The Chain Rule and Applications

Composition of Functions

DEFINITION: Suppose $y = f(x)$ and $z = g(y)$ are two real-valued functions. If the intersection of the Range of f and the Domain of g is not empty, is not empty, then the **composition** $g \circ f$ **of** g **with** f is defined as

$$z = (g \circ f)(x) = g(f(x)).$$

The **domain of** $g \circ f$ consists of those values of x for which $y = f(x)$ is in domain of g.

Example

1. Let $y = f(x) = e^x$ and $z = g(y) = \ln(y)$. Find the natural domains of f and g. Find the formula for $z = g(f(x))$ and its natural domain. Find the formula for $w = f(g(y))$ and its natural domain.

2. The following function can be thought of as the composition of two other functions, f and g. Identify these other functions.

   $h(x) = \cos(2x)$

Differentiating Compositions: The Chain Rule

Recollection

Recall that by Taylor's Theorem, if f is differentiable at a, then

$$f(x) = f(a) + f'(a)(x-a) + E(x).$$

3. What are the values of the following limits?

$$\lim_{x \to a} E(x) = \qquad\qquad \lim_{x \to a} \frac{E(x)}{x-a} =$$

An Idea

Suppose you know the derivative of $y = f(x)$ at $x = a$ and the derivative of $z = g(y)$ at $y = f(a)$. The following line of reasoning suggests how we can find the derivative of $z = g(f(x))$ at $x = a$, if this exists.

On the tangent line to the graph of f at $(a, f(a))$: $\quad dy = f'(a)dx$

On the tangent line to the graph of g at $(f(a), g(f(a)))$: $\quad dz = g'(f(a))dy$

If we set the output dy on the tangent line for f equal to the input dy on the tangent line for g, then

$$dz = g'(f(a))f'(a)dx.$$

This reasoning SUGGESTS that the derivative of $z = g(f(x))$ at $x = a$ is $g'(f(a))f'(a)$.

Remark: Why does this reasoning only suggest and not prove this result?

a) We don't know for sure that a tangent line to the graph of $g \circ f$ exists at the point $(a, g(f(a)))$, and

b) even if this tangent line does exist, we don't know if dz and dx correspond to increments on this tangent line.

THEOREM: (**Chain Rule**)

Suppose $y = f(x)$ is differentiable at $x = a$ and $z = g(y)$ is differentiable at $y = f(a)$. Then the composition $z = (g \circ f)(x)$ is differentiable at $x = a$, and

$$(g \circ f)'(a) = g'(f(a))f'(a).$$

Proof: By Taylor's Theorem,

$$\begin{aligned} y &= f(x) = f(a) + f'(a)(x-a) + E_f(x), && \text{for } x \text{ near } a, \text{ and} \\ z &= g(y) = g(f(a)) + g'(f(a)) \cdot (y - f(a)) + E_g(y), && \text{for } y \text{ near } f(a), \end{aligned}$$

Therefore,

$$\begin{aligned} (g \circ f)'(a) &= \lim_{x \to a} \frac{g(f(x)) - g(f(a))}{x-a} \\ &= \lim_{x \to a} \frac{g(y) - g(f(a))}{x-a} \\ &= \lim_{x \to a} \frac{g(f(a)) + g'(f(a)) \cdot (y - f(a)) + E_g(y) - g(f(a))}{x-a} \\ &= \lim_{x \to a} \frac{g'(f(a)) \cdot (y - f(a)) + E_g(y)}{x-a} \\ &= \lim_{x \to a} \frac{g'(f(a)) \cdot (f(x) - f(a)) + E_g(f(x))}{x-a} \\ &= \lim_{x \to a} \frac{g'(f(a)) \cdot \Big(f'(a)(x-a) + E_f(x) \Big) + E_g(f(x))}{x-a} \\ &= \lim_{x \to a} \left(g'(f(a))f'(a) + g'(f(a)) \cdot \frac{E_f(x)}{x-a} + \frac{E_g(f(x))}{x-a} \right) \\ &= g'(f(a))f'(a) + g'(f(a)) \cdot \lim_{x \to a} \frac{E_f(x)}{x-a} + \lim_{x \to a} \frac{E_g(f(x))}{x-a} \\ &= g'(f(a))f'(a) + \lim_{x \to a} \frac{E_g(f(x))}{x-a}. \end{aligned}$$

Now

$$\frac{E_g(f(x))}{x-a} = \begin{cases} 0 & , \quad \text{if } f(x) = f(a), \\ \dfrac{E_g(f(x))}{f(x)-f(a)} \cdot \dfrac{f(x)-f(a)}{x-a}, & \quad \text{if } f(x) \neq f(a). \end{cases}$$

Therefore,

$$\lim_{x \to a} \frac{E_g(f(x))}{x-a} = \begin{cases} 0 & , \quad \text{if } f(x) = f(a), \\ \displaystyle\lim_{x \to a} \frac{E_g(f(x))}{f(x)-f(a)} \cdot \lim_{x \to a} \frac{f(x)-f(a)}{x-a}, & \quad \text{if } f(x) \neq f(a). \end{cases}$$

But if $f(x) \neq f(a)$, then $\displaystyle\lim_{x \to a} \frac{E_g(f(x))}{f(x)-f(a)} \cdot \lim_{x \to a} \frac{f(x)-f(a)}{x-a} = \lim_{y \to f(a)} \frac{E_g(y)}{y - f(a)} \cdot f'(a) = 0$,

so in any case, $\displaystyle\lim_{x \to a} \frac{E_g(f(x))}{x-a} = 0$. Therefore, $(g \circ f)'(a) = g'(f(a))f'(a)$.

4. Give reasons justifying each step of the proof of the Chain Theorem.

Practice with the Chain Rule

5. Find the derivative of the composite function with the formula given below. *Explicitly state the functions making up the composition*, then and apply the Chain Rule.

   $$z = (g \circ f)(x) = 3^{6x}$$

Applications of the Chain Rule

The Derivative of the Natural Logarithm Function

If we assume that $\ln(x)$ is differentiable for $x > 0$, then we can use the Chain Rule to find its formula:

$$\begin{aligned} e^{\ln(x)} &= x, \qquad x > 0 \\ \frac{d}{dx}\, e^{\ln(x)} &= \frac{d}{dx}\, x \\ e^{\ln(x)} \cdot \ln'(x) &= 1 \\ x \cdot \ln'(x) &= 1 \\ \ln'(x) &= \frac{1}{x}, \quad x > 0. \end{aligned}$$

Generalizing the Power Rule to Negative Integer Exponents

We know that if n is a positive integer, then $\frac{d}{dx}x^n = n \cdot x^{n-1}$. By applying the quotient rule, we also know that $g(y) = y^{-1} = 1/y$ is differentiable for $y \neq 0$ and has the derivative $g'(y) = -1/y^2$. Therefore, for a positive integer n,

$$\begin{aligned} x^{-n} &= (x^n)^{-1} = g(x^n), \qquad x \neq 0 \\ \frac{d}{dx}x^{-n} &= \frac{d}{dx}g(x^n) = g'(x^n) \cdot n\, x^{n-1} \\ &= \frac{-1}{(x^n)^2} \cdot n\, x^{n-1} = -n\, \frac{x^{n-1}}{x^{2n}} = -n\, x^{-n-1}, \quad x \neq 0. \end{aligned}$$

Thus the power rule applies to x^n whether n is a positive or negative integer.

Generalizing the Power Rule to Rational Nonzero Exponents

Let p, q be nonzero integers. Let $g(x) = x^{1/q}$, with the domain restricted to values of x for which this formula gives a real number. If we assume that g is differentiable, then we can use the Chain Rule to find its derivative.

$$\begin{aligned}
(g(x))^q &= (x^{1/q})^q = x^{q/q} = x^1 = x \\
\frac{d}{dx}(g(x))^q &= \frac{d}{dx}x \\
q \cdot (g(x))^{q-1} g'(x) &= 1 \\
q \cdot (x^{1/q})^{q-1} g'(x) &= 1 \\
q \cdot (x^{(q-1)/q}) g'(x) &= 1 \\
g'(x) &= \frac{1}{q \cdot (x^{1-1/q})} = \frac{1}{q} \cdot x^{\frac{1}{q}-1}.
\end{aligned}$$

Continuing, let $f(x) = x^{p/q} = (x^p)^{1/q} = g(x^p)$, with the domain restricted to values of x for which this formula gives a real number. Then by the Chain Rule, f is differentable and

$$\begin{aligned}
f(x) &= g(x^p) \\
f'(x) &= g'(x^p) \cdot \frac{d}{dx}x^p \\
&= \frac{1}{q} \cdot (x^p)^{\frac{1}{q}-1} \cdot p\,x^{p-1} \\
&= \frac{p}{q} \cdot x^{\frac{p}{q}-p+p-1} \\
&= \frac{p}{q} \cdot x^{\frac{p}{q}-1}
\end{aligned}$$

Thus the Power Rule applies to x^r, where r is any nonzero rational number:

$$\frac{d}{dx}x^r = r\,x^{r-1}, \qquad \text{rational } r \neq 0,$$

where the domain is restricted to values of x for which this formula gives a real number.

Practice with the Power Rule and the Chain Rule

Find derivatives of the following functions:

a) $f(x) = x^{3/2}$

b) $h(x) = \tan(\sqrt{\theta})$

On the Geometric Interpretation of the Chain Rule

The function $h(x) = \cos(3x+7)$ is the composition $f(g(x))$, where $f(x) = \cos(x)$ and $g(x) = 3x+7$. The Chain Rule states that $h'(a) = f'(g(a))g'(a)$, so in this case

$$h'(x) = -\sin(3x+7) \cdot 3.$$

The purpose of this note is to explain this formula *geometrically*.

1. First note that $h(x) = \cos(3x+7) = \cos(3(x+7/3))$, so the graph of $h(x)$ is just the graph of the following function shifted 7/3 units to the left:

$$c(x) = \cos(3x).$$

2. Suppose we want to find $h'(-2)$, i.e. the slope of the tangent line to the curve $y = h(x)$ at $x = -2$. The point on the graph of $y = c(x)$ that corresponds to $(-2, h(-2))$ is found by first adding 7/3 to $x = -2$, then evaluating c at that value:

$$\big(-2+7/3, c(-2+7/3)\big) = (1/3, c(1/3)).$$

3. Now the question arises: what is the slope of the tangent line to $y = c(x)$ at $x = 1/3$? In other words, what is $c'(1/3)$? To answer this question, we will use the graph of $y = \cos(x)$. The point on $y = \cos(x)$ that corresponds to $(1/3, c(1/3))$ on $y = c(x)$ is found by frist multiplying $x = 1/3$ by 3, then evaluating cos at that value:

$$\big(3 \cdot \frac{1}{3}, \cos(3 \cdot \frac{1}{3})\big) = (1, \cos(1)).$$

The slope of the line tangent to $y = \cos(x)$ at $x = 1$ is easy to determine, since the derivative of $\cos(x)$ is $-\sin(x)$: this slope is thus $-\sin(1)$.

4. However, in terms of changes along the $x-$axis, $\cos(x)$ achieves its function values at 1/3 of the rate of $\cos(3x)$. This is because the graph of $\cos(3x)$ is compressed in the x-direction relative to the graph of $\cos(x)$. To make the comparison from the other direction, $\cos(3x)$ achieves its function values 3 times as quickly as $\cos(x)$; so, at corresponding points, the rate of change of $y = \cos(3x)$ will be 3 times the rate of change of $y = \cos(x)$. Thus we multiply the value the slope of the tangent line to $y = \cos(x)$ at $x = 1$ by a factor of 3. Our conclusion is:

$$h'(-2) = 3 \cdot (-\sin(1)) = -3\sin(1).$$

5. Notice that none of the conclusions above depended specifically on $x = -2$, so we can draw the same conclusions for $x = a$:

$$h'(a) = 3 \cdot (-\sin(3(a+7/3))) = -3\ \sin(3a+7).$$