will denote by $\exp(x)$, or less frequently just e^x. exp or e is a number, approximately equal to 2.718. This function has the peculiar property (defined more below) that $\exp(x)$ is its own derivative. Below are some properties:

$$\ln(\exp(x)) = x$$
$$\exp(\ln x) = x$$
$$\ln x^\alpha = \alpha \ln x$$
$$\ln(xy) = \ln x + \ln y$$
$$\ln\left(\frac{x}{y}\right) = \ln x - \ln y$$
$$\ln 1 = 0$$
$$\ln 0 = -\infty$$

4   Growth Rates

The growth rate of a variable x is defined as its change between two periods of time divided by the value in the "base" period. Sometimes we also call this the "percentage change," the change in the variable divided by its base. Most often when I use the term growth rate I mean the percentage change between two adjacent periods of time, but you could define growth rates over longer horizons.

The period-over-period growth rate of a variable x in period t is:

$$g_t^x = \frac{x_t - x_{t-1}}{x_{t-1}}$$

One can re-arrange this to get:

3

$$1 + g_t^x = \frac{x_t}{x_{t-1}}$$

A very useful fact is that the natural log of one plus a "small" number is approximately equal to the small number. Since growth rates as we have defined them are percentage changes, these are "small" numbers (e.g. 1 percent is 0.01). Taking natural logs of the above, we have:

$$\ln\left(1 + g_t^x\right) = \ln x_t - \ln x_{t-1} \approx g_t^x$$

For example, suppose that the growth rate is 0.02. $\ln(1 + .02) = 0.0198$, which is very close to 0.02. Suppose the growth rate is -0.02 (i.e. x declined between $t - 1$ and t). Here $\ln(1 - 0.02) = -0.0202$, which is again very close. This approximation is sufficiently good that when economists talk about growth rates we use the log first difference interchangeably with the actual formula.

Another useful application of this is to look at a product of two variables. Suppose that $y_t = x_t z_t$, so that y is the product of x and z. Taking natural logs, we have:

$$\ln y_t = \ln x_t + \ln z_t$$

Now first differencing, we get:

$$\ln y_t - \ln y_{t-1} = \ln x_t - \ln x_{t-1} + \ln z_t - \ln z_{t-1} \Rightarrow g_t^y \approx g_t^x + g_t^z$$

In other words, the growth rate of a product is approximately equal to the sum of the growth rates.

5   Calculus

Suppose we have a continuous (e.g. no kinks or breaks) function of x, $y = f(x)$. The derivative is the slope of this function, which is itself usually a function (i.e. the derivative can be different at different values of x, unless the function is linear). A derivative is therefore a measure of how y changes as x changes.

We will use the following notation to refer to a derivative of a univariate function, which in words says "change in y given a change in x":

$$\frac{dy}{dx} = f'(x)$$

The second derivative is just the derivative of the derivative; it is a measure of how the slope (the first derivative) changes as x changes (recall again the derivative is a function itself, so you can take the derivative of a derivative). The notation we use for the second derivative is:

$$\frac{d^2y}{dx^2} = f''(x)$$

The following are some popular functions and their derivatives. In what follows x is taken to be a variable; $f(\cdot)$, $g(\cdot)$, and $h(\cdot)$ are functions; and α is a constant.

4

$$f(x) = \alpha, \qquad f'(x) = 0$$
$$f(x) = \alpha x, \qquad f'(x) = \alpha$$
$$f(x) = x^\alpha, \qquad f'(x) = \alpha x^{\alpha-1}$$
$$f(x) = \ln x, \qquad f'(x) = \frac{1}{x}$$
$$f(x) = \exp(x), \qquad f'(x) = \exp(x)$$
$$f(x) = \exp(\alpha x), \qquad f'(x) = \alpha \exp(x)$$

Note that the derivative is a linear operator; this means that the derivative of a sum is equal to the sum of the derivatives:

$$f(x) = h(x) + g(x), \qquad f'(x) = h'(x) + g'(x)$$

Also the derivative of a constant times a function is just the constant times the derivative of the function:

$$f(x) = \alpha g(x), \qquad f'(x) = \alpha g'(x)$$

If you have a composite function that is a product of two functions, we can use the so-called "product rule," which in words says that the derivative of the composite is the derivative of the first function times the second plus the derivative of the second function times the first:

$$f(x) = h(x)g(x), \qquad f'(x) = h'(x)g(x) + g'(x)h(x)$$

Suppose instead that you have a composite function that is the ratio (or quotient) of two functions of the same variable. To find the derivative we can use the "quotient rule," which says that the derivative of the composite is the derivative of the top function times the bottom minus the top function times the derivative of the bottom, all divided by the bottom function squared:

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

Suppose instead you have that $f(\cdot)$ is a function, $h(\cdot)$, of another function, $g(\cdot)$. To calculate the derivative of $f(\cdot)$, we use the "chain rule," which says that derivative of the function is the derivative of the outside function times the derivative of the inside function:

$$f(x) = h(g(x)), \qquad f'(x) = h'(g(x))g'(x)$$

To see this clearly, let's take two examples. In the first, suppose that $h(\cdot) = (\cdot)^\gamma$, and that $g(\cdot) = (\cdot)^\alpha$:

$$f(x) = (x^\alpha)^\gamma$$

The chain rule can be used to compute the derivative:

$$f'(x) = \gamma (x^\alpha)^{\gamma-1} \alpha x^{\alpha-1}$$

The $\gamma (x^\alpha)^{\gamma-1}$ is the "derivative of the outside" part, while the $\alpha x^{\alpha-1}$ is the "derivative of the inside" part. You can simplify this:

$$f'(x) = \alpha\gamma x^{\alpha\gamma-1}$$

If you look at this hard enough, you'll see that this is the same thing we'd have gotten had we just simplified the function first by writing it $f(x) = x^{\alpha\gamma}$. Suppose instead the function were:

$$f(x) = (1 + \alpha x)^\gamma$$

The "outside" function raises to the power γ, while the inside function is just linear in x. Using the chain rule, we get:

$$f'(x) = \gamma(1 + \alpha x)^{\gamma-1}\alpha$$

Where here the $\gamma(1 + \alpha x)^{\gamma-1}$ is the derivative of the outside part, while α is the derivative of the inside.

We can extend the basic calculus rules to multivariate functions as well. Suppose that $y = f(x, z)$. The partial derivative of y with respect to x is the change in y given a small change in x, holding z fixed. The partial derivative of y with respect to z is defined similarly. Basically, for a multivariate function, to find the partial derivatives treat all but the one variable you are differentiating with respect to as a constant, and use the same rules for differentiating univariate functions shown above. The notation we'll use is:

$$\frac{\partial y}{\partial x} = f_x(x, z)$$
$$\frac{\partial y}{\partial z} = f_z(x, z)$$

In words, the partial derivative measures the "change in y for a change in x, holding z fixed." It is thus basically the same thing as the derivative of a univariate function, we just use the notation ∂y instead of dy to denote that we are holding all other variables fixed.

Partial derivatives are easiest to think of with a couple of examples:

$$f(x, z) = \ln x + z^\alpha$$
$$f_x(x, z) = \frac{1}{x}$$
$$f_z(x, z) = \alpha z^{\alpha - 1}$$

$$f(x, z) = x^\alpha z^\gamma$$
$$f_x(x, z) = \alpha x^{\alpha - 1} z^\gamma$$
$$f_z(x, z) = \gamma x^\alpha z^{\gamma - 1}$$

All the rules for derivatives discussed above apply to the multivariate functions we've just discussed above. In particular, the derivative of a sum is the sum of the derivatives, the chain rule still holds, etc.

A useful concept that sometimes comes up is the "total derivative." The total derivative says that the total change in the left hand side variable is approximately equal to the sum of the partial derivatives times the change in the variables on the right hand side. Let dx stand for "change in x." Here we use the dx to denote that we are not necessarily holding all other variables fixed. The total derivative for a multivariate function is:

$$dy \approx f_x(x, z)dx + f_z(x, z)dz$$

For example, suppose that the function is $y = x^\alpha z^\gamma$. The total derivative says that:

$$dy \approx \alpha x^{\alpha - 1} z^\gamma dx + \gamma x^\alpha z^{\gamma - 1} dz$$

Let's use some numbers to make sense of this approximation. In particular, assume $\alpha = \gamma = 0.5$. Suppose $x = 1$ and $z = 1$ initially. This means the initial value of $y = 1$. Let's suppose that x increases to 1.2 and z increases to 1.1. Then the new value of the function is 1.1489. The total derivative would use the equation above, where the partial derivatives are evaluated at the initial points:

$$dy \approx 0.5 \times 1^{-0.5} \times 1^{0.5} \times (1.2 - 1) + 0.5 \times 1^{0.5} \times 1^{-0.5} \times (1.1 - 1) = 0.15$$

Here we can see that the total derivative gives a good approximation to the true change in the function. This approximation will be less good the (i) bigger the changes in x and z are and (ii) the more non-linear the function is.

A useful application of the total derivative is to look at the growth rate of a sum. Suppose that we have the following relationship, where the variables are indexed by time:

$$y_t = x_t + z_t$$

Totally differentiating, we get (here the total differentiation is relatively trivial, as the partials are both 1):

$$dy_t = dx_t + dz_t$$

This just says that the total change in the left hand side must equal the total change in the right hand side. Divide both sides by the previous period's value of y, y_{t-1}:

$$\frac{dy_t}{y_{t-1}} = \frac{dx_t}{y_{t-1}} + \frac{dz_t}{y_{t-1}}$$

Multiply and divide the right hand side by x_{t-1} and z_{t-1} where appropriate to get:

$$\frac{dy_t}{y_{t-1}} = \frac{x_{t-1}}{y_{t-1}} \frac{dx_t}{x_{t-1}} + \frac{z_{t-1}}{y_{t-1}} \frac{dz_t}{z_{t-1}}$$

Note that $\frac{dy_t}{y_{t-1}}$ can be interpreted as the growth rate between $t-1$ and t: the change in y, dy_t, divided by the base. Re-writing:

$$g_t^y = \frac{x_{t-1}}{y_{t-1}} g_t^x + \frac{z_{t-1}}{y_{t-1}} g_t^z$$

In other words, the growth rate of the sum is equal to the "share-weighted" sum of the growth rates, where the share weights are the previous period's ratio of x and z to y.

As written, this statement holds with equality. A useful approximation is to treat the shares as constants and use the long run averages of the shares, as opposed to the previous period's shares. For example, in a closed economy with no government spending, total GDP must equal the sum of consumption and investment spending: $y_t = c_t + I_t$. The growth rate of GDP is approximately equal to the share-weighted growth rates of consumption and investment, where we use the historical average shares.

6   Optimization

In economics we are often interested in finding optimums of functions. The optimum of a function, $f(x)$, is the value of x, x^*, at which $f(x^*)$ is either as large (the maximum) or as small (the minimum) as possible on the feasible set of values of x.

Provided certain regularity conditions are satisfied, we can characterize optima using calculus. A necessary condition for x^* to be an interior optimum of $f(x)$ is that $f'(x^*) = 0$. By "interior" I mean that I am not considering values of x that are on the "endpoints" of the feasible set of x values. This condition is what is called a first order condition. The intuition for this is straightforward – for the case of a maximum, if a function were either increasing or decreasing at x^*, then x^* could not possibly be an maximum. If $f'(x^*) > 0$, you could increase $f(x)$ by increasing x^*. If $f'(x^*) < 0$, you could increase $f(x)$ by decreasing x^*.

We refer to points at which the first order condition is satisfied as "critical points" – these are

values of x at which the derivative of $f(\cdot)$ is equal to zero. Not all critical points are "global" optima – you could have multiple points where the first order condition is satisfied, but only when represents the "global" optimum. We would refer to such points as "local" maxima and minima. For most of this course we will not be dealing with functions like this.

The first derivative being zero is necessary for either a maximum or a minimum. So how do we tell whether the critical point is a max or a min? The answer lies in looking at the second derivative. If the second derivative is negative, then the critical point is a maximum. The second derivative being negative means that, moving away from x^*, the derivative is getting smaller. Since the derivative at the critical point x^* is zero, moving away from x^* must make the value of the function smaller, since the derivative will go negative away from x^*. If the value of the function gets smaller away from x^*, then the critical point x^* must be a maximum. The converse is true for a minimum. For a critical point to be a minimum, the second derivative would be positive.

Let's work through a couple of examples. Suppose that we have a function $f(x) = x^2$. The first derivative is $f'(x) = 2x$. The first order condition only holds if $x = 0$. Is this point a maximum or a minimum? The second derivative is just the derivative of the derivative: $f''(x) = 2 > 0$. Since the second derivative is positive, this point must be a minimum. We can clearly see this by plotting the function. $x^* = 0$ is clearly a minimum.
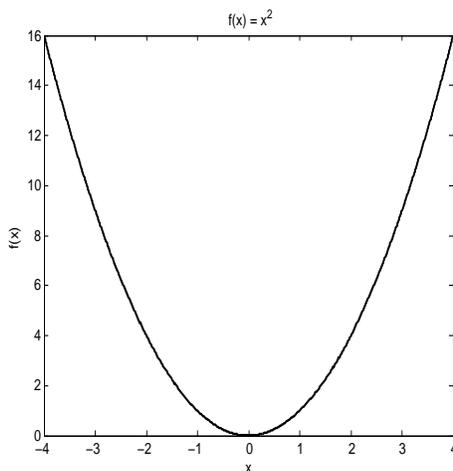
Next consider a more complicated function, $f(x) = \ln x - 2x$. The first derivative is $f'(x) = \frac{1}{x} - 2$. To find the critical point set this equal to zero and solve for x^*, which yields: $x^* = \frac{1}{2}$. The second derivative is $f''(x) = -\frac{1}{x^2} < 0$. Hence, this point represents a maximum. We can see that it is a maximum by just looking at a plot of the function:

f(x) = ln x – 2x
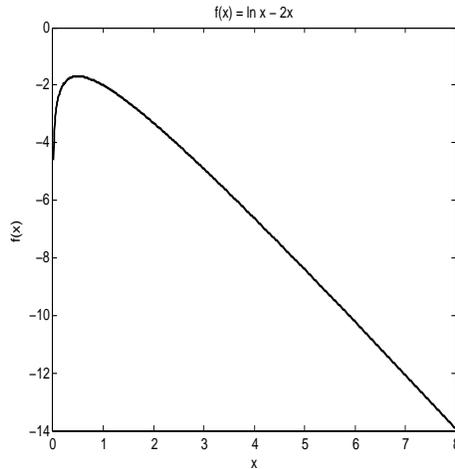
Here the maximum occurs at $x^* = \frac{1}{2}$, which is what the calculus told us.

The basic rules of optimization that we have encountered apply equally well to multivariate problems. Suppose you have a function of two variables, $f(x, z)$. The first order conditions are to set the partial derivatives with respect to both arguments equal to zero: $f_x(x, z) = 0$ and $f_z(x, z) = 0$. The second order conditions are a little more complicated, but basically get at the same point. Technically the second order conditions place restrictions on the Hessian, which is a matrix of second derivatives. We won't concern ourselves with any of that in this course.

In multivariate problems we'll often be constrained optimization – maximizing or minimizing an objective function subject to a constraint. In some sense the basic definition of economics is constrained optimization – how people maximize their well-being subject to the scarcity they face. The basic idea is that you want to maximize $f(x, z)$, subject to some constraint on the relationship between x and z. Most often in economics the constraints are linear and take the form of simple budget constraints.

Let's look at constrained optimization using an example. Suppose that x and z are goods, with prices p_x and p_y, respectively. The household has total income y, which it takes as given. The problem is to pick x and z to maximize its utility, subject to the constraint that it not spend more than its income. I assume that utility is log and additive in the two "goods" x and z:

$$\max_{x,z} \quad U = \ln x + \ln z$$

$$\text{s.t.}$$

$$p_x x + p_z z \leq y$$

Note that I've written the constraint with a weak inequality sign instead of an equal sign. In virtually every application we'll encounter in this course the constraint will hold with an equality – you wouldn't want to leave income unused, for example. There are circumstances in which the

constraint might not bind, and nothing forces consumers to spend all of their income. Therefore writing it with a weak inequality is a bit more general.

There are two ways to handle a constrained optimization problem. The method of Lagrange multipliers is one that you would likely learn in a multivariate calculus class. We won't do that here. Rather, our approach will be to (i) assume the constraint holds with equality (i.e. the constraint "binds"), and (ii) eliminate one of the variables, solving for it from the constraint and then plugging into the objective function, and (iii) finding the optimum in the normal way as if the problem were unconstrained.

Let's work this out. I'm going to eliminate z by solving for it from the budget constraint:

$$z = \frac{1}{p_z}y - \frac{p_x}{p_z}x$$

Now take this and plug it into the objective function:

$$\max_{x} \quad U = \ln x + \ln\left(\frac{1}{p_z}y - \frac{p_x}{p_z}x\right)$$

This is now just an unconstrained problem. Take the derivative with respect to x:

$$\frac{dU}{dx} = \frac{1}{x} - \frac{p_z}{y - p_x x}\frac{p_x}{p_z} = \frac{1}{x} - \frac{p_x}{y - p_x x}$$

Set this equal to zero:

$$\frac{1}{x} = \frac{p_x}{y - p_x x}$$

Now, to go further, it is helpful to note that $y - p_x x = p_z z$ from the constraint. So we can plug this back in to this first order condition to get:

$$\frac{1}{x} = \frac{p_x}{p_z}\frac{1}{z}$$

Simplifying a little bit:

$$\frac{z}{x} = \frac{p_x}{p_z}$$

If you've taken intermediate micro, this condition will look familiar – it's the "Marginal rate of substitution = price ratio" condition. The left hand side is the marginal utility of x, equal to $\frac{1}{x}$, divided by the marginal utility of z, equal to $\frac{1}{z}$. The right hand side is the price ratio. To solve for the actual allocation, take this condition, plug it into the budget constraint, and simplify:

$$z = x\frac{p_x}{p_z}$$

$$p_x x + p_z x\frac{p_x}{p_z} = y$$

$$x = \frac{y}{2p_x}$$

$$z = \frac{y}{2p_z}$$

So the optimal bundle here is to spend half of your income on x and half on y.

We will be working a lot in this course with a basic optimization problem like this. The basic strategy is to suppose that the constraint binds, plug the constraint into the objective function, and then take the first order condition. Then do some algebra with the first order condition and the constraint to back out the solution. That cookbook approach will work 99.9% of the time in this course.