The Chain Rule

Statement of the Chain Rule

In the last lecture, we learned how to compose two functions. In this lecture, we learn how to differentiate the composition of two functions. The rule which allows us to differentiate the composition of two functions is called the Chain Rule.

We state the Chain Rule as follows: let $f(x)$ and $g(x)$ be two real-valued functions whose domains are the real line and which have derivatives. Then the composition of $f(x)$ and $g(x)$, $(g \circ f)(x)$, has a derivative given by the following formula:

$$\frac{\mathrm{d}(g \circ f)}{\mathrm{d}x}(x) = \frac{\mathrm{d}g}{\mathrm{d}x}(f(x)) \cdot \frac{\mathrm{d}f}{\mathrm{d}x}(x).$$

Another way to write the rule above is

$$(g \circ f)'(x) = g'(f(x)) \cdot f'(x).$$

Let us analyze the statement of the Chain Rule step by step. First, we need two real-valued functions whose domains are the real line and which have derivatives. We have plenty of functions in this class already which fit this description, including polynomial functions and sine and cosine functions. For example, let $f(x) = 7x^2$ and let $g(x) = \cos x$. Next, we find the composition of $g(x)$ after $f(x)$:

$$(g \circ f)(x) = g(f(x)) = \cos(f(x)) = \cos(7x^2).$$

The Chain Rule then tells us that $(g \circ f)(x)$ has a derivative, and it gives us a formula with which to compute that derivative. The formula is a product of two derivatives. The first is $g'(f(x))$, which means that we first take the derivative of $g(x)$, and then we replace x with the formula for $f(x)$ in the formula for $g'(x)$. The other derivative is simply the derivative of $f(x)$. Once we have formulae for $g'(f(x))$ and $f'(x)$, we multiply them together to get a formula for $(g \circ f)'(x)$. So, in our example above, we have that $g'(x) = -\sin x$, so we get that

$$g'(f(x)) = -\sin(f(x)) = -\sin(7x^2).$$

We also have that $f'(x) = 14x$. Therefore the derivative of $(g \circ f)(x) = \cos(7x^2)$ is

$$(g \circ f)'(x) = g'(f(x)) \cdot f'(x) = (-\sin(7x^2)) \cdot (14x) = -14x \sin(7x^2).$$

For another example, suppose we have two functions $f(x)$ and $g(x)$ which satisfy the conditions of the Chain Rule. Suppose further that we have the following table of values:

x	$f(x)$	$f'(x)$	$g(x)$	$g'(x)$
1	4	3	-2	7
2	6	2	0	9
3	8	1	4	5
4	8	1	3	4
5	7	-1	2	4

Even though we cannot find the formula for $(g \circ f)(x)$ using this table, we know that $(g \circ f)(x)$ exists and is differentiable by the Chain Rule, and we can find $(g \circ f)'(1)$ by plugging in values from the table into the formula from the Chain Rule:

$$(g \circ f)'(1) = g'(f(1)) \cdot f'(1) = g'(4) \cdot f'(1) = 4 \cdot 3 = 12.$$

Notice that we did not simply multiply $g'(1)$ and $f'(1)$ together. This example should demonstrate to you that you do not need formulae for your functions in order to apply the Chain Rule at a point.

Applying the Chain Rule

The next skill we need to discuss is using the Chain Rule when we are given a function which is the composition of two other functions. Again, the best way to learn this skill is by doing many examples.

First, let $h(x) = \cos(\sin x)$. It should be clear to you that this function is the composition of two other functions $f(x)$ and $g(x)$. Specifically, if $f(x) = \sin x$ and $g(x) = \cos x$, then $h(x) = (g \circ f)(x)$, as detailed below:

$$(g \circ f)(x) = g(f(x)) = \cos(f(x)) = \cos(\sin x) = h(x).$$

Since we know that $h(x)$ is the composition of $g(x)$ after $f(x)$, and we know how to differentiate $g(x)$ and $f(x)$, we can apply the Chain Rule to find the derivative of $h(x)$:

$$h'(x) = (g \circ f)'(x) = g'(f(x)) \cdot f'(x) = -\sin(f(x)) \cdot \cos x = -\sin(\sin x) \cdot \cos x = -(\cos x)(\sin(\sin x)).$$

Next, let $h(x) = (5x^2 + 6x - 9)^{11}$. We already know how to differentiate this function using the Power Rule, but let us use the Chain Rule instead here (the Power Rule is simply a special case of the Chain Rule). First, find two functions $f(x)$ and $g(x)$ such that $h(x) = (g \circ f)(x)$. The function $f(x)$ will be the inner function, the first function applied, and in this case we get that a formula for $f(x)$ is the expression being raised to the eleventh power, that is, $f(x) = 5x^2 + 6x - 9$. A formula for $g(x)$, the outer function, is then $g(x) = x^{11}$. Both of these functions have derivatives, so, applying the Chain Rule, we get that the derivative of $h(x)$ is

$$\begin{aligned} h'(x) &= (g \circ f)'(x) = g'(f(x)) \cdot f'(x) = 11(f(x))^{10} \cdot (10x + 6) \\ &= 11(5x^2 + 6x - 9)^{10} \cdot (10x + 6) = (110x + 66)(5x^2 + 6x - 9)^{10}. \end{aligned}$$

Now let us try $h(x) = \cos((2x + 5)^2)$. This is a case where we can break up $h(x)$ into the composition of three functions. The best way to apply the Chain Rule in a case like this is to proceed from outside to inside. We demonstrate this technique below. First, we find $f(x)$ and $g(x)$ so that $h(x) = (g \circ f)(x)$ and $g(x)$ is as simple as possible. A good choice for $f(x)$ and $g(x)$ is then $f(x) = (2x + 5)^2$ and $g(x) = \cos x$. Applying the Chain Rule, we get

$$h'(x) = (g \circ f)'(x) = g'(f(x)) \cdot f'(x) = -\sin(f(x)) \cdot f'(x) = -\sin((2x + 5)^2) \cdot f'(x).$$

To complete this expression for $h'(x)$, we need to find a formula for $f'(x)$. The function $f(x)$ is the composition of two functions. Specifically, we can take $p(x) = 2x + 5$ and $q(x) = x^2$, so that $f(x) = (q \circ p)(x)$. Now we can find a formula for $f'(x)$:

$$f'(x) = (q \circ p)'(x) = q'(p(x)) \cdot p'(x) = 2(p(x)) \cdot 2 = 2(2x + 5) \cdot 2 = 8x + 20.$$

This allows us to complete the formula for $h(x)$:

$$h'(x) = -\sin((2x + 5)^2) \cdot f'(x) = -\sin((2x + 5)^2) \cdot (8x + 20) = -(8x + 20)\sin((2x + 5)^2).$$

If you are comfortable enough with the Chain Rule, you may do most of this work in your head. You do not have to explicitly find $f(x)$ and $g(x)$ and substitute into the Chain Rule formula step by step. That said, this is a complicated rule of differentiation, and it is easy to make mistakes in applying the Chain Rule. You should always be able to fall back to the step-by-step method above if it is necessary.