

Eric Rasmusen

November 20, 2018

NOTES ON SLOPES AND TAKING THE DERIVATIVE OF x^2 , $1/x$, \sqrt{x} ,
AND 5^x .

These notes are rather long, but mathematics often has the perverse feature that if someone writes a long explanation, the reader can read it much more quickly than if he writes a short explanation, because all the steps are written out and the reader doesn't have to think of them himself.

Separate notes cover the rules for taking derivatives. These notes are in case you want to understand *why* the formulas are the way they are. One of online MBA students came from a Great Books undergrad background, so he said he'd like to understand why the calculus formulas are true rather than just memorize them. I decided to try, especially since I can teach these to my teenagers too, and learn to understand derivatives better myself. I feel inspired to write some notes on Continuity next— Lipschitz, uniform, weak*, absolute, etc.. (I've also written notes on Exponents and Cournot that I've written this fall.)

What's I've done here is use analogical reasoning, the kind used in law. I don't know if that's correct reasoning here, and am eager for comment. I get to the right answers, so I know it's correct in some sense, but they might be rationalizations rather than reasons for the calculus rules. I hope this is a good approach, though, because I find I don't really understand the conventional approach to derivatives (the $\frac{f(x+h)-f(x)}{h}$ limit, which I discuss at the end of these notes). Mathematics usually uses a step by step careful approach to proofs. That's a safe way to do it, because you can think about and understand each step. But the more steps you have to keep in your mind, the harder it is to understand the overall proof, which means the harder it is to understand why the proposition is true. Paradoxically, a careful proof

makes it easier to know something is correct but harder to understand why it is correct. The brighter you are, the less you need to use math, but the dimmer you are, the harder is it for you to follow math. The implication is that if you aren't bright, your only hope is to use careful mathematics, but you might not be able to understand the proposition at all. I hope that the step by step approach I use here has all the good features one could want— correctness, ease, and understandability— but it might be that I fail in all three features. Comments welcomed on each of them.

So let's proceed. The basic formula for all the derivatives we'll be covering except for the derivative of 5^x is

$$\frac{d}{dx}ax^b = bax^{b-1} \quad (1)$$

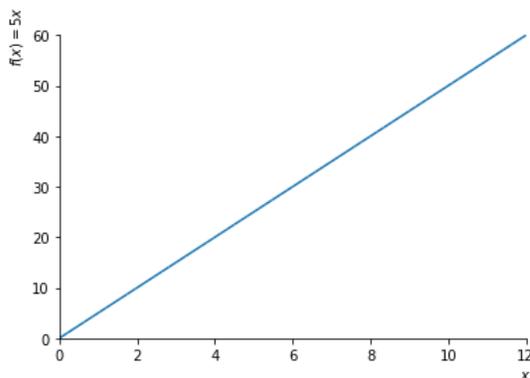
so that

$$\begin{aligned} \frac{d}{dx}5x^2 &= 2 \cdot 5x^{2-1} \\ &= 10x \end{aligned} \quad (2)$$

To start, remember that the derivative of a function is the slope of the curve it draws, and the slope is “rise over run”. Thus, the derivative of $f(x) = 4$ is zero, because the slope of the flat line $f = 4$ is zero. The derivative of $f(x) = 5x$ is 5, because its slope is 5. If x goes from 2 to 3 (the run), then $f(x)$ goes from 10 to 15 (the rise), and the slope is rise over run, or $5/1$, which is 5. That's easy to see for a straight diagonal line like $f(x) = 5x$ or $f(x) = 4 + 5x$. So our first rule, illustrated in Figure 1, is

$$\frac{d}{dx}4 + 5x = 5. \quad (3)$$

FIGURE 1: THE FUNCTION $f(x) = 5x$

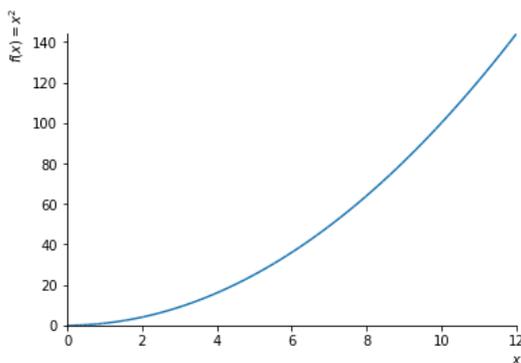


Suppose we have $f(x, z) = zx$. The slope would be z , because this would be a straight diagonal line. If we increase x by 1, that increases $f(x, z)$ by amount z . But also, if we change z by 1, that increases $f(x, z)$ by amount x . We can take the derivative with respect to either x or z . There are two slopes, depending on whether we graph $f(x, z)$

in (x, f) space or in (z, f) space. Or, of course, we could have a three-dimensional diagram in (x, f, z) space, and then we'd be talking about the slope of a particular direction down the mountain of f values. The slopes is called a “partial derivative” in this case, but it really the same thing as a regular old derivative, though tradition calls for it to be written with a curly d instead of a regular one, so

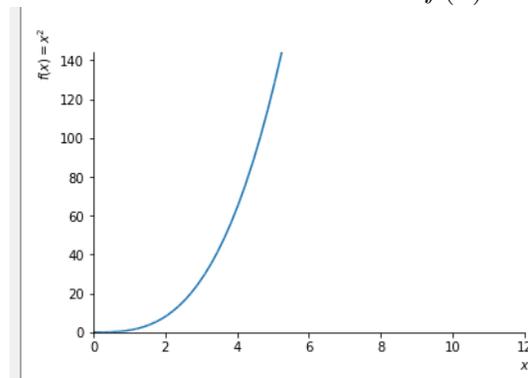
$$\frac{\partial}{\partial x}zx = z. \quad (4)$$

FIGURE 2: THE FUNCTION $f(x) = x^2$



Where calculus becomes useful is when the function is nonlinear, like $f(x) = x^2$ or $f(x) = 1/x$. Then, the slope keeps changing whenever x changes, and it isn't obvious what the slope is at any point. Suppose we have $f(x) = x^2$, as in Figure 2. That's the same as $f(x) = x \cdot x$. Let's label each of these variables, even though since they're the same variable, they have to take the same value: $f(x) = x_1 \cdot x_2$, where $x_1 = x_2$. Think about increasing the value of x_2 . That will increase the value of $f(x)$ at a rate of x_1 , just like increasing x would increase the value of $5x$ at a rate of 5. But when we increase x_2 , we also are increase x_1 , and that increases the value of $f(x)$ at a rate of x_2 . So the total rate of increase when we increase x is $x_1 + x_2$, which equals $2x$, and we have found that the slope of the function $f(x) = x^2$ is $2x$, just what we get from the calculus rule:

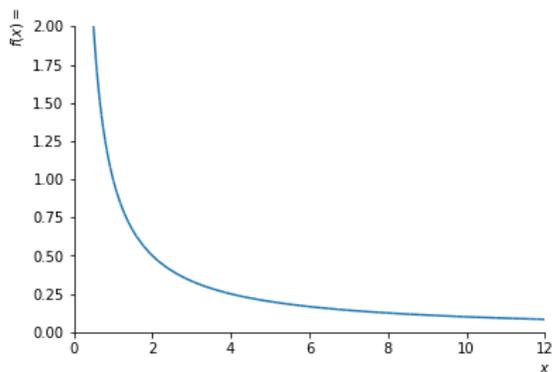
$$\frac{d}{dx}x^2 = 2x. \quad (5)$$

FIGURE 3: THE FUNCTION $f(x) = x^3$ 

How about $f(x) = x^3$, as in Figure 3? Well, let's think of that as $f(x) = x_1x_2x_3$. When we change x_1 , that increases f at a rate of x_2x_3 . When we change x_2 , that increases f at a rate of x_1x_3 . When we change x_3 , that increases f at a rate of x_1x_2 . If we increase all of them at once, $f(x)$ thus increases by $x_2x_3 + x_1x_3 + x_1x_2$. But since $x_1 = x_2 = x_3$, that means $f(x)$ increases at a rate of $x^2 + x^2 + x^2$, which equals $3x^2$. Thus,

$$\frac{d}{dx}x^3 = 3x^2. \quad (6)$$

We could use the same reasoning to derive the slopes of $f(x) = x^4$, $f(x) = x^5$, and so forth.

FIGURE 4: THE FUNCTION $f(x) = \frac{1}{x}$ 

How about $f(x) = x^{-1}$, as in Figure 4? We need different reasoning for that. Recall that $x^{-1} \equiv 1/x$. Next, $1/x = \frac{x}{x^2}$. Let's again use subscripts to distinguish between different places where x enters $f(x)$, so

$$1/x = \frac{x}{x^2} = x_1 \frac{1}{x_2} \frac{1}{x_3} \quad (7)$$

Let us denote $Z \equiv \frac{d}{dx} x^{-1}$. Then

$$Z = (1) \frac{1}{x_2} \frac{1}{x_3} + x Z \frac{1}{x_3} + x \frac{1}{x_2} Z = \frac{1}{x^2} + 2Z, \quad (8)$$

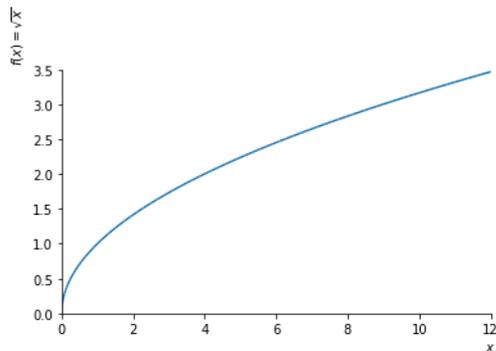
in which case

$$-Z = \frac{1}{x^2} \quad (9)$$

and

$$Z = \frac{d}{dx} x^{-1} = -x^{-2} = -\frac{1}{x^2}, \quad (10)$$

which is the standard formula. I should add more explanation, but I don't have time right now.

FIGURE 5: THE FUNCTION $f(x) = \sqrt{x}$ 

How about $f(x) = x^{.5}$, as in Figure 5? A trick similar to the $f(x) = 1/x$ trick will work. Recall that $x^{.5} \equiv \sqrt{x}$, and $\sqrt{x} \cdot \sqrt{x} = x$ by the definition of square root. Then,

$$\frac{d}{dx}x_1 = \frac{d}{dx}(x_2^{.5}x_3^{.5}) \quad (11)$$

and if we let $W \equiv \frac{d}{dx}x^{.5}$ then

$$\frac{d}{dx}x_1 = Wx_3^{.5} + x_2^{.5}W \quad (12)$$

so

$$1 = 2Wx^{.5} \quad (13)$$

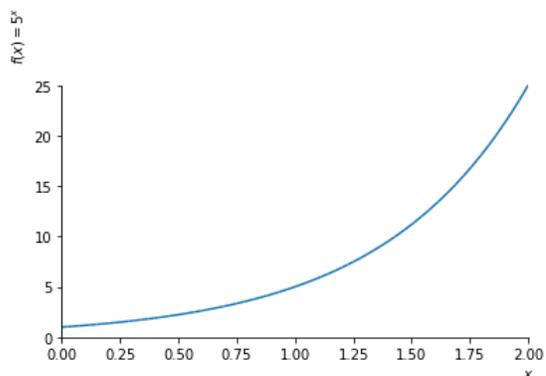
and

$$W = \frac{d}{dx}x^{.5} = \frac{1}{2x^{.5}} = .5x^{-.5} \quad (14)$$

or

$$\frac{d}{dx}\sqrt{x} = \frac{1}{2\sqrt{x}} \quad (15)$$

I should add more explanation, but I don't have time right now.

FIGURE 6: THE FUNCTION $f(x) = x^5$ 

How about $f(x) = 5^x$? First, recall how exponents are multiplied: $5^x \cdot 5^2 = 5^{x+2}$. For the slope, we are thinking of making x a little bigger. Let's make x bigger by amount h . Thus, we are going from 5^x to 5^{x+h} . By the multiplication rule, though, that means we are going from 5^x to $5^x \cdot 5^h$. Notice that whatever value we pick for h , this means the amount of increase is going to be $5^x \cdot 5^h - 5^x$, which equals $5^x(5^h - 1)$, which is proportional to 5^x . Thus, the slope is $5^x \cdot w$, for some constant number w . In other words, as $f(x)$ gets bigger, the slope gets bigger at exactly the same rate: $f'(x) = w f(x)$; or, put a little differently, the ratio of the function's slope to the function itself is constant: $\frac{f'(x)}{f(x)} = w$.

What is that constant number w ? It depends on whether we have $f(x) = 5^x$ or $f(x) = 45^x$. If the base number is 5, the slope is going to be a lot lower than if it is 45. Here's why. Remember the exponent rules: if $x = 0$, then $5^x = 5^0 = 1$ and $45^x = 45^0 = 1$ also. But we know the slope of 45^x , which is $w_{45} 45^x$, should be a lot bigger than the slope of 5^x , which is $w_5 5^x$, so it must be that w_{45} is bigger than w_5 . To see this, let's be concrete and use the numbers .0001 and 992, though what matters is picking *some* small enough and *some* large enough base amount. The slope of $.0001^x$ must be tiny. If $x = 0$, then the slope of $.0001^x$ would be $w_{.0001} .0001^0$, which equals $w_{.0001}$, so $w_{.0001}$ must be tiny too. In particular, it must be that $w_{.0001} < 1$, whereas $w_{992} > 1$. Somewhere in between .0001 and 992 there will be a base number e such that $w_e = 1$ exactly. That number has been calculated. It is

about 2.7, and it is indeed known as “ e ”, just as the ratio of a circle’s circumference to its diameter, which is about 3, is known as pi, π .

What’s most important is that for any exponential function, whatever the base, the slope changes with x but is always proportional to $f(x)$. In other words, the slope is always $wf(x)$ for some number w that depends on the base. But we can then come up with notation to describe what the w ’s are for other base numbers besides e , using the exponent rules and the idea of the “logarithm”. Define $\log(z)$ as the function such that $e^{\log(z)} = z$. Thus, $\log(5)$ would be the number such that $e^{\log(5)} = 5$. That number is about 2, because $2.7^2 \approx 5$. Then notice that $5^x = (e^{\log(5)})^x$. By the exponent multiplication rule, $5^x = (e^{\log(5)})^x = e^{\log(5) \cdot x}$. So if we can find

$$\frac{d}{dx} e^{\log(5) \cdot x} \tag{16}$$

we’re done.

Let's start with something simpler looking than $\frac{d}{dx}e^{\log(5)\cdot x}$, the problem of finding

$$\frac{d}{dx}e^{2x} \tag{17}$$

Suppose we start with e^{2x} and we change x by a little. That will have two effects. First, $2x$ will grow at rate of 2. Let's define the function $h(x) = 2x$. We know that

$$\frac{d}{dh}e^h = e^h. \tag{18}$$

Thus, whenever h increases by a little bit, e^h increases at a rate of e^h . But when we are changing x , and $h = 2x$, the little bit we are changing x changes h at a rate of 2. So the total increase in e^h from increasing x a little bit is the 2 directly from the increase in x times the e^h from increasing h , which makes $2e^h$, which is $2e^{2x}$. Thus,

$$\frac{d}{dx}e^{2x} = 2e^{2x}. \tag{19}$$

The same reasoning would apply if $h(x) = \log(5) \cdot x$ instead of $h(x) = 2x$. Thus,

$$\frac{d}{dx}e^{\log(5)\cdot x} = \log(5)e^{\log(5)x} \tag{20}$$

so

$$\frac{d}{dx}5^x = \log(5)5^x \tag{21}$$

The Conventional Approach

These notes use an unconventional, verbal, approach to deriving the derivative rules. The conventional approach is more mathematical, but easier, maybe, for getting the rules. See <http://tutorial.math.lamar.edu/Classes/CalcI/DiffExpLogFcns.aspx> for a nice explanation of it. It's not as good for understanding the rules, I think. In the conventional approach, we look at

$$f'(x) = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h} \quad (22)$$

Thus, for example, if $f(x) = 5x^2$,

$$\begin{aligned} f'(x) &= \lim_{h \rightarrow 0} \frac{5(x+h)^2 - 5x^2}{h} \\ &= \lim_{h \rightarrow 0} \frac{5x^2 + 5h^2 + 5xh + 5xh - 5x^2}{h} \\ &= \lim_{h \rightarrow 0} \frac{10xh}{h} \\ &= \lim_{h \rightarrow 0} 10x \\ &= 10x \end{aligned} \quad (23)$$

We ask whether there is a single limit for each sequence of h that gets closer and closer to 0. Note that this means we need to look at approaching from above ($h = 2, 1, .1, .001, .0001, \dots$) but also at approaching from below ($h = -2, -1, -.1, -.001, -.0001, \dots$). An example of where you might get fooled that way is the function

$$\begin{aligned} f(x) &= x^2 \quad \text{if } x \leq 9 \\ &= 3x \quad \text{if } x \geq 9, \end{aligned} \quad (24)$$

where the limit is $f'(x) = 3$ approaching from above, but $f' = 2x$ approaching from below, so $f(3) = 9$ either way, but $f'(3) = 3$ or $f'(3) = 6$ depending on whether you make h positive or negative.

It also means not just finding one special sequence like $h = 1, .1, .01, .001, \dots$, because there are some functions for which you could find a limit using that special sequence but not with other sequences. Here I will get into higher math, but it still is intuitive. Think of the following function:

$$\begin{aligned} f(x) &= x^2 && \text{if } x \text{ is a multiple of } \frac{1}{10^n} \text{ for any } n \\ &= 17 && \text{if } x \text{ is } \textit{not} \text{ a multiple of } \frac{1}{10^n} \end{aligned} \tag{25}$$

We could find a sequence of h 's that would make $\frac{d}{dx}f(x) = 2x$, I think. Try $\{1, .1, .01, .001, \dots\}$. Most convergent sequences wouldn't get us that result, though. In fact, $f(x) = 17$ for almost all values of x , and $f(x) = 0$ — if you randomly picked a value of x , there is 100% probability that $f(x) = 17$. I'll have to ask somebody who knows real analysis better whether there is also a 100% probability that $f'(x)$ exists and equals zero, but that might be true too, since $f(x)$ is differentiable except at a countable number of points. Moreover, if we require a derivative to be the same for all sequences of h 's, then $f'(x)$ won't exist at the values of x such that $f(x) = x^2$, because $f(x)$ isn't continuous at those values. (Note that $x = \sqrt{17}$ is not an except, because you can't find n to make $\sqrt{17}$ a multiple of $\frac{1}{10^n}$.) The function $f(x)$ jumps up and down around those points, so the slope is infinite; as the change in x gets smaller and approaches zero, the change in $f(x)$ stays the same so the ratio goes to infinity. So that's why we require that the derivative be the same however we pick the sequence of h 's. One reason I mention this is because it shows that though the conventional approach is in some sense safer than analogical reasoning, it, too, can trick you. I think that usually when it's taught, the teacher skips the step of showing that the derivative comes out the same regardless of the sequence of h 's, so in that sense it's no more rigorous or general than my analogical approach.