One-Dimensional Chain Rule: Demonstration

Aurelien Lourot

2014-12-30

This article demonstrates the *chain rule*, a formula for computing the derivative of the composition of two functions, first used by Leibniz (around 1700). It states that

given two real functions f and g ($\mathbb{R} \to \mathbb{R}$) and $x_0 \in \mathbb{R}$,

if g is differentiable at x_0 and f is differentiable at $g(x_0)$,

then

$$(f \circ g)'(x_0) = (f' \circ g)(x_0) \cdot g'(x_0)$$

1 Alternate notations

The *chain rule* is often written in the following more accessible way:

$$\frac{d}{dx}(f(g(x))) = f'(g(x)) \cdot g'(x) \tag{1}$$

where the free variable x_0 has been replaced by x.

The word *chain* suddenly springs into focus when using *Leibniz's notation*: if one defines $u = g(x)$ and $y = f(u)$,

- f' represents the variation of y due to the variation of u, i.e. $\frac{dy}{du}$;
- g' represents the variation of u due to the variation of x, i.e. $\frac{du}{dx}$.

Substituting these terms in (1) results in

$$\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx}$$

2 Demonstration

The definition of g being differentiable at x_0 is

$$\lim_{a \to 0} \frac{g(x_0 + a) - g(x_0)}{a}$$

exists. Let's name it b_0.

$$\Leftrightarrow \exists h : \mathbb{R} \to \mathbb{R}, \forall a \in \mathbb{R}^*, \frac{g(x_0 + a) - g(x_0)}{a} = b_0 + h(a) \tag{2}$$

with

$$\lim_{a \to 0} h(a) = 0$$ [1]

$$\begin{aligned} (2) \quad &\Leftrightarrow \quad g(x_0 + a) = g(x_0) + a \cdot b_0 + a \cdot h(a) \\ &\Rightarrow \quad f(g(x_0 + a)) = f(g(x_0) + a \cdot b_0 + a \cdot h(a)) \end{aligned}$$

[1] Note that there might also be one or several non null values of a, even "far" from 0, for which $h(a)$ is exactly null.

Given $k(a) = a \cdot b_0 + a \cdot h(a)$, we have

$$f(g(x_0 + a)) = f(g(x_0) + k(a)) \tag{3}$$

Thanks to f being differentiable at $g(x_0)$, we know that **for** $k(a)$ **non null**

$$\lim_{k(a)\to 0} \frac{f(g(x_0) + k(a)) - f(g(x_0))}{k(a)}$$

exists. Let's name it c_0.

$$\Leftrightarrow \exists l : \mathbb{R} \to \mathbb{R}, \frac{f(g(x_0) + k(a)) - f(g(x_0))}{k(a)} = c_0 + l(k(a)) \tag{4}$$

when $k(a) \neq 0$ with

$$\lim_{b\to 0} l(b) = 0 \quad ^{(2} \tag{5}$$

When $k(a) \neq 0$, $\quad (4) \quad \Leftrightarrow \quad f(g(x_0) + k(a)) = f(g(x_0)) + k(a) \cdot c_0 + k(a) \cdot l(k(a))$

Let's define $m : \mathbb{R} \to \mathbb{R}$ as

$$m(c) = \begin{cases} l(c), & c \neq 0; \\ 0, & c = 0. \end{cases}$$

We now have, for any value of $k(a)$, $f(g(x_0) + k(a)) = f(g(x_0)) + k(a) \cdot c_0 + k(a) \cdot m(k(a))$

Substituting this result into (3) gives

$$f(g(x_0 + a)) = f(g(x_0)) + k(a) \cdot c_0 + k(a) \cdot m(k(a))$$

$$\Leftrightarrow f(g(x_0 + a)) - f(g(x_0)) = [a \cdot b_0 + a \cdot h(a)] \cdot c_0 + [a \cdot b_0 + a \cdot h(a)] \cdot m(k(a))$$

$$\Leftrightarrow \frac{f(g(x_0 + a)) - f(g(x_0))}{a} = b_0 \cdot c_0 + h(a) \cdot c_0 + [b_0 + h(a)] \cdot m(k(a))$$

$$\Rightarrow \lim_{a\to 0} \frac{f(g(x_0 + a)) - f(g(x_0))}{a} = b_0 \cdot c_0$$

Substituting b_0 and c_0 by their respective definitions gives

$$\lim_{a\to 0} \frac{f(g(x_0 + a)) - f(g(x_0))}{a} = \left(\lim_{a\to 0} \frac{g(x_0 + a) - g(x_0)}{a}\right) \cdot \left(\lim_{k(a)\to 0} \frac{f(g(x_0) + k(a)) - f(g(x_0))}{k(a)}\right) \tag{6}$$

Note that all three limit arguments (a, a and $k(a)$) are independant free variables, so that (6) can be re-written as

$$\lim_{a\to 0} \frac{f(g(x_0 + a)) - f(g(x_0))}{a} = \left(\lim_{b\to 0} \frac{g(x_0 + b) - g(x_0)}{b}\right) \cdot \left(\lim_{c\to 0} \frac{f(g(x_0) + c) - f(g(x_0))}{c}\right)$$

$$\Leftrightarrow \boxed{(f \circ g)'(x_0) = g'(x_0) \cdot f'(g(x_0))}$$

■

[2] $(5) \Rightarrow \lim_{a\to 0} l(k(a)) = 0$