

The Chain Rule

There is only one Chain Rule for a function of several variables. Assume $f(x_1, x_2, x_3, \dots, x_n)$ is a function of n variables. Then assume for each fixed i such that $1 \leq i \leq n$ that $x_i = g_i(t_1, t_2, t_3, \dots, t_m)$ where g_i is a function of m variables. Then f may also be considered as a function of the m variables $t_1, t_2, t_3, \dots, t_m$ and thus has m distinct partial derivatives with respect to those variables. In general, for each fixed j such that $1 \leq j \leq m$ we may write

$$\frac{\partial f(g_1, g_2, g_3, \dots, g_n)}{\partial t_j} = \sum_{k=1}^n \frac{\partial f(x_1, x_2, x_3, \dots, x_n)}{\partial x_k} \left| \begin{array}{l} x_1 = g_1(t_1, t_2, \dots, t_m) \\ x_2 = g_2(t_1, t_2, \dots, t_m) \\ x_3 = g_3(t_1, t_2, \dots, t_m) \\ \vdots \\ x_n = g_n(t_1, t_2, \dots, t_m) \end{array} \right. \cdot \frac{\partial g_k(t_1, t_2, t_3, \dots, t_m)}{\partial t_j}$$

When either n or m is 1 then it is more appropriate to write special cases of the above formula in which one or more partial derivatives get replaced by regular derivatives. There are exactly three special cases that can arise in practice.

When $n = 1$ and $m \geq 2$ and we let $x = x_1$ and we let $g = g_1$ the above expression simplifies to

$$\frac{\partial f(g(t_1, t_2, t_3, \dots, t_m))}{\partial t_j} = \frac{df(x)}{dx} \Big|_{x=g(t_1, t_2, t_3, \dots, t_m)} \cdot \frac{\partial g(t_1, t_2, t_3, \dots, t_m)}{\partial t_j}$$

When $n \geq 2$ and $m = 1$ and we let $t_1 = x$ and we let $x_i = g_i(x)$ then the first expression above simplifies to

$$\frac{df(g_1, g_2, g_3, \dots, g_n)}{dx} = \sum_{k=1}^n \frac{\partial f(x_1, x_2, x_3, \dots, x_n)}{\partial x_k} \left| \begin{array}{l} x_1 = g_1(x) \\ x_2 = g_2(x) \\ x_3 = g_3(x) \\ \vdots \\ x_n = g_n(x) \end{array} \right. \cdot \frac{dg_k(x)}{dx}$$

When $n = 1$ and $m = 1$ and we let $x_1 = x = g(t)$ the above expression simplifies to

$$\frac{df(g(t))}{dt} = \frac{df(x)}{dx} \Big|_{x=g(t)} \cdot \frac{dx}{dt} = f'(g(t)) \cdot g'(t)$$

which is the normal Chain Rule for a function of one variable.