Appendix 1: Calculus Cookbook

This appendix provides a very fast review of those parts of calculus that are used in the textbook. It is *not* a substitute for a real course in the subject and should not be regarded as such. But it may serve to remind you of those parts of calculus that are required to read the text.

Why Calculus?

Calculus is one of the fundamental tools of mathematical analysis. It comes in two closely related pieces, differential calculus and integral calculus.

- *Differential calculus* provides answers to questions like, Given the position of car traveling along a highway as a function of time, what is the speed that the car is traveling, also as a function of time? It takes a function and tells you the *rate of change* of the function.
- *Integral calculus* provides answers to questions like, Given the speed of a car as a function of time, what is its position as a function of time? It takes a rate of change and tells you the value of the function (sort of).

Engineers, scientists, and motorists are all interested in these types of questions, so calculus is of interest to them. Economists find calculus interesting and useful because of its close tie to marginal analysis and optimization. Economic models of consumers and firms typically assume that the individual consumer or firm acts in a purposeful fashion, maximizing some numerical objective function. For firms, this objective function typically is profit; for consumers, it is something called utility. Calculus comes in handy because it gives us the ability, in a mathematical model, to express the rate of increase or decrease in this objective function, as marginal changes are made in production quantities by firms or in amounts of goods consumed by consumers. It allows us to find levels of production and consumption that maximize profit and utility.

Calculus is not necessary for the study of economics, especially in this age of spreadsheets and Solver. Moreover, because it is scary to some people, it can be pedagogically counterproductive. But calculus is often more efficient than searching, and it gives a convenient and simple way to express simple ideas. For this reason, you should conquer any fears you may have and use calculus; it is not hard and it is awfully useful.

A1.1. The Derivative

In calculus, the key concept is that of the *derivative* of a function. Imagine a function f that associates to every number x another number $f(x)$, in the way that functions do. To carry around a concrete example, I use the function

$$f(x) = x^2 - x + 2.$$

In the usual fashion, we can graph the function f. In Figure A1.1, I graph four functions: In panel a, I graph $f(x) = x^2 - x + 2$; the other three panels show functions for which formulas do not exist. Those other three functions are there to illustrate what functions might look like; in particular, focus on the behavior of these functions around the value $x = 2$. The function in panel b is discontinuous at $x = 2$ (it jumps there); the function in panel c is continuous but kinked at $x = 2$; the function in panel d is both continuous and smooth (neither jumps nor kinks). (If you are not quite sure what I mean by a kink, wait a bit; it becomes clearer.)

For any function f, I can ask, *What is the rate of change in the function over some discrete interval?* For example, what is the rate of change in the function $f(x) = x^2 - x + 2$, over the interval from 2 to 4? Since we have a formula for this function, this is easy to answer: It is the proverbial "rise over the run," or

$$\frac{f(4) - f(2)}{4 - 2} = \frac{(4^2 - 4 + 2) - (2^2 - 2 + 2)}{4 - 2} = \frac{14 - 4}{2} = 5.$$

Over this interval, the function increases at a rate of 5 units per 1 unit increase in the variable. Over the interval from 0 to 2, the function increases at a rate $(4 - 2)/(2 - 0) = 1$. Over the interval from 2 to 2.5, it increases at a rate

$$\frac{f(2.5) - f(2)}{2.5 - 2} = \frac{5.75 - 4}{2.5 - 2} = \frac{1.75}{0.5} = 3.5.$$

Over the interval from 2 to 2.1, it increases at a rate

$$\frac{f(2.1) - f(2)}{2.1 - 2} = \frac{4.31 - 4}{2.1 - 2} = \frac{0.31}{0.1} = 3.1.$$

Over the interval from 1.99 to 2, it increases at the rate

$$\frac{f(2) - f(1.99)}{2 - 1.99} = \frac{4 - 3.9701}{2 - 1.99} = \frac{0.0299}{0.01} = 2.99.$$

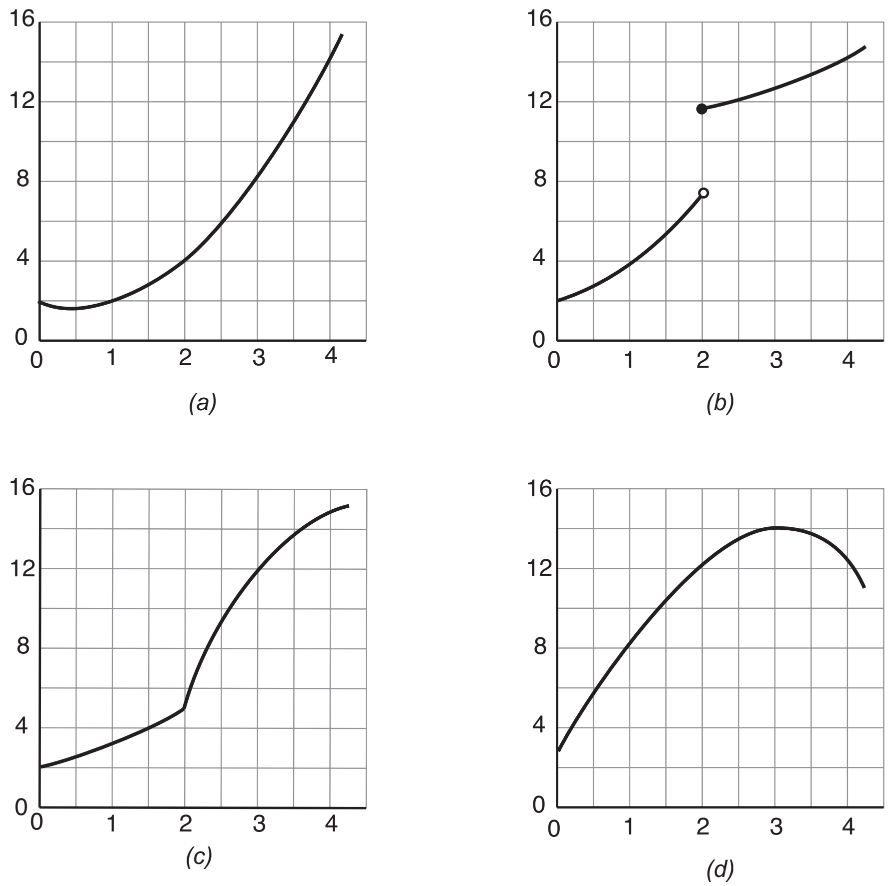


Figure A1.1. Graphs of four functions.

These are all *rates of change* in the function over some discrete interval. In all the examples, the function increases over the interval in question, so the answer is always a positive number. But, if we consider the interval from 0 to 0.5, we get a negative rate of change, because the function decreases over this interval.[1]

Suppose I ask instead for the "instantaneous rate of change" of the function f at the point 2. The function is rising at 2, so the rate of change at 2 is a positive number. We find the instantaneous rate of change at 2 by taking smaller and smaller discrete intervals starting or ending at 2 or simply bracketing 2, measuring the rates of increase over those discrete intervals, and seeing what is the limit as the intervals get smaller. We actually did this

[1] To be precise, we get $[f(0.5) - f(0)]/(0.5 - 0) = (1.75 - 2)/(0.5 - 0) = -0.25/0.5 = -0.5$.

to some extent in the examples:

1. Over the interval from 2 to 4, the rate of increase was 5.
2. Over the interval from 2 to 2.5, the rate of increase was 3.5.
3. Over the interval from 2 to 2.1, the rate of increase was 3.1.
4. Over the interval from 1.99 to 2, the rate of increase was 2.99.

The trend is clear. As the intervals get smaller, the rates of increase get closer and closer to 3.

Alternatively, lay a ruler along the function f at the value 2 and measure the slope of the *tangent line*. This is illustrated for you in Figure A1.2. The slope of this tangent line is 3, to the limits of our measurement abilities.

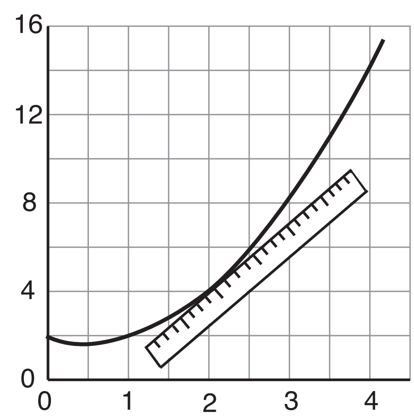
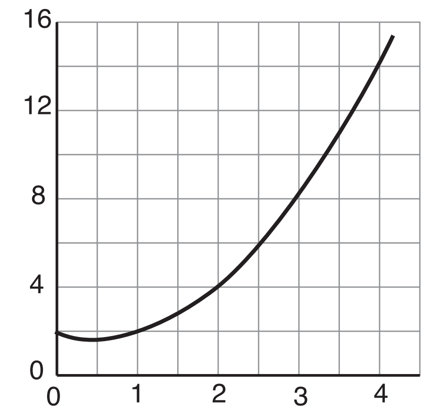


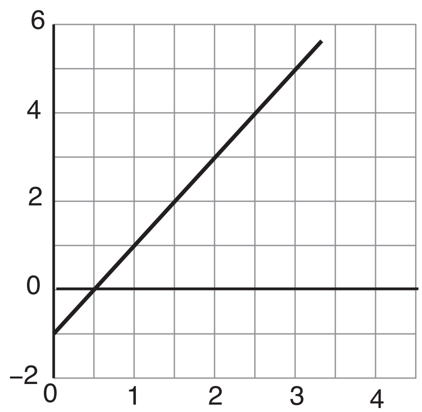
Figure A1.2. Measuring the slope of a function at a point. To find the slope or instantaneous rate of change of a function at a point, you place a ruler tangent to the function at that point and calculate the slope of the line described by the ruler's edge.

This instantaneous rate of increase (or decrease) of the function is the *derivative* of the function. It is defined formally, mathematically, as the limit of the rates of increase (or decrease) over smaller and smaller discrete intervals encompassing the point at which you wish to find the derivative. Do not worry too much about that; it is just the slope of the tangent line to the function at the point you are interested in.

Note that the derivative changes as we change the point we are looking at. In the function $f(x) = x^2 - x + 2$, the tangent line has slope 3 at $x = 2$; but at $x = 3$ the tangent line has slope 5 (get your ruler out if you do not believe me); at $x = 0$, the slope is -1 (the function is decreasing); and at $x = 0.5$,



(a) The function $f(x) = x^2 - x + 2$



(b) The derivative of the function f

Figure A1.3. The function and its derivative. Panel a shows the function $f(x) = x^2 - x + 2$, and panel b shows its derivative.

the slope is 0. In Figure A1.3, I graph the function f in panel a and directly below it I graph its derivative function.

Note that the derivative is negative for $x < 0.5$, the function f decreases below $x = 0.5$, but the derivative is getting closer and closer to 0, the function is decreasing more and more slowly. The derivative is 0 at $x = 0.5$ (the function has bottomed out), then the derivative is positive and increasingly so for $x > 0.5$ (the function now increases and at an increasing rate). We see why I drew the derivative as a linear function in a bit.

Not every function has a derivative at every point. If the function jumps discountinuously at a point, it cannot have a derivative there. And, even if it is continuous, if its slope changes at a point, as in Figure A1.1(c) at $x = 2$, then the function has no derivative there. Note that the function in Figure A1.1(c) has an identifiable rate of increase just to the left of $x = 2$ and a different rate of increase just to the right, but since these rates are not the same, the function kinks there: its slope changes discontinuously; it has no derivative at $x = 2$.

The function in Figure A1.1(c) does have nice slopes at all points except for $x = 2$. In other words, this function has a derivative everywhere but at $x = 2$. In Figure A1.4(a), I graph the function from Figure A1.1(c) again, and below it, in Figure A1.4(b), I graph its derivative function, where the jump at $x = 2$ indicates that the function has no derivative at that one point.[2] In fancy math-talk, we say that the function from Figure A1.1(c) is *not differentiable* at $x = 2$, but it is differentiable everywhere else. In less fancy talk, we say that this function is *kinked* at $x = 2$, but is *smooth* everywhere else.

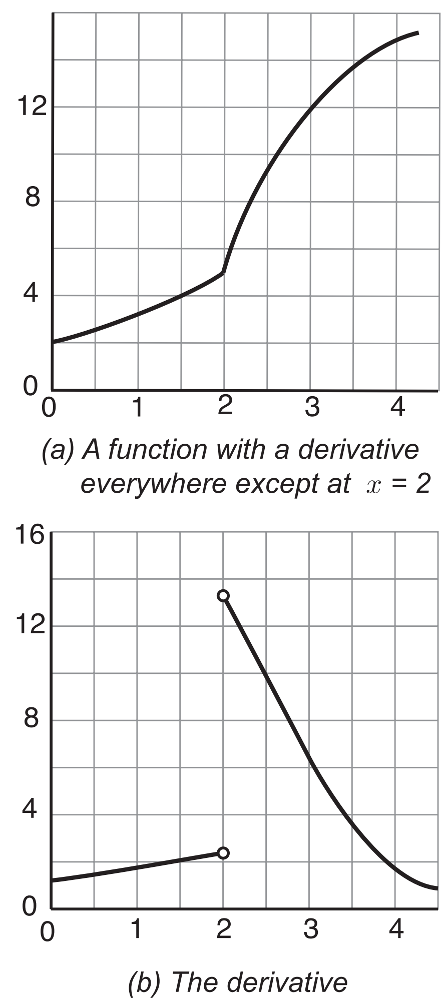


Figure A1.4. A function with a derivative except at $x = 2$, and that derivative.

[2] This drawing has the right shape, but may be off a bit in terms of values.

It should be clear that the shape of a function and that of its derivative are closely related. Where the function is increasing, its derivative is positive. If the function is decreasing (and decreasing at an increasing rate), its derivative is negative and is decreasing. In Figure A1.5, panels a, b, and c, I draw three functions. In panels x, y, and z, I draw three more functions, which are meant to be the derivatives of the first three functions, where the functions have derivatives. Can you match the functions to their derivatives? (The answer is at the end of this appendix.)

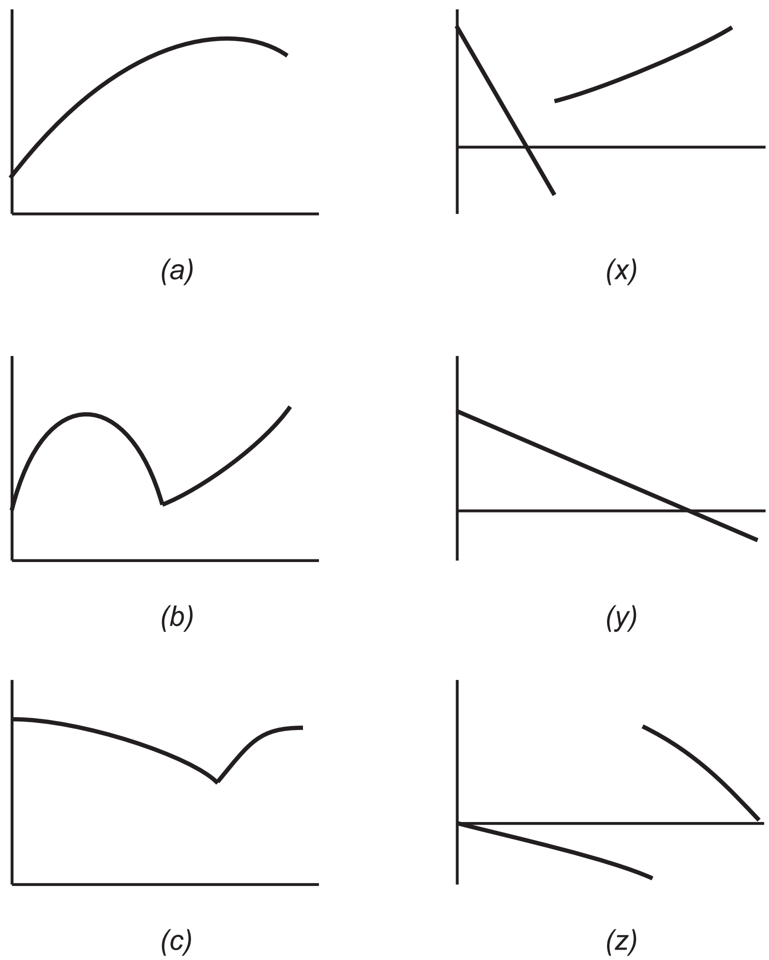


Figure A1.5. Three functions and three derivatives. Match each function in the left-hand column with its derivative (from the right-hand column). The answers are given at the end of the appendix.

For a function f that has a derivative, we write both f' and $\frac{\mathrm{d}f}{\mathrm{d}x}$ for that derivative. The f' notation is quite convenient in that it allows us to write $f'(x)$ to denote the derivative at the value x. When we want to write the value of derivative function at x using the $\frac{\mathrm{d}f}{\mathrm{d}x}$ symbology for the derivative, we write $\frac{\mathrm{d}f}{\mathrm{d}x}|_x$. But we will try to avoid this double use of the letter x.

A1.2. Maximization and Minimization

Suppose we want to find where a function f is maximized or minimized. Since the function cannot be at either a maximum or a minimum where it is either rising (has positive slope) or falling (has negative slope), if the function has a derivative at its max or min, that derivative must be 0. In other words, *a **necessary** condition for a max or a min, if the function has a derivative there, is that the derivative is 0.* Turning this around, if you have a function f that has derivatives at all points, to find its maxes and mins, you look at places where its derivative is 0.

This procedure is not free of problems.

- If you are looking for the point where a function attains its maximum or minimum over an interval—suppose, for instance, you want to know where the function $f(x) = x^2 - 4x - 5$ is maximized for x between 0 and 10—the endpoints of the interval could be maxes or mins without having a zero derivative. This function has derivative $f'(x) = 2x - 4$ (trust me for now that this is so), and it is true that the one point where the derivative is zero, namely $x = 2$, is a minimum. But $x = 10$ is a maximum. Think of it this way: The derivative at $x = 10$ is 16, so the function is going up there, and you could increase the value of the function if you could go beyond $x = 10$, to (say) $x = 10.1$. But you are *constrained* to stay inside the interval from 0 to 10. A mathematician would say that the italicized statement is true for *interior* points—points where you are unconstrained to move in either direction. A lot of economics deals with optimization subject to constraints, and the text spends considerable time describing how we generalize the italicized statement to such circumstances. (This is the main subject of Part III of the book, employing what is euphemistically known as the *bang for the buck*.)
- Just because the derivative is 0 does not mean we are at a max or a min. For example, the function $f(x) = x^3$ has derivative $f'(x) = 3x^2$. (Again, trust me.) This derivative is 0 at $x = 0$. But the function $f(x) = x^3$ is neither maximized nor minimized at $x = 0$, it is rising for all values of x. The derivative of 0 at $x = 0$ simply means (for this function) that the instantaneous rate of change in the function is 0 at $x = 0$; it is momentarily flat at $x = 0$. But the function has positive slope everywhere else (the function $f'(x) = 3x^2$ is strictly positive for x different from 0), and so the function never achieves either max or min. (A mathematician would say, "That's why, in the italicized statement, we said that a zero derivative is *necessary*. No one said it was sufficient.")
- Also, the necessity of a zero derivative at a max or min depends on the

function having a derivative there. For functions with kinks, you have to worry about the kink. This can be difficult, but there are easy cases; for example, suppose that a function f is kinked at the value x_0 but has a derivative everywhere else, moreover, $f'(x) > 0$ for $x < x_0$ and $f'(x) < 0$ for $x > x_0$. Because $f'(x) > 0$ for $x < x_0$, we know the function is increasing for $x < x_0$. And because $f'(x) < 0$ for $x > x_0$, we know it is falling for $x > x_0$. As long as f is continuous (it does not jump itself), this means it must achieve *a* maximum at x_0 precisely.

- And, finally, we must distinguish between global and local maxima and minima. Derivatives tell you, essentially, whether the function is going up or down or is flat for small changes in the variable; that is, locally. A function might achieve a local maximum at a point x, but it could get still higher some distance away from x. Think, for instance, of getting to the peak of the second highest hill in a chain of hills.

Second-Order Conditions

Suppose the function f has derivatives everywhere, and at a point x that is not up against any boundaries, $f'(x) = 0$. Are you at a (local) max or a min (or something else)? One way to tell is to look at what are called the *second-order conditions*. These involve the second derivative of f, denoted f'', which is just the derivative of the derivative. You need not know about these things to read the text, but it is a shame not to indicate how simple they really are.

Suppose you are looking at a function f and a point x where $f'(x) = 0$ and $f''(x) < 0$.[3] Then f achieves a local maximum at x. Why? If $f''(x_0) < 0$, this means that f' is decreasing at x. If $f'(x) = 0$, this (in addition) means that $f'(x') > 0$ for x' a bit less than x, so f rises as we approach x from below; and $f'(x') < 0$ for x' a bit more than x, which means that f falls as we go beyond x. This is just what it means for x to be a local maximum.[4]

A1.3. Derivatives of Some Functions

In many cases, the derivatives of functions that are expressed algebraically can also be expressed algebraically. Three important and useful examples of this are:

$$\text{For } f(x) = x^k,\ f'(x) = kx^{k-1}.$$

[3] To be precise: it is implicit here that f has a derivative $f'(x')$ for all x' close to x and that its derivative f' has a derivative f'' at x.

[4] Let me remind those of you who have seen this before: The second-order condition is sufficient for a local max; it is not necessary; for instance, consider $f(x) = -x^4$ at $x = 0$.

This works if k is an integer, so (for example), if $f(x) = x^3$, then $f'(x) = 3x^2$. But it works equally well for fractional powers of x; for instance, if $f(x) = x^{1/2}$, then $f'(x) = (1/2)x^{-1/2}$. And it works for negative powers of x: if $f(x) = x^{-2.76}$, then $f'(x) = -2.76x^{-3.76}$.

(Do you know about negative and fractional powers? Let me remind you: $x^{1/2}$ means the square root of x; $x^{2/3}$ means the square of the cube-root of x; $x^{5.76}$ means x^5 times the 100th root of x raised to the 76th power; x^{-5} means $1/x^5$; and so on.)

$$\text{For } f(x) = e^x,\, f'(x) = e^x.$$

This is purely for cultural enrichment. The exponential function is used in places in the text, but we do not take its derivative. (You will, however, have to know how to work with this function and to have Excel compute it for you.)

$$\text{For } f(x) = \ln(x),\, f'(x) = 1/x = x^{-1}.$$

Here $\ln(x)$ means the natural logarithm of x, which is sometimes written $\log(x)$ or $\log_e(x)$. We use the natural logarithm function in the text and exercises, and you will need to know at what it looks like, how to get your calculator (or Excel) to compute it for you, and its derivative.

A1.4. Five Important Rules

In addition to these formulas, derivatives of complicated functions like $f(x) = x^2 - x + 3$ are computed by using the following five rules.

Adding a Constant

If $f(x) = g(x) + c$, where c is some constant, then $f'(x) = g'(x)$. If you just want to memorize the rule, be my guest, but this is really quite sensible. If we add (or subtract) a constant to a function for every value of x, we certainly do not change its rate of increase or decrease at any point.

To see this rule in action, consider $f(x) = x^2 - x + 3$. Since 3 is a constant, $f'(x)$ is the derivative of $x^2 - x$; we can forget about the 3.

The Addition Rule

If $f(x) = g(x) + h(x)$, then $f'(x) = g'(x) + h'(x)$. Hence, for the function $x^2 - x$, its derivative is the sum of the derivatives of x^2, which we now know is $2x$, and the derivative of the function $-x$, for which we need the next rule.

Multiplication by a Constant

If $f(x) = kg(x)$ for some constant k, then $f'(x) = kg'(x)$. Intuitively, suppose $g(x)$ is the profit earned by a firm that produces and sells x units of output, measured in dollars. Suppose 110 yen equal 1 dollar; so if $f(x)$ is the profit measured in yen, $f(x) = 110g(x)$. Since $g'(x)$ is the rate of change in profit, measured in dollars, at production level x, the rate of change in yen, $f'(x)$, is just 110 times $g'(x)$. Which is just what the rule says.

Hence the function $-x$, which is $-1 \times x$, has derivative -1 times the derivative of x, which is 1. Therefore, for $f(x) = x^2 - x + 2$, $f'(x) = 2x - 1$. (Now you know why I draw the derivative of $x^2 - x + 2$ as a linear function in Figure A1.3(b).)

The Product Rule

If $f(x) = g(x)h(x)$, then $f'(x) = g'(x)h(x) + h'(x)g(x)$. The most intuitive explanation for this I know of runs as follows.[5] Think in terms of changing x a little bit, to $x + \delta$. Then, approximately, $f(x + \delta) = f(x) + \delta f'(x)$. And $g(x + \delta) = g(x) + \delta g'(x)$. So, approximately,

$$\begin{aligned} f(x+\delta)g(x+\delta) &= [f(x) + \delta f'(x)][g(x) + \delta g'(x)] \\ &= f(x)g(x) + \delta f'(x)g(x) + \delta g'(x)f(x) + \delta^2 f'(x)g'(x)] \\ &= f(x)g(x) + \delta[f'(x)g(x) + g'(x)f(x)] + \delta^2 f'(x)g'(x). \end{aligned}$$

The last term is very, very small if δ is small; so the "slope" of $f(x)g(x)$ for small changes δ in x is $f'(x)g(x) + g(x)f'(x)$, just as the product rule says.

The Chain Rule

If $f(x) = g[h(x)]$, then $f'(x) = g'[h(x)]h'(x)$. Think of a firm that transforms raw material into a saleable product. Let x be the amount of raw material the firm uses, and suppose $Q(x)$ tells us the amount of final product that can be produced from x units of raw material. Suppose as well that, if the firm produces and sells q units of final output, its total revenues are given by the function $\mathrm{TR}(q)$. Then, the total revenues of the firm *as a function of the amount of raw material it uses* is given by the composite function $\mathrm{TR}[Q(x)]$.

What is the rate of change in total revenue as a function of x? First we ask, what is the rate of change in final output as a function of x, what is $Q'(x)$? Suppose that, at some level of raw material input x_0, we get (on

[5] This explanation may make a bit more sense to you after you read the next section. And it may not make sense at all, if you are seeing calculus for the first time or for the first time in a long time.

the margin) two units more of output for every unit of input we use, or $Q'(x_0) = 2$.

Suppose as well that, at x_0 units of input, we have $Q(x_0) = q_0$ units of final product, and total revenues for the marginal extra unit of final product rise by \$3. That is, $\mathrm{TR}'(q_0) = 3$.

Now, at x_0 units of input, what is the approximate impact on total revenue of a marginal additional unit of input? On the margin, one more unit of input raises output by approximately two units. Each of those two (marginal) units raises total revenues by approximately \$3. Hence, the marginal impact of an extra unit of input is \$6 in revenue, which is just what the chain rule tells us.

A1.5. Derivatives and Discrete Changes

We said earlier that the derivative of the function f at a point x is the limit of the rates of change in the function over smaller and smaller intervals that encompass the point x. In a sense then, the rates of change over intervals give an *approximation* to the derivative of the function, an approximation that is better the smaller is the interval.

Turning this around, if we can compute derivatives analytically, we can use them to approximate discrete changes over intervals of positive length, approximations that improve the smaller is the interval.

To give an example, consider the function

$$f(x) = 4x^2 - x^{-3} + 5\ln(x).$$

Suppose for some reason I want to know what are $f(13) - f(10)$ and then $f(10.3) - f(10)$. One way to get these discrete differences would be to compute f for the arguments 13, 10.3, and 10 and subtract. But, if I am in a hurry, I can get an approximation by (a) computing the derivative of f at $x = 10$, and then (b) multiply this by the length of the interval.

The derivative of $f(x)$ is

$$f'(x) = 4 \times 2x - (-3) \times x^{-4} + \frac{5}{x} = 8x + 3x^{-4} + \frac{5}{x}.$$

Evaluated at $x = 10$, this is

$$f'(10) = 80 + \frac{3}{10^4} + \frac{5}{10} = 80.5003.$$

Therefore, my estimate of $f(13) - f(10)$ is $80.5003 \times 3 = 241.5009$, and my estimate of $f(10.3) - f(10)$ is $80.5003 \times .3 = 24.15009$. In fact, according to EXCEL, $f(13) - f(10) = 277.312$, and $f(10.3) - f(10) = 24.507$. The error is around 15% on the larger interval and 1.5% on the smaller interval. We have a fairly good quick approximation in both cases, but the approximation is a lot better on the smaller interval.

A1.6. Integrals and Integration

For a function $f(x)$, its derivative $f'(x)$ is the slope of f at x, for each x. The *integral* of $f(x)$, on the other hand, is "the" function whose derivative is $f(x)$.

Why are there quotes around "the"? Because many functions have the derivative $f(x)$. Suppose that $g(x)$ is a function whose derivative is $f(x)$. The rules about derivatives tell us that for any constant k, $h(x) = g(x) + k$ is another function whose derivative is $f(x)$.

For example, consider $f(x) = 3x^2 - 4x + 3$ and $g(x) = x^3 - 2x^2 + 3x$. By the rules and formulas for derivatives, $g'(x) = f(x)$. But $f(x)$ is the derivative as well of $h(x) = x^3 - 2x^2 + 3x + 101$ and of $x^3 - 2x^2 + 3x - 101010101$.

Because of this, when we speak of the integral of $f(x)$, we either mean the entire class of functions whose derivatives are $f(x)$, all of which are constant translates of one another,[6] or we specify some single value of the integral we are interested in. For instance, we might say something like this: Let $g(x)$ be that function whose derivative is $f(x) = 3x^2 - 4x + 3$ and that takes on the value 10 at $x = 1$. If I am looking for this function, I proceed as follows: First, I note that integrals of $f(x)$ are functions of form $x^3 - 2x^2 + 3x + k$, for a constant k. Thus, at $x = 1$, the function has the value $1 - 2 + 3 + k = 2 + k$. If I want this to equal 10, then $k = 8$.

Actually, it is rare that one specifies an integral g of f with $g(1) = 10$ or $g(x) =$ anything except 0. That is, when we pin down an integral, normally we specify some argument x_0 and say that we want an integral whose value at x_0 is 0. We write this as

$$g(x) = \int_{x_0}^{x} f(x)\,\mathrm{d}x.$$

[6] Might two functions $g(x)$ and $h(x)$ both have the derivative f yet not be constant translates of one another? The answer is no. Suppose g and h both have f as their derivative. Look at the function $D(x) = g(x) - h(x)$. The rules for finding derivatives tell us that the derivative of D is $g'(x) - h'(x)$, which, by our hypothesis that $g'(x) = f(x) = h'(x)$, is the constant 0. But this means that $D(x)$ never increases or decreases; it is a constant k. This implies that $g(x) = h(x) + k$, just as we claim.

This brings us to the characterization of integrals with which you are probably most familiar (if you have any familiarity with integrals). Suppose I want

$$g(x) = \int_{x_0}^{x} f(x)\,\mathrm{d}x.$$

That is, I want g to be the function whose derivative is $f(x)$ and value 0 at x_0. The *Fundamental Theorem of Calculus* tells us that:

> *For $x > x_0$, the integral $g(x)$ is the area under the function f over the range from x_0 to x; for $x < x_0$, it is the negative of this area.*

Consider Figure A1.6, the function f graphed there, and the point x_0. We claim that $g(x) = \int_{x_0}^{x} f(x)\,\mathrm{d}x$ is the shaded area, for $x > x_0$. This is not simply a definition, instead it is a mathematical, logical assertion that the function $g(x)$ constructed in this fashion satisfies $g(x_0) = 0$ (which should be obvious) and $g'(x) = f(x)$ (which is anything but obvious).

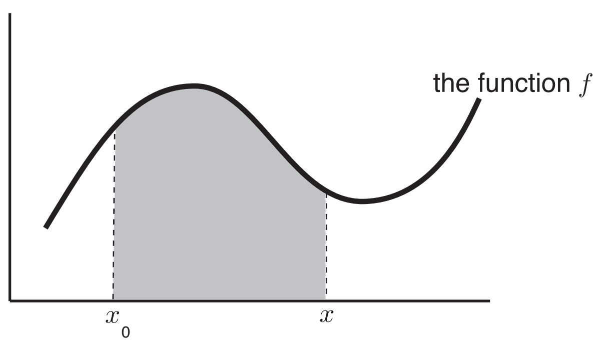


Figure A1.6. The integral of f.

Why is $g'(x) = f(x)$? Recall that $g'(x)$ is the limit of the discrete rates of change of g over smaller and smaller intervals that begin, or end, or bracket x. Suppose we look at an interval from x to $x + \delta$, for small values of δ. The difference between $g(x)$ and $g(x + \delta)$ is the difference in the areas that define these two, which is the shaded "rectangular sliver" in Figure A1.7. This is almost a rectangle whose base is δ and whose height is $f(x)$. It is only almost a rectangle because the height of the rectangle varies between $f(x)$ and $f(x + \delta)$. Suppose the function f is continuous (no jumps) and δ is small. Then, the height of f over this interval does not vary too much and the area of the shaded region is approximately $f(x) \times \delta$. Therefore, the

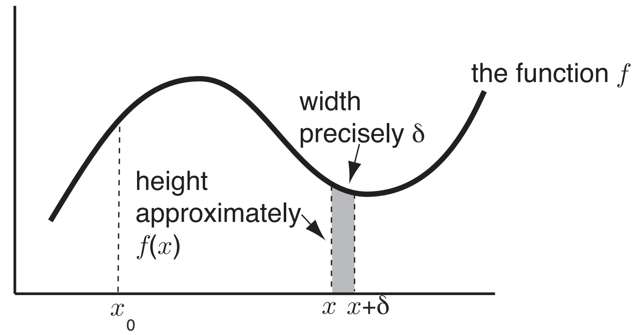


Figure A1.7. Why the area under f is the integral of f.

rate of increase of the area function at x is $f(x)$; the derivative of the "area" function is f.

Textbooks on calculus go on from this point to do several things, such as the following:

1. They give formulas for various important integrals, such as

$$\int_{x_0}^{x} x^k \,\mathrm{d}x = \frac{x^{k+1}}{k+1} - \frac{x_0^{k+1}}{k+1} \text{ for } k \neq -1.$$

2. They give rules, such as

$$\int_{x_0}^{x} [f(x) + g(x)] \,\mathrm{d}x = \int_{x_0}^{x} f(x) \,\mathrm{d}x + \int_{x_0}^{x} g(x) \,\mathrm{d}x.$$

3. They do all manner of other fancy stuff.

For purposes of reading the textbook, you don't need to bother with this "integral" stuff, except to be clear on three things:

1. The integral of the function f is a function whose derivative is f.
2. You have to pin down the value of an integral at some point before you are sure what it is, and the usual thing is to specify some point x_0 and say that the integral should have the value 0 at x_0.
3. In which case, the integral of f is the area under the curve f, from x_0 up to (or, with a negative sign, down to) the argument x of the integral.

A1.7. Partial Derivatives

So far we have dealt with functions of a single variable; the argument of the function is a number, and the value of the function is another number. We continue with functions whose values are numbers, but in many places we need to discuss functions whose arguments are vectors of numbers. Such functions are called *multivariate functions*.

For example, consider the function $F(x, y) = x^2 + 3xy - 2y^2 + 7x - y\ln(x) + 3$. If I give you values for x and y (where $x > 0$ is required, so that $\ln(x)$ makes sense), you can plug them in (or use Excel or your calculator) to evaluate the function.

Multivariate functions come up in economics because decision makers in economic problems often control more than one variable. Multiproduct firms have to decide how much of each sort of product to produce and sell. Even if the firm has a single product, this product can often be produced with varying combinations of several inputs. Consumers have to decide how much milk, bread, cheese, beef, beer, wine, and the like to consume. Hence, we have functions to be maximized or minimized that have more than one variable. How do we deal with these using calculus?

The first step involves the concept of a *partial derivative*. This gives the instantaneous rate of change of the function as we vary one of its arguments, leaving the other arguments at some fixed level.

For instance, suppose we are interested in how $F(x, y) = x^2 + 3xy - 2y^2 + 7x - y\ln(x) + 3$ changes as we vary x around the value $x = 2$, with y fixed at 4. If y is fixed at 4, the function, as a function of x alone, is $F(x, 4) = x^2 + 12x - 32 + 7x - 4\ln(x) + 3$. And the instantaneous rate of change of this function in x is just its derivative with respect to x, which is

$$\frac{dF(x,4)}{dx} = 2x + 12 + 7 - \frac{4}{x} = 2x + 19 - \frac{4}{x}.$$

At $x = 2$, this is 21. For small increases in x (at $x = 2$, $y = 4$), the function rises at a rate of approximately 21 times the amount that x is increased. For example,

$$F(2.001, 4) - F(2, 4) \approx 21 \times 0.001 = 0.021.$$

At the same time, if we fix $x = 2$ and vary y slightly, we are looking at the function (in y) $4 + 6y - 2y^2 + 14 - y\ln(2) + 3$, whose derivative in y is $6 - 4y - \ln(2)$. At $y = 4$, this is $-10 - \ln(2)$. Thus, if we increase y from 4 to 4.01, keeping x fixed at 2, the function *decreases* by approximately $0.01 \times [10 + \ln(2)]$.

What we have done is to fix one of the variables and look at the derivative in the other. More generally, when we have functions of more than two variables, we fix all but one variable and look at the derivative or instantaneous rate of change in the one not fixed.

If we do this in general, we have what is known as the *partial derivative* of the function. For our two-variable function F, we would write for the partial derivative of F with respect to x the symbols

$$\frac{\partial F(x,y)}{\partial x},$$

where a curly d indicates that this is a partial derivative. This is just the derivative of F with respect to x, treating y as a constant. All the rules you know from differentiation of a function of a single variable apply, so that

$$\frac{\partial F(x,y)}{\partial x} = \frac{\partial[x^2 + 3xy - 2y^2 + 7x - y\ln(x) + 3]}{\partial x} =$$

$$\frac{\partial x^2}{\partial x} + \frac{\partial(3xy)}{\partial x} + \frac{\partial(-2y^2)}{\partial x} + \frac{\partial(7x)}{\partial x} + \frac{\partial[-y\ln(x)]}{\partial x}$$

by the addition rule for derivatives, which term by term is

$$= 2x + 3y + 0 + 7 - \frac{y}{x},$$

applying the rules and formulas for differentiation, *always treating y as a constant*. That is,

$$\frac{\partial F(x,y)}{\partial x} = 2x + 3y + 7 - \frac{y}{x}.$$

At $x = 2$ and $y = 4$, this is just $2 \times 2 + 3 \times 4 + 7 - (4/2) = 21$. (This is the same 21 we computed three paragraphs ago.)

Just to check, what is

$$\frac{\partial F(x,y)}{\partial y}?$$

The answer is given at the end of this appendix.

We use partial derivatives in much the same way we use regular derivatives. For one thing, they can be handy for computing approximate discrete changes in a function. Suppose, for example, we want to know, for the function $F(x, y)$, what is

$$F(2.2, 4.1) - F(2, 4).$$

This amounts to an increase in x by 0.2 and an increase in y by 0.1. Hence, the discrete difference we are looking for is approximately

$$\left.\frac{\partial F(x, y)}{\partial x}\right|_{(x,\, y)=(2,\, 4)} \times 0.2 \quad + \quad \left.\frac{\partial F(x, y)}{\partial y}\right|_{(x,\, y)=(2,\, 4)} \times 0.1 =$$

$$21 \times 0.2 \quad + \quad [-10 + \ln(2)] \times 0.1 \quad = \quad 4.2 - 1 - 0.1\ln(2) = 3.13,$$

where I evaluated $\ln(2)$ on my calculator. (The term

$$\left.\frac{\partial F(x, y)}{\partial x}\right|_{(x,\, y)=(2,\, 4)}$$

means the partial derivative of F with respect to x, evaluated at the values $x = 2$ and $y = 4$.) I asked Excel to compute this difference by evaluating $F(2.2, 4.1)$ and $F(2, 4)$, and Excel told me that the precise difference is 3.22. This is around a 3% error, which is not too bad.

A1.8. Maximization and Minimization of Multivariate Functions

The main way in which we use partial derivatives is in maximization and minimization problems. Suppose we are looking for values of x, y, and z that maximize the function $G(x, y, z)$. At any maximum, all three partial derivatives of G must be 0. Why? If, say $\partial G/\partial z < 0$, then a small decrease in z, leaving x and y fixed, increases the value of the function. If $\partial G/\partial x > 0$, then a small increase in x, leaving y and z fixed increases the value of the function. If the function is maximized at some point, it must be (instantaneously) flat in all three directions.

As with regular derivatives of functions of a single variable, having partial derivatives equal to 0 is only necessary for finding a max or a min (as long as the function is differentiable in all its arguments); it is not sufficient.

There are generalizations of the second-order conditions for multivariate optimization problems, involving second-partial derivatives of the function to be maxed or minned, but we have no occasion to bother with those, so I do not review them here.

A1.9. One more time...

This appendix is in now way a substitute for a textbook in calculus. If you have never had a course in calculus (or otherwise studied a serious textbook on the subject), you may need help when calculus is employed in the text. That said, microeconomics is a great context for learning how to use calculus, so whether this is all new to you, or you learned calculus so long ago that this feels like it is all new to you, if you follow the applications of calculus in the text, *making sure that you understand the intuition behind those applications* you will get a bonus from studying microeconomics, namely a (better) sense of what (differential) calculus is and how it can be employed.

A1.10. Exercises

Following are some problems that review some of the important ideas in this appendix. Some are a bit tricky, but to understand the use of calculus in the text, at least in terms of differentiation, you should be able to do problems A1.1, A1.3(a), A1.4(a), A1.5(a), and A1.6.

A1.1 Evaluate the derivatives of the following functions:

(a) $f(x) = x^7 - 4.5x^{3.2} + 7\ln(x) + 100$

(b) $g(x) = x^{-3}$

(c) $h(x) = 1/x^3$ (Do not look for tricks that are not there.)

(d) $F(z) = (z - 5)^2 z + 3.14159$

(e) $A(r) = 3.14159r^2$ (This is a famous formula, if 3.14159 is replaced by the mathematical constant π. If you remember what it is the famous formula for and want a challenge, try to draw a picture that "explains" the answer you are getting. If you can do that, here is a hard question. You may remember from high school geometry that the formula for the volume of a sphere of radius r is $4\pi r^3/3$. A much less well-known formula is the formula for the surface area of a sphere. What is this formula? If you can work this out, you are way ahead of the game.)

A1.2 Two more derivatives to take (these are harder than anything we use in the book, so do not worry if you have some problems with them):

(a) $G(y) = y^2 \ln(y)$

(b) $H(y) = \ln(y^6 + 2y^3 + 10)$

A1.3 Part a should be fairly easy. Part b is tougher.

(a) Suppose $F(x) = x^3 + 2x^2 + 3x + 4$. Using calculus, compute (approximately) $F(2.08) - F(2)$.

(b) The answer to Problem A1.2(b) is

$$\frac{1}{y^6 + 2y^3 + 10} \times (6y^5 + 6y^2)$$

(in case you did not get it). What then is $H(1.05) - H(0.95)$, for this function H? (I obviously want an approximate answer, calculated using calculus.)

A1.4 Once again, part b may strain your capabilities, if you are not used to this stuff.

(a) If $M(x, y) = x^4 + x^2y^2 + y^4$, what is the partial derivative of M in the variable x?

(b) If $H(a, b) = (a^2 + b^2)^{1/2}$, what is the partial derivative of H in the variable a?

A1.5 You should have no problems with part a, but part b might be hard.

(a) If $M(x, y) = x^3 + 3x^2y + 2xy^2 + y^2$, approximately what is $M(2.1, 0.9) - M(2, 1)$?

(b) Suppose we have a right triangle whose "short" sides have lengths 3 and 4, respectively. If we enlarge each of these sides by 0.01, how much longer (approximately, using calculus) is the hypotenuse. (I should add that this is a problem where using the calculus approximation is harder than computing the difference directly. But do not let that stop you from using the calculus approximation, and please look at the answer to A1.4(b) if you need to before doing this one.)

A1.6 Is it possible that the function $F(x, y, z) = 2x^2 - 2xy + yz + y^2 - 3xz + 3z^2$ hits a (local) minimum at the point $(x = 2, y = 1, z = 2)$? If yes, why? If not, can you find a point nearby this point for which F is smaller?

Solutions

Concerning Figure A1.5

Panel a goes with panel y, b with x, and c with z.

Concerning the Exercise on Page C19

$$\frac{\partial F(x,y)}{\partial y} = 3x - 4y - \ln(x).$$

Solution to Problem A1.1

(a) $7x^6 - 14.4x^{2.2} + 7/x$; (b) $-3x^{-4}$; (c) same as (b), which can be written $-3/x^4$; (d) You can use the product rule to get $2(z-5)z + (z-5)^2 = 2z^2 - 10z + z^2 - 10z + 25 = 3z^2 - 20z + 25$, or you can expand this polynomial into $z^3 - 10z^2 + 25z + 3.14159$ and take the derivative of that, getting the same $3z^2 - 20z + 25$; (e) I use the symbol π for 3.14159 (which is not quite accurate), and the derivative is $2\pi r$. Then πr^2 is the formula for the area of a circle with radius r, while $2\pi r$ is its circumference, and you just learned that the rate of change of the area of a circle, as you change its radius, is the circumference of the circle. Now, draw a picture. If you want to know, the formula for the surface area of a sphere, it is $4\pi r^2$. Why?

Solution to Problem A1.2

(a) This one needs the product rule. The answer is

$$2y\ln(y) + y^2/y = 2y\ln(y) + y.$$

(b) Here you need the rule for taking the derivative of a composition of two functions. The answer is

$$\frac{1}{y^6 + 2y^3 + 10} \times (6y^5 + 6y^2).$$

Solution to Problem A1.3

(a) The derivative of F is $3x^2 + 4x + 3$, which at the value of $x = 2$ is $3 \times 4 + 4 \times 2 + 3 = 23$. Over this range, x changes by 0.08, so the approximate change in the function is $0.08 \times 23 = 1.84$. (In fact, the exact change, calculated using Excel, is 1.891712.)

(b) This is a change of 0.1 in the argument, in a range where the argument of the function is around 1. The derivative of H at the value of 1 (determined by plugging in the formula) is $\frac{12}{13}$, so the answer is $\frac{12}{13} \times 0.1 = 0.0923077$. Note that, if I use Excel to evaluate the function H at 1.05 and then at 0.95 and subtract, I get a difference of 0.09242314, so the calculus-based approximation is pretty good.

Solution to Problem A1.4

(a) $4x^3 + 2xy^2$

(b)

$$(1/2)(a^2 + b^2)^{-1/2} \times 2a \quad = \quad \frac{a}{(a^2 + b^2)^{1/2}}$$

Solution to Problem A1.5

(a) First we calculate the two partial derivatives and evaluate them at the point $(2, 1)$:

$$\frac{\partial M}{\partial x} = 3x^2 + 6xy + 2y^2,$$

which is 26 at $(x = 2, y = 1)$. And

$$\frac{\partial M}{\partial y} = 3x^2 + 4xy + 2y,$$

which is 22 at $(2, 1)$. So if we increase x by 0.1 and simultaneously decrease y by 0.1, we should see a net change in the value of the function of $26 \times 0.1 + 22 \times (-0.1) = 0.4$. (In fact, the exact change is 0.38.)

(b) The formula for the length of the hypotenuse of a right triangle is $(a^2 + b^2)^{1/2}$, where a and b are the lengths of the two short sides, so that the two partial derivatives of the length of the hypotenuse as functions of a and b are, respectively, $a/(a^2+b^2)^{1/2}$ and $b/(a^2+b^2)^{1/2}$. At the point $a = 3$ and $b = 4$, these are $\frac{3}{5}$ and $\frac{4}{5}$, respectively, so if you enlarge each side by 0.01, the length of the hypotenuse will increase by (approximately) $\frac{3}{5}\times 0.01 + \frac{4}{5}\times 0.01 = 0.014$. (In fact, the exact answer is .014000399…, so we are not off by much at all.)

Solution to Problem A1.6

The partial derivatives of F and their values at the point $(x = 2, y = 1, z = 2)$, are

$$\frac{\partial F}{\partial x} = 4x - 2y - 3z, \quad \text{hence} \quad \left.\frac{\partial F}{\partial x}\right|_{(x=2,y=1,z=2)} = 8 - 2 - 6 = 0;$$

$$\frac{\partial F}{\partial y} = -2x + z + 2y, \quad \text{hence} \quad \left.\frac{\partial F}{\partial y}\right|_{(x=2,y=1,z=2)} = -4 + 2 + 2 = 0;$$

and

$$\frac{\partial F}{\partial z} = y - 3x + 6z, \quad \text{hence} \quad \left.\frac{\partial F}{\partial z}\right|_{(x=2,y=1,z=2)} = 1 - 6 + 12 = 7.$$

To be a candidate for a minimum, all three partials must be 0; since they are not, this point cannot be a minimum. In fact, the partials tell us that decreasing z a bit, starting at $(x = 2, y = 1, z = 2)$ while holding x and y at 2 and 1, respectively, causes the function to decrease; more specifically, by moving to (say) $(x = 2, y = 1, z = 1.999)$, the function decreases by approximately $0.001 \times 7 = 0.007$.