planetmath.org

Math for the people, by the people.

chain rule (several variables)

Canonical name	ChainRuleseveralVariables
Date of creation	2013-03-22 12:33:10
Last modified on	2013-03-22 12:33:10
Owner	rmilson (146)
Last modified by	rmilson (146)
Numerical id	10
Author	rmilson (146)
Entry type	Theorem
Classification	msc 26B12
Related topic	ChainRule
Related topic	Jacobian
Related topic	JacobianMatrix

The chain rule is a theorem of analysis that governs derivatives of composed functions. The basic theorem is the chain rule for functions of one variables (`http://planetmath.org/ChainRule`see here). This entry is devoted to the more general version involving functions of several variables and partial derivatives. Note: the symbol D_k will be used to denote the partial derivative with respect to the kth variable.

Let $F(x_1, \ldots, x_n)$ and $G_1(x_1, \ldots, x_m), \ldots, G_n(x_1, \ldots, x_m)$ be differentiable functions of several variables, and let

$$H(x_1, \ldots, x_m) = F(G_1(x_1, \ldots, x_m), \ldots, G_n(x_1, \ldots, x_m))$$

be the function determined by the composition of F with $G_1, \ldots, G_n$ The partial derivatives of H are given by

$$(D_k H)(x_1, \ldots, x_m) = \sum_{i=1}^{n} (D_i F)(G_1(x_1, \ldots, x_m), \ldots)(D_k G_i)(x_1, \ldots, x_m).$$

The chain rule can be more compactly (albeit less precisely) expressed in terms of the Jacobi-Legendre partial derivative symbols (`http://members.aol.com/jeff570/calc` note). Just as in the Leibniz system, the basic idea is that of one quantity (i.e. variable) depending on one or more other quantities. Thus we would speak about a variable z depends differentiably on $y_1, \ldots, y_n$, which in turn depend differentiably on variables $x_1, \ldots, x_m$. We would then write the chain rule as

$$\frac{\partial z}{\partial x_j} = \sum_{i=1}^{n} \frac{\partial z}{\partial y_i} \frac{\partial y_i}{\partial x_j}, \qquad j = 1, \ldots m.$$

The most general, and conceptually clear approach to the multi-variable chain is based on the notion of a differentiable mapping, with the Jacobian matrix of partial derivatives playing the role of generalized derivative. Let, $X \subset \mathbb{R}^m$ and $Y \subset \mathbb{R}^n$ be open domains and let

$$\mathbf{F} : Y \to \mathbb{R}^l, \qquad \mathbf{G} : X \to Y$$

be differentiable mappings. In essence, the symbol $\mathbf{F}$ represents l functions of n variables each:

$$\mathbf{F} = (F_1, \ldots, F_l), \qquad F_i = F_i(x_1, \ldots, x_n),$$

whereas $\mathbf{G} = (G_1, \ldots, G_n)$ represents n functions of m variables each. The derivative of such mappings is no longer a function, but rather a matrix of partial derivatives, customarily called the Jacobian matrix. Thus

$$D\mathbf{F} = \begin{pmatrix} D_1 F_1 & \ldots & D_n F_1 \\ \vdots & \ddots & \vdots \\ D_1 F_l & \ldots & D_n F_l \end{pmatrix} \qquad D\mathbf{G} = \begin{pmatrix} D_1 G_1 & \ldots & D_m G_1 \\ \vdots & \ddots & \vdots \\ D_1 G_n & \ldots & D_m G_n \end{pmatrix}$$

The chain rule now takes the same form as it did for functions of one variable:

$$D(\mathbf{F} \circ \mathbf{G}) = ((D\mathbf{F}) \circ \mathbf{G}) (D\mathbf{G}),$$

albeit with matrix multiplication taking the place of ordinary multiplication.

This form of the chain rule also generalizes quite nicely to the even more general setting where one is interested in describing the derivative of a composition of mappings between manifolds.