Math in Econ 205

These notes review the math concepts and notation needed in Econ 205, explaining what you need to know and how to apply it. It is designed to be a non-technical, user-friendly primer and reference.

1 Exponent Rules

When an exponent is raised to another exponent, you *multiply* the exponents,

$$(3^2)^3 = 3^{2\cdot 3} = 3^6.$$

This is useful when you have an expression, say $q^{3/2} = 3k$, and you want to solve it for q. You need to raise both sides to the 2/3 power (since $\frac{2}{3} \cdot \frac{3}{2} = 1$) to get $q = (3k)^{2/3}$.

When two numbers with a common base are multiplied together, you *add* the exponents,

$$3^2 \cdot 3^3 = (3 \cdot 3) \cdot (3 \cdot 3 \cdot 3) = 3^{2+3} = 3^5.$$

2 Functions

What are they? A *function* is a relationship defining exactly one output for each set of inputs. The inputs are also called *arguments*. They are usually noted as a letter with arguments in parentheses, like $C(w, v, q)$.

What do I need to know? You need to be comfortable working with generic functions, for example $C(w, v, q)$ instead of a specific example like $C(w, v, q) = q^2wv$.

Two other small points. First, note that $f(x, y)$ can refer to both the function f, which gives the relationship between x, y, and the output of f obtained by plugging in the specific values x and y. You can tell between them based on whether the inputs x and y are variables or known values. Second, functions only depend on their arguments. The function $f(x, y)$ does not depend on z, so a change in z will not affect it (assuming that z does not affect x or y).

How will we use them? We will use functions for anything and everything because they describe the answers to economic questions. You can think of any function as answering the question, what is the output given the inputs? For example, a cost function $C(w, q)$ tells us how much it costs a firm to produce q units when the wage is w.

We will also use functions to indicate dependence. Perhaps we have some some variable x, but we learn from our model that x can be determined from p_x and p_y. In that case, we might write x as a function of p_x and p_y, $x(p_x, p_y)$, to denote that dependence. This means that we can learn everything we need to know about the value of x from the values of p_x and p_y.

3 The Derivative

What is it? The instantaneous rate of change of a function, which you can think of as the slope of a function at a single point. The derivative of the function f at a point x is noted as either $\frac{df(x)}{dx}$ or $f'(x)$.

What do I need to know? You need to be very comfortable taking and interpreting derivatives. Common derivative rules are shown in Table 1.

How will we use them? Anywhere and everywhere. The derivative is one of our most basic methods of analysis for two reasons. First, they are necessary for maximizing functions (covered later). Second, once we have a function answering an economic question, the derivative tells us how the result changes when we change an input, like finding the marginal cost from a cost function.

$f(x)$	$f'(x)$	Example
x^b	bx^{b-1}	$\frac{d}{dx}6x^{2/3} = 6(\frac{2}{3}x^{-1/3}) = 4x^{-1/3}$
$\ln(x)$	$1/x$	
$f(g(x))$	$f'(g(x))g'(x)$	$\frac{d}{dx}2\ln(3x^2) = \frac{2}{3x^2}(6x) = \frac{4}{x}$

Table 1: Formulas for common derivatives: the power rule (row 1), derivative of natural logarithm (row 2), and chain rule (row 3).

A note on interpretation. Since the derivative tells us how a function changes, it's easy to interpret it as the change in the output when the input changes by one unit. For example, we discuss marginal cost as the increase in cost needed to produce one more unit. This is not exactly true for a one-unit change in the input. It is only accurate for very small changes in the input.

For instance, the derivative of x^2 is $2x$, so the rate of change at $x = 3$ is 6. But $4^2 - 3^2 = 7 > 6$, a change greater than the one given by the derivative. It is not that the derivative is wrong, but that a one-unit change is too large. As we move away from $x = 3$, the rate of change of the function changes and the derivative no longer approximates the change accurately. We use the one-unit change language because it is easy to interpret, but the more precise version is that if the input x increases by a very small amount dx, the function increases by the amount of the increase dx times the derivative, $\frac{df}{dx}\,dx$.

3.1 The Partial Derivative

What is it? A derivative for functions of several variables. It only considers the instantaneous rate of change with respect to one variable, holding all others constant. Constant means that other variables are treated like the number 5. It is noted with curly d's, $\frac{\partial f(x,y)}{\partial x}$.

The partial derivative of the cost function $C(q, w)$ with respect to w, noted $\frac{\partial C}{\partial w}$, would be interpreted as the change in the cost of producing the quantity q (held constant!) when the wage increases slightly.

What do I need to know? As before, how to take and interpret them. If you have a function $f(x, y)$ that depends on x and y, you can find the partial with respect to x by treating all appearances of y as constants (e.g. as if they were the number 5) and taking the derivative.

How will we use them? The same way as normal derivatives. They are exceptionally important because most of our functions will be of several variables. For example, utility functions depend on consumption of several goods.

Example 1. Find all of the partial derivatives of $f(x, y, z) = x^2y^2 \ln(z) + 2xy + 2y + 3$.

Start with $\frac{\partial f}{\partial x}$. In the partial with respect to x, all appearancs of y and z are treated like constants, so the function is the same as $f(x) = ax^2 + bx + c$, where $a = y^2 \ln(z)$, $b = 2y$, and $c = 2y + 3$. Thus

$$\frac{\partial f}{\partial x} = 2x(y^2 \ln(z)) + 2y + 0 + 0 = 2xy^2 \ln(z) + 2y.$$

Similarly for y and z,

$$\frac{\partial f}{\partial y} = 2y(x^2 \ln(z)) + 2x + 2 + 0$$
$$\frac{\partial f}{\partial z} = \frac{x^2y^2}{z} + 0 + 0 + 0.$$

3.2 The Total Derivative

What is it? The instantaneous change in a function $f(x_1, \ldots, x_k)$ when *all* of its inputs change by a small amount. This is different from the partial derivative because all inputs—not just one—are changing, but to find it you just need to take a bunch of partial derivatives. It is defined as

$$df = \frac{\partial f}{\partial x_1}\, dx_1 + \cdots + \frac{\partial f}{\partial x_k}\, dx_k.$$

Essentially, multiply each partial derivative by the change in the relevant variable dx_i (which is usually one) and add them together.

What do I need to know? What it is, how to take it, and how it applies in cases with indirect effects (see below).

How will we use it? To find how indirect effects change a function, and to find the tradeoff between input variables that leaves the output unchanged. For example, how large does dx_1 have to be relative to dx_2 to leave the output unchanged? (See the second example below.)

Example 2. Consider the utility function $U(x, y) = xy + 2y$. Find the total derivative.

Start with the partials.

$$\frac{\partial U}{\partial x} = y + 0$$
$$\frac{\partial U}{\partial y} = x + 2.$$

Now assemble them into a total derivative.

$$\begin{aligned} dU &= y(1) + (x+2)(1) \\ &= y + x + 2. \end{aligned}$$

In the example above, we found the change in U when x and y both change by the same small amount, dx and dy. In one case, illustrated in the example below, we leave dx and dy instead of replacing them with one. The reason is so that we can find the relative size of dx and dy needed to make the change in the function dU equal to zero. We will use this trick in class several times.

Example 3. What ratio of x to y leaves utility constant (e.g. what trade of x for y leaves us on the same indifference curve) in $U(x, y) = xy + 2y$?

From the example above, we know the total derivative. The difference is that we want to leave in dx and dy rather than replacing them with one (which we do so that they can have different relative sizes) and set dU equal to zero. This gives

$$\begin{aligned} dU &= y\,dx + (x+2)\,dy \\ 0 &= y\,dx + (x+2)\,dy \\ dy &= \frac{-y}{x+2}\,dx \\ \frac{dy}{dx} &= \frac{-y}{x+2}. \end{aligned}$$

The answer can be interpreted as: at point (x, y), the consumer's utility is constant if he trades one unit of x for $\frac{y}{x+2}$ units of y. (The "one-unit" change interpretation is subject to the caveat in the section on the derivative.)

Indirect Effects. We use the total derivative to analyze indirect effects, where several arguments to a function depend on the same underlying parameter. For example, consider $f(x(a), y(a))$, where both the inputs x and y are functions of a. A change in a affects both $x(a)$ and $y(a)$, and the changes in $x(a)$ and $y(a)$ both affect the output, requiring a total derivative:

$$\frac{\partial f}{\partial a} = \frac{\partial f}{\partial x}x'(a) + \frac{\partial f}{\partial y}y'(a).$$

Note that the dx term here is now replaced by $\frac{dx}{da}$, the change in x when we change a. This is just an application of the chain rule.

Example 4. A consumer has utility $U(x, y)$ and purchases the bundle $(x(p_x, p_y), y(p_x, p_y))$ when prices are (p_x, p_y). How does utility respond to a change in p_x?

The price does not directly affect utility, but it affects the consumer's choice of x and y (which are functions of price) and hence affects utility indirectly. Using the rule outlined above,

$$\frac{\partial U}{\partial p_x} = \frac{\partial U(x,y)}{\partial x}\frac{\partial x(p_x,p_y)}{\partial p_x} + \frac{\partial U(x,y)}{\partial y}\frac{\partial y(p_x,p_y)}{\partial p_x}.$$

4 Integrals

What are they? For our purposes, the integral is the area between a function and the axis.

What do I need to know? You only need to know two things. First, the integral of a function from a to b equals the area under the curve (between the curve and the axis) from a to b. Second,

$$\int_a^b f'(x)\,dx = f(b) - f(a).$$

This second result relates the integral to the derivative. In the few cases where the lecture notes mention the fundamental theorem of calculus, this is the result it's referring to. The crux of the statement is that, to integrate, you need to find the *antiderivative* of the function you're integrating, then take its value at the upper limit of the integral minus its value at the lower limit. The antiderivative is just a function whose derivative is the original function: the antiderivative of $f(x)$ has the derivative $f(x)$. The only antiderivatives we'll use are in Table 2.

$f'(x)$	$f(x)$
x^b	$\frac{1}{b}x^{b+1}$
$1/x$	$\ln(x)$

Table 2: Common antiderivatives.

Example 5. Calculate $\int_4^5 10 - p\,dp$.

$$\begin{aligned}\int_2^4 10 - p\,dp &= 10p - \frac{1}{2}p^2\Big|_2^4 \\ &= (10\cdot 4 - \frac{1}{2}\cdot 16) - (10\cdot 2 - \frac{1}{2}\cdot 4) = 14.\end{aligned}$$

How will we use them? To find the area under the curve in a few cases. In general, you'll be given straight lines so that you can find the area of a triangle rather than having to integrate.

5 Optimization and the Lagrangian

Maximizing and minimizing functions are essential parts of the course because we want to describe agents who maximize utility and firms that minimize production costs. There are two types of optimization problems:

1. *Unconstrained* optimization problems, which place no limits on what values the variables maximized over can take.

2. *Constrained* optimization problems, which force the values of the variables maximized over to satisfy certain conditions, like $x + y = 10$ or $U(x, y) = 5$.

5.1 Notation

The mathematical notation for "maximize this function" is

$$\max_{x,y} f(x, y),$$

where the function is maximized by choice of x and y (the variables in the subscript of max) and $f(x, y)$ is the function to be maximized. This expression includes no constraint. If there is a constraint, it is added on after a "subject to" statement,

$$\max_{x,y} f(x, y) \text{ subject to } x + y \leq 10,$$

where $x+y \leq 10$ is an example of a constraint. To describe a minimization problem, replace max with min.

The maximizing value of a variable x is usually referred to as x^*, denoting the specific value of x that solves the problem.

5.2 Unconstrained Optimization

How do I do it? Follow the two steps below for the generic problem $\max_{x,y} f(x, y)$.

1. Take partial derivatives of $f(x, y)$ with respect to each variable we're maximizing over and set them equal to zero. These are called first-order conditions (FOCs).

    - When setting equal to zero, replace the choice variables with their starred version (x^* instead of x). This is because the partial derivatives equal zero at the maximizing values x^* and y^*, but not at generic values of x and y.

2. Solve the system. We want expressions for the variables we're maximizing over that do not depend on each other—the solution for x^* cannot include y^*.

Example 6. Maximize $f(x, y, t) = t(x + y) - x^2 - y^2$ by choice of x and y.

The function also depends on t, but we only care about choosing x and y. Our optimal x and y can thus depend on t. The problem is

$$\max_{x,y} txy - x^2 - y^2.$$

For step 1, take the partials and set them equal to zero. Note that when we set them equal to zero, we are using the maximizing values x^* and y^*.

$$\frac{\partial f}{\partial x} = t - 2x^* = 0$$
$$\frac{\partial f}{\partial y} = t - 2y^* = 0.$$

For step 2, solve for x^* and y^* in terms of t and constants. The variable t can be part of the solution because we are not maximizing by choice of t.

$$x^* = t/2$$
$$y^* = t/2.$$

The example above skips some details. For instance, are we sure that x^* and y^* are maximizers, not minimizers? To be sure, we would check the matrix of second derivatives. In this class, we will avoid second derivative matrices, although we will occasionally mention second derivatives when they are scalars.

What's the intuition? Start with a function of one variable, $f(x)$. If the derivative is positive at x, $f'(x) > 0$, then we can increase the function a bit by increasing x. If it is negative, $f'(x) < 0$, then we can increase the function a bit by decreasing x. In either case, the function can be increased or reduced by changing x. The only way to have a maximizer (or minimizer) is to have $f'(x) = 0$.

The intuition generalizes nicely with several variables. Now, if the partial derivative with respect to x is positive, we can still increase the function by increasing x, interpreted as a move in the x direction holding all other variables constant. Thus, we need the partial derivatives with respect to all variables to equal zero, or else we could increase the function by adjusting one variable.

To make sure that it's a maximum, we need the derivatives to be decreasing (a negative second derivative), or some similar condition on a matrix of second derivatives. If they are not, then the derivative will become positive again and the function will increase more. (Analogously, a minimum requires increasing derivatives, a positive second derivative.)

5.3 Constrained Optimization: Lagrangians

How do I do it? Make sure you know what your constraint is and follow these steps, outlining the method of Lagrange multipliers.

1. Rewrite the constraint so that one side is zero, $g(x, y) = 0$.

2. Form a function called the *Lagrangian* as

$$\mathcal{L} = (\text{Function to maximize}) + \lambda(\text{constraint set equal to zero}).$$

For example, to maximize $f(x,y)$ subject to $g(x,y)=0$,

$$\mathcal{L} = f(x,y) + \lambda g(x,y).$$

3. Take the partial derivatives of $\mathcal{L}$ with respect to x, y, and λ and set them equal to zero. (The derivative with respect to λ is always just $g(x,y)$.) Again, replace x, y, and λ with x^*, y^*, and λ^*.

4. Solve the system for x^*, y^*, and λ^*.

   - In general, solve the partial derivatives for λ and set them equal to each other,

$$\begin{aligned} \frac{\partial f}{\partial x} + \lambda \frac{\partial g}{\partial x} &= 0 \\ \Rightarrow \lambda &= -\frac{\partial f}{\partial x} / \frac{\partial g}{\partial x} \\ \Rightarrow -\frac{\partial f}{\partial x} / \frac{\partial g}{\partial x} = \lambda &= -\frac{\partial f}{\partial y} / \frac{\partial g}{\partial y} \end{aligned}$$

   - Solve the resulting expression for x^* in terms of y^* (or vice versa) and substitute into the budget constraint,

$$g(x^*(y^*), y^*) = 0.$$

   You can solve this expression for y^* because it only depends on y^* and constants.

   - Get x^* from the expression for x^* in terms of y^*, $x^*(y^*)$.

Example 7. A consumer maximizes utility $U(x,y) = x^2y$. The price of x is $p_x = 3$, the price of y is $p_y = 1$, and the consumer has income I to spend. Find the consumer's utility-maximizing bundle (x^*, y^*).

Step 1: rewrite the constraint. The consumer can't spend more than I, so $3x + y = I$. The constraint is then $I - 3x - y$.

Step 2: form the Lagrangian,

$$\mathcal{L} = x^2y + \lambda(I - 3x - y).$$

Step 3: take first-order conditions and solve. The conditions are

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial x} &= 2x^*y^* - 3\lambda^* = 0 \\ \frac{\partial \mathcal{L}}{\partial y} &= (x^*)^2 - \lambda^* = 0 \\ g(x,y) = \frac{\partial \mathcal{L}}{\partial \lambda} &= I - 3x^* - y^* = 0. \end{aligned}$$

Now, use the first two partials to solve for x^* in terms of y^*,

$$\frac{2}{3}x^*y^* = \lambda = (x^*)^2$$
$$\frac{2}{3}y^* = x^*.$$

Substitute back into the constraint and solve for all variables.

$$0 = I - 3(\frac{2}{3}y^*) - y^*$$
$$y^* = \frac{I}{2}$$
$$x^* = \frac{2}{3}y^* = \frac{I}{3}.$$

The consumer thus purchases $I/2$ units of y and $I/3$ units of x. We usually do not care about the value of λ^*, but we can solve for it. In this example, it is $\frac{I^2}{9}$.

Intuition. You can get through this class without caring how Lagrange's method works, but there's not too much mystery. With no constraint, we wanted to exhaust all gains from changing a variable—the partial derivatives had to equal zero. Now, we might not be able to exhaust all gains. In the example above, utility has no unconstrained maximum; it always gets bigger when x and y get bigger.

This means the partial derivatives will be positive at the constrained maximum. We just need to make sure that we can't increase utility by trading some x for some y. To judge whether that trade would increase the function, we need to consider both how much the function would change and how much the constraint would change. For instance, more x might cause the function to increase a lot, but it wouldn't be worthwhile if it tightened the constraint a lot too.

The constrained maximum thus requires all variables to make the same marginal contribution to the function after we account for their effects on the constraint. This is exactly the expression derived above, $\frac{\partial f}{\partial x}/\frac{\partial g}{\partial x} = \frac{\partial f}{\partial y}/\frac{\partial g}{\partial y}$.

What's with the λ? The multiplier λ can be interpreted as the change in the function to be maximized if the constraint were loosened slightly. In our example, λ answers the question: by how much would utility increase if we gave the consumer a little bit more income?

To see why this is the interpretation, recall that

$$-\frac{\partial f}{\partial x}/\frac{\partial g}{\partial x} = \lambda^*.$$

If the constraint were loosened slightly, we could increase x by up to $1/\frac{\partial g}{\partial x}$. The function f would increase by $\frac{\partial f}{\partial x}$ for each unit change in x, producing a total change of $\frac{\partial f}{\partial x}/\frac{\partial g}{\partial x}$, exactly the value of λ^* above. (The explanation misses a -1, but this is the result of how we write the constraint. The constraint could be $I - p_x x - p_y y$ or $p_x x + p_y y - I$, which would change the sign of $\frac{\partial g}{\partial x}$.)

6 Homogeneity

What is it? A property of functions. A function $f(x_1, \ldots, x_n)$ is *homogeneous of degree* k if for any number t

$$f(t \cdot x_1, \ldots, t \cdot x_n) = t^k f(x_1, \ldots, x_n).$$

This means: if we multiply all of the inputs by some amount t, then the output is multiplied by some amount t^k. Some functions are not homogeneous of any degree.

What do I need to know? You need to be able to tell if a function is homogeneous of degree 0 or 1. If you double the inputs to a function and the output does not change (e.g. it changes by $2^0 = 1$), it is homogeneous of degree 0. If the output doubles ($2^1 = 2$), it is homogeneous of degree 1.

You can check if a function is homogeneous and, if so, of what degree by

1. Multiplying all arguments of the function by a constant t.
2. Trying to factor the output into the original function $f(x, y)$ without any t terms and a single t term.
    - If you can factor it this way and the single t term is raised to some power k, then the function is homogeneous of degree k.
    - If you can't, the function is not homogeneous.

Example 8. Find if the following functions are homogeneous and, if so, of what degree.

1. $f(x, y) = \sqrt{xy}$.

    Multiplying x and y by t gives

    $$f(tx, ty) = \sqrt{(tx)(ty)} = \sqrt{t^2 xy} = t\sqrt{xy} = tf(x, y).$$

    The function is homogeneous of degree 1.

2. $f(x, y) = \frac{3x}{y}$.

    $$f(tx, ty) = \frac{3(tx)}{(ty)} = \frac{t}{t}\frac{3x}{y} = t^0 f(x, y).$$

    The function is homogeneous of degree 0.

3. $f(x, y) = xy + \sqrt{y}$.

    $$f(tx, ty) = t^2 xy + \sqrt{ty}.$$

    It is impossible to pull out a single t^k term multiplied by $xy + \sqrt{y}$, so the function is not homogeneous of any degree.

How will we use it? Homogeneity will show up as a property of functions. We'll prove that the functions are homogeneous and you should be able to interpret what homogeneity means in economic terms.

We care about homogeneity of degrees 0 and 1 because they tell us how economic outcomes responds to scaling some feature of the environment. For example, does the best bundle you can purchase change if all prices and income double?

7 Expected Value

What is it? A way to describe a representative outcome when different outcomes are possible, used when outcomes are random. If $x_1, \dots, x_k$ are some outcomes and outcome x_i occurs with probability α_i, the *expected value* EV is

$$\mathrm{EV} = \alpha_1 x_1 + \cdots + \alpha_k x_k.$$

It just weights each outcome by the probability that it occurs and adds them up. It is interpreted as the average result you would get if you could repeat the random process many times. For example, you might buy a lottery ticket that pays either \$0 or \$100, but the expected value would describe its average payoff.

What do I need to know? How to calculate and interpret it.

Example 9. Suppose that a lottery ticket pays \$0 with probability .5, \$20 with probability .25, and \$100 with probability .25. What is the expected value of the lottery ticket?

$$\mathrm{EV} = 0 \cdot (.5) + 20 \cdot (.25) + 100 \cdot (.25) = 5 + 25 = 30.$$

How will we use it? To study decision making under risk and uncertainty.

8 Other Topics

This section discusses minor points and topics that might be useful, but are not essential.

Notation for Constants. Sometimes, we want to work with an unknown but fixed level of one variable. This is usually noted with a bar or a subscript zero, e.g. $\bar{k}$ or k_0.

Strong versus Weak. The lecture notes occasionally have a "strong" and "weak" definition for a concept. This applies when there are inequalities. Weak versions usually allow equality, $\geq$ ("weak" inequalities), whereas strong ones, $>$ ("strict" inequalities), do not.

The Envelope Theorem. The theorem is an advanced topic that shows up in the lecture notes. I will refer to it when introducing a few findings, but will not prove it or discuss it in depth. Interested students can consult the lecture notes.