planetmath.org

Math for the people, by the people.

proof of chain rule (several variables)

Canonical name	ProofOfChainRuleseveralVariables
Date of creation	2013-03-22 16:05:07
Last modified on	2013-03-22 16:05:07
Owner	paolini (1187)
Last modified by	paolini (1187)
Numerical id	6
Author	paolini (1187)
Entry type	Proof
Classification	msc 26B12

We first consider the case $m = 1$ i.e. $G\colon I \to \mathbb{R}^n$ where $I \subset \mathbb{R}$ is a neighbourhood of a point $x_0 \in \mathbb{R}$ and $F\colon U \subset \mathbb{R}^n \to \mathbb{R}$ is defined on a neighbourhood U of $y_0 = G(x_0)$ such that $G(I) \subset U$. We suppose that both G is differentiable at the point x_0 and F is differentiable in y_0. We want to compute the derivative of the compound function $H(x) = F(G(x))$ at $x = x_0$.

By the definition of derivative (using Landau notation) we have

$$F(y_0 + k) = F(y_0) + DF(y_0)k + o(|k|).$$

Choose any $h \neq 0$ such that $x_0 + h \in I$ and set $k = G(x_0 + h) - G(x_0)$ to obtain

$$\begin{aligned}
\frac{H(x_0+h) - H(x_0)}{h} &= \frac{F(G(x_0+h)) - F(G(x_0))}{h} \\
&= \frac{F(G(x_0) + k) - F(G(x_0))}{h} = \frac{F(y_0 + k) - F(y_0)}{h} \\
&= \frac{DF(y_0)(G(x_0+h) - G(x_0)) + o(|G(x_0+h) - G(x_0)|)}{h} \\
&= DF(y_0)\frac{G(x_0+h) - G(x_0)}{h} + \frac{o(|G(x_0+h) - G(x_0)|)}{h}.
\end{aligned}$$

Letting $h \to 0$ the first term of the sum converges to $DF(y_0)G'(x_0)$ hence we want to prove that the second term converges to 0. Indeed we have

$$\left|\frac{o(|G(x_0+h) - G(x_0)|)}{h}\right| = \left|\frac{o(|G(x_0+h) - G(x_0)|)}{|G(x_0+h) - G(x_0)|}\right| \cdot \left|\frac{G(x_0+h) - G(x_0)}{h}\right|.$$

By the definition of $o(\cdot)$ the first fraction tends to 0, while the second fraction tends to the absolute value of $G'(x_0)$. Thus the product tends to 0, as needed.

Consider now the general case $G\colon V \subset \mathbb{R}^m \to U \subset \mathbb{R}^n$. Given $v \in \mathbb{R}^m$ we are going to compute the directional derivative

$$\frac{\partial F \circ G}{\partial v}(x_0) = \frac{dF \circ g}{dt}(0)$$

where $g(t) = G(x_0 + tv)$ is a function of a single variable $t \in \mathbb{R}$. Thus we fall back to the previous case and we find that

$$\frac{\partial F \circ G}{\partial v}(x_0) = DF(G(x_0))g'(0). = DF(G(x_0))\frac{\partial G}{\partial v}(x_0)$$

In particular when $v = e_k$ is the k-th coordinate vector, we find

$$g'(0) = D_{x_k} F \circ G(x_0) = DF(G(x_0)) D_{x_k} G(x_0) = \sum_{i=1}^{n} D_{y_i} G(x_0) D_{x_k} G^i(x_0)$$

which can be compactly written

$$DF \circ G(x_0) = DF(G(x_0)) DG(x_0).$$