110.108 CALCULUS I

Week 5 Lecture Notes: September 26 - September 30

---

1. Lecture 3: The Chain Rule

It should be obvious now that the derivative of a product of functions is NOT the product of the derivatives. However, there is a construction involving two functions where the derivative IS the product of the functions. This happens when the construction is the composition of two functions, and the rule governing the derivative of a composition is called the Chain Rule:

Definition 1. Let g be a function differentiable at a point $x = a$ and f a function differentiable at the point $g(a)$. The the function $F = f \circ g$ is differentiable at $x = a$, and

$$F'(x) = \frac{d}{dx}\left[f(g(x))\right]\bigg|_{x=a} = f'(g(x))g'(x).$$

Notes:

- The derivative of a composition IS the product of the derivatives, but there is a twist; the derivative of the outside function is evaluated at the inside function value, so it is not simply the product of the two individual function evaluated at x.
- In Leibniz notation, let $y = f(u)$ and $u = g(x)$ (so that we can recover $y = f(g(x))$). Then the Chain Rule can be expressed as a product
  $$\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx},$$
  and at the point $x = a$,
  $$\frac{dy}{dx}\bigg|_{x=a} = \frac{dy}{du}\bigg|_{u=g(a)} \cdot \frac{du}{dx}\bigg|_{x=a}.$$
  Notice that in the first general expression, the derivative looks like a standard product, but in the second expression one can more easily see that the first derivative in the product is actually evaluated at $u = g(a)$ and not simply a.
- So what is the idea of the proof? In general, the proof of this formula will take some work. However, in a specific case, we can at least see how the formula comes out.

partial proof idea. Suppose that we have a composition of functions of the form $f(g(x))$, where $g(x)$ is differentiable at $x = a$ and $f(x)$ is differentiable at $g(a)$. Then if the derivative of $f(g(x))$ exists at $x = a$, we would have

$$\frac{d}{dx}\left[f(g(x))\right]\bigg|_{x=a} = \lim_{x\to a} \frac{f(g(x)) - f(g(a))}{x - a}.$$

In general, working with this limit will be difficult. However, let's make an extra restriction on the kinds of functions we allow for now: Let assume that, near $x = a$, we have that $g(x)$ is not equal to $g(a)$ (except at a, of course). This is not a huge restriction, but it makes the proof not usable for ANY function. That doesn't matter, though, as I am only interested in demonstrating the formula

---

Date: October 3, 2011.

using the derivative definition. Under the assumption, we can readily alter the derivative definition by a clever form of 1:

$$\begin{aligned}
\frac{d}{dx}\left[f(g(x))\right]\bigg|_{x=a} &= \lim_{x\to a}\frac{f(g(x))-f(g(a))}{x-a} \\
&= \lim_{x\to a}\frac{f(g(x))-f(g(a))}{x-a}\left(\frac{g(x)-g(a)}{g(x)-g(a)}\right) \\
&= \lim_{x\to a}\frac{f(g(x))-f(g(a))}{g(x)-g(a)}\left(\frac{g(x)-g(a)}{x-a}\right) \\
&= \left(\lim_{x\to a}\frac{f(g(x))-f(g(a))}{g(x)-g(a)}\right)\left(\lim_{x\to a}\frac{g(x)-g(a)}{x-a}\right)
\end{aligned}$$

at least when these limits both exist. Notice immediately that the second limit DOES exist, and is just the definition of $g'(a)$. The first, however, is a bit trickier.

Recall that $g(x)$ is differentiable at $x = a$. Thus it is continuous at a, and so $\lim_{x\to a} g(x) = g(a)$ implies that $\lim_{x\to a} g(x) - g(a) = 0$. Hence as $x \to a$, we have $g(x) \to g(a)$. Thus

$$\lim_{x\to a}\frac{f(g(x))-f(g(a))}{g(x)-g(a)} = \lim_{g(x)\to g(a)}\frac{f(g(x))-f(g(a))}{g(x)-g(a)} = f'(g(a))$$

exists. Thus the formula works the way it does because of the structure of the definition of a limit, and

$$\frac{d}{dx}\left[f(g(x))\right]\bigg|_{x=a} = \left(\lim_{g(x)\to g(a)}\frac{f(g(x))-f(g(a))}{g(x)-g(a)}\right)\left(\lim_{x\to a}\frac{g(x)-g(a)}{x-a}\right) = f'(g(a))g'(a).$$

□

Example 2. For $h(x) = \dfrac{1}{x^2+1}$, find $h'(x)$.

First, we do not need the Chain Rule for this calculation. $h(x)$ is a rational function, and we can simply use the Quotient Rule. Here

$$h'(x) = \frac{d}{dx}\left[\frac{1}{x^2+1}\right] = \frac{\frac{d}{dx}\left[1\right](x^2+1)-(1)\frac{d}{dx}\left[x^2+1\right]}{(x^2+1)^2} = \frac{0-2x}{(x^2+1)^2} = \frac{-2x}{(x^2+1)^2}.$$

However, if we rewrite $h(x) = (x^2+1)^{-1}$ instead, then it looks like $h(x)$ is the composition of the power function $f(x) = x^{-1}$ and the quadratic function $g(x) = x^2+1$. Hence, using the Chain Rule, with $f'(x) = -x^{-2}$, and $g'(x) = 2x$, we get

$$h'(x) = \frac{d}{dx}\left[(x^2+1)^{-1}\right] = \frac{d}{dx}\left[f(g(x))\right] = f'(g(x))g'(x) = -(x^2+1)^{-2}(2x) = \frac{-2}{(x^2+1)^2}.$$

Notice that the result is the same either way. This is no accident and is true even if the two calculations resulted in two functions that do not look like each other. This is because when two functions are equal, their graphs are equal. When two functions have the same graph, then the collection of slopes of tangent lines must also be equal. Put this together to say that when two functions are equal, their derivative MUST be equal. In other words, the derivative of a function is unique. This will be important later when we begin to integrate.

Example 3. Let $h(x) = \sin e^x$.

Let $f(x) = \sin x$, and $g(x) = e^x$. Then $f'(x) = \cos x$, $f'(g(x) = \cos e^x$, and $g'(x) = e^x$. And so

$$h'(x) = \frac{d}{dx}\left[f(g(x))\right] = f'(g(x))g'(x) = (\cos e^x)\, e^x = e^x \cos e^x.$$

Example 4. Let $i(x) = e^{\sin x^2}$.

Here, we have actually three functions nested together, and can denote them $f(x) = e^x$, $g(x) = \sin x$, and $h(x) = x^2$. Then $i(x) = f\left(g\left(h(x)\right)\right)$. Finding the derivative of this function, however, is still very straightforward: Simply treat the two inner functions ($g(x)$ and $h(x)$) as a single function during the first application of the Chain Rule. The re-apply the Chain Rule to these two functions: Indeed,

$$\begin{aligned} i'(x) = \frac{d}{dx}\left[f\left(g\left(h(x)\right)\right)\right] &= f'\left(g\left(h(x)\right)\right)\frac{d}{dx}\left[g\left(h(x)\right)\right] \\ &= f'\left(g\left(h(x)\right)\right)\left(g'\left(h(x)\right)h'(x)\right) \\ &= f'\left(g\left(h(x)\right)\right)g'\left(h(x)\right)h'(x). \end{aligned}$$

One final note here: When using the Chain Rule, always differentiate from the outside in. This may be why it is called the Chain Rule. The derivative of a nested set of functions is a chain of derivatives of the functions. It is a product, but each function is evaluated at the value of the remaining function in the chain.

Example 5. Let $g(x) = a^x$, where $a > 0$, and $a \neq 1$.

Here, we previous calculated that $g'(x) = \frac{d}{dx}\left[a^x\right] = a^x g'(0)$, althou9gh we had no way to actually calculate the derivative at $x = 0$. We "measured" it, but really, this is insufficient. However, recall the identity give by a function and its inverse: Let $f(x) = e^x$. Then $f^{-1}(x) = \ln x$, and

$$\left(f \circ f^{-1}\right)(x) = e^{\ln x} = x = \ln e^x = \left(f^{-1} \circ f\right)(x).$$

This allows us to differentiate $g(x)$ by first rewriting it using this last set of identities: $g(x) = a^x = e^{\ln a^x} = e^{x \ln a}$, so

$$g'(x) = \frac{d}{dx}\left[a^x\right] = \frac{d}{dx}\left[e^{x\ln a}\right] = e^{x \ln a} \cdot (\ln a) = a^x \ln a.$$

Hence, we have found two things: a way to calculate the derivative of a general exponential function, and what the derivative of that function is at $x = 0$. As a last note, recall when $a = e$, that the derivative at $x = 0$ was 1. With this new calculation, we get

$$\frac{d}{dx}\left[e^x\right] = e^x \ln e = e^x(1) = e^x.$$