Chain rule

Demonstration

Abdelwahid Benslimane

wahid.benslimane@gmail.com

The derivative of a compound function $g \circ f(x) = g(f(x))$ is obtained by the following formula: $(g \circ f)'(x) = g'(f(x)).\, f'(x)$. This is what we call the chain rule

This is something we see and manipulate from a very young age, and it is on the same principle that the mechanism of the backpropagation of the gradient in an artificial neural network is based. I propose to demonstrate this formula.

We assume that the respective definition sets of f and g are I and J. It is also assumed that $f(I) \subset J$.

We suppose that the conditions are met so that the derivative of $g \circ f$ exists at the point x which belongs to I.

Demonstration

$$\begin{aligned}
(g \circ f)'(x) &= \lim_{h \to 0} f(x) \frac{g(f(x+h)) - g(f(x))}{h} \\
&= \lim_{h \to 0} f(x) \frac{g(f(x+h)) - g(f(x))}{h} \times \frac{f(x+h) - f(x)}{f(x+h) - f(x)} \\
&= \lim_{h \to 0} f(x) \frac{g(f(x+h)) - g(f(x))}{f(x+h) - f(x)} \times \frac{f(x+h) - f(x)}{h} \\
&= \lim_{h \to 0} f(x) \frac{g(f(x+h)) - g(f(x))}{f(x+h) - f(x)} \times \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} \\
&= \lim_{h \to 0} f(x) \frac{g(f(x+h)) - g(f(x))}{f(x+h) - f(x)} \times f'(x).
\end{aligned}$$

If we write $k = f(x+h) - f(x)$, it is obvious that as h tends to 0, k tends to 0. So studying the limit when h tends to 0, is the same as studying the limit when k tends to 0 (after changing the variable). Therefore, it comes that:

$$\lim_{h \to 0} f(x) \frac{g(f(x+h)) - g(f(x))}{f(x+h) - f(x)} = \lim_{k \to 0} f(x) \frac{g(f(x)+k) - g(f(x))}{k} = g'(f(x)).$$

We then obtain:

$$(g \circ f)'(x) = \lim_{h \to 0} f(x) \frac{g(f(x+h)) - g(f(x))}{f(x+h) - f(x)} \times f'(x) = g'(f(x)).\, f'(x).$$

We have just demonstrated that $(g \circ f)'(x) = g'(f(x)).\, f'(x)$.