The Chain Rule

The Chain Rule for functions of one variable says that if $y = y(x)$ and $f = f(y)$ are functions of one variable and $F(x) = f(y(x))$ is the composition of f and g, then F is differentiable at x if y is differentiable at x and f is continuously diffrentiable at $y(x)$, then F is differentiable at x and the derivative $F'(x)$ is given by

$$F'(x) = f'(y(x)).y'(x),$$

or

$$\frac{dF}{dx}(x) = \frac{df}{dy}(y(x))\frac{dy}{dx}(x).$$

There is a similar formula for functions of several variables. If $y_i = y(x_1, \cdots x_n)$ is a function of n variables for $1 \leq i \leq m$ and $f = f(y_1, \cdots y_m)$ is a function of m variables and F is the composition:

$$F(x_1, \cdots x_n) = F(y_1(\underline{x}, \cdots y_m(\underline{x}))$$

where $\underline{x} = (x_1, \cdots x_n)$, $\underline{y}(\underline{x}) = (y_1(\underline{x}, \cdots \underline{y}_m(\underline{x}))$ and $\frac{\partial f}{\partial y_i}(\underline{y}(x))$ exist and are continuous for $1 \leq i \leq m$, and $\frac{\partial f}{\partial x_j}(\underline{x})$ exist for $1 \leq j \leq n$, then $\frac{\partial F}{\partial x_j}(\underline{x})$ exist for $1 \leq j \leq n$ and

$$\frac{\partial F}{\partial x_j}(\underline{x}) = \sum_{i=1}^{m} \frac{\partial f}{\partial y_i}(\underline{y}(\underline{x}))\frac{\partial y_i}{\partial x_j}(\underline{x}).$$

Gradient

If $f = f(x, y)$ is differentiable then the *gradient* $\nabla f(x_0, y_0)$ of f at (x_0, y_0) is given by

$$\nabla f(x_0, y_0) = \frac{\partial f}{\partial x}(x_0, y_0)\mathbf{i} + \frac{\partial f}{\partial y}(x_0, y_0)\mathbf{j}.$$

Similarly if $f = f(x, y, z)$ then

$$\nabla f(x_0, y_0, z_0) = \frac{\partial f}{\partial x}(x_0, y_0, z_0)\mathbf{i} + \frac{\partial f}{\partial y}(x_0, y_0, z_0)\mathbf{j}$$

$$+\frac{\partial f}{\partial z}(x_0, y_0, z_0)\mathbf{k}.$$

There is a similar formula if f is a function of n variables

Directional derivatives

If $\underline{v}$ is any vector (in $\mathbb{R}^2$ or $\mathbb{R}^3$ or $\mathbb{R}^n$) and F is a function of 2 or 3 or n variables, then the *directional derivative of F at $\underline{a}$ in the direction $\underline{v}$* is $f'(0)$ where

$$f(t) = F\left(\underline{a} + t\frac{\underline{v}}{|\underline{v}|}\right),$$

where $|v|$ is the length of v. So if $|v| = 1$ and $\underline{v} = (v_1, v_2) \in \mathbb{R}^2$, for example, the chain rule gives

$$f'(t) = v_1 \frac{\partial F}{\partial x}(\underline{a} + t\underline{v}) + v_2 \frac{\partial F}{\partial y}(\underline{a} + t\underline{v}).$$

Then

$$f'(0) = v_1 \frac{\partial F}{\partial x}(\underline{a}) + v_2 \frac{\partial F}{\partial y}(\underline{a}) = \underline{v}.\nabla F(\underline{a}).$$

In general, if $\underline{v} = (v_1, v_2)$ does not necessarily have unit length, we get

$$f'(0) = \frac{\underline{v}.\nabla F(\underline{a})}{|\underline{v}|}.$$

So this is the general formula for the directional derivative in the direction of $\underline{v}$, if $\underline{v}$ is a vector in $\mathbb{R}^2$. There is a similar formula if $\underline{v}$ is a vector in $\mathbb{R}^3$.

Directions of maximal increase and decrease

If F is a function of two or three (or more) variables the directions of maximal increase and decrease of F at $\underline{a}$ are the directions of the vectors $\underline{v}$ for which

$$\frac{\underline{v}.\nabla F(\underline{a})}{|\underline{v}|}$$

is as large as possible and as negatively large as possible. Now

$$\frac{\underline{v}.\nabla F(\underline{a})}{|\underline{v}|} = |\nabla F(\underline{a})| \cos\theta$$

where θ is the angle between $\nabla F(\underline{a})$ and $\underline{v}$. But $-1 \leq \cos\theta \leq 1$ and $\cos\theta = 1$ if and only if $\underline{v}$ is a positive multiple of $\nabla F(\underline{a})$ and $\cos\theta = -1$ if and only if $\underline{v}$ is a negative multiple of $\nabla F(\underline{a}$. So the derivatives in the directions of maximal increase and decrease are

$$|\nabla F(\underline{a})|$$

and

$$-|\nabla F(\underline{a})|.$$

Normals and Tangent Planes

Let $f = f(x, y, z)$ be a function of three variables. Then the set of (x, y, z) such that

$$f(x, y, z) = c$$

is a surface. At any point (x_0, y_0, z_0) in the surface, $\nabla f(x_0, y_0, z_0)$ is in the direction of the normal to the surface so the normal line to the surface at this point is given by

$$\mathbf{r}(t) = x_0\mathbf{i} + y_0\mathbf{j} + z_0\mathbf{k} + t\nabla f(x_0, y_0, z_0).$$

The tangent plane to the surface at this point is

$$\nabla \mathbf{f}(x_0, y_0, z_0).((x - x_0)\mathbf{i} + (y - y_0)\mathbf{j} + (z - z_0)\mathbf{k}) = 0.$$

The reason for this is that if (x, y, z) is close to (x_0, y_0, z_0) on the surface,

$$0 = c - c = f(x, y, z) - f(x_0, y_0, z_0)$$

$$\approx f(x_0, y_0, z_0) + \nabla f(x_0, y_0, z_0).((x - x_0)\mathbf{i} + (y - y_0)\mathbf{j} + (z - z_0)\mathbf{k}) - f(x_0, y_0, z_0).$$

So

$$0 \approx \nabla f(x_0, y_0, z_0).((x - x_0)\mathbf{i} + (y - y_0)\mathbf{j} + (z - z_0)\mathbf{k})$$

This means that $\nabla f(x_0, y_0, z_0)$ is perpendicular to the tangent plane at (x_0, y_0, z_0), that is, is in the direction of the normal.