The Chain Rule and Directional Derivatives

April 7, 2010

The chain rule in two variables

$f : X \subseteq \mathbb{R}^2 \to \mathbb{R}$ differentiable at $\mathbf{x_0} = (a, b)$
$\mathbf{x} : T \subseteq \mathbb{R} \to \mathbb{R}^2$ differentiable at $t = t_0$.

$$\boxed{\frac{df}{dt}(t_0) = \frac{\partial f}{\partial x}(\mathbf{x}_0)\frac{dx}{dt}(t_0) + \frac{\partial f}{\partial y}(\mathbf{x_0})\frac{dy}{dt}(t_0)}$$

This can be rewritten (vector notation):

$$\frac{df}{dt}(t_0) = \left(\frac{\partial f}{\partial x}(\mathbf{x}_0), \frac{\partial f}{\partial y}(\mathbf{x_0})\right) \cdot \left(\frac{dx}{dt}(t_0), \frac{dy}{dt}(t_0)\right)$$

Or using the gradient:

$$\boxed{\frac{df}{dt}(t_0) = \nabla f(\mathbf{x_0}) \cdot \mathbf{x}'(t_0)}$$

Generalization to functions $f : X \subseteq \mathbb{R}^n \to \mathbb{R}$

Let $\mathbf{x} : T \subseteq \mathbb{R} \to \mathbb{R}^n$ and $f : X \subseteq \mathbb{R}^n \to \mathbb{R}$

$$\boxed{\frac{df}{dt}(t_0) = \nabla f(\mathbf{x_0}) \cdot \mathbf{x}'(t_0)}$$

And in matrix notation:

$$\boxed{\frac{df}{dt}(t_0) = Df(\mathbf{x_0}) D\mathbf{x}(t_0)}$$

The general chain rule

Let $f : X \subseteq \mathbb{R}^m \longrightarrow \mathbb{R}^p$ and $\mathbf{x} : T \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^m$

$$\boxed{D(f \circ \mathbf{x})(\mathbf{t_0}) = Df(\mathbf{x_0})D\mathbf{x}(\mathbf{t_0})}$$

Here: $\mathbf{x_0} = (x_1(\mathbf{t_0}), x_2(\mathbf{t_0}), \ldots, x_n(\mathbf{t_0}))$.

The gradient

Let $f : X \subseteq \mathbb{R}^n \to \mathbb{R}$ be a scalar valued function. Then the gradient

$$\boxed{\nabla f = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n}\right).}$$

Directional Derivative

Let f be a differentiable function and $\mathbf{a}$ be a point in the domain of f then

$$\boxed{D_{\mathbf{v}} f(\mathbf{a}) = \nabla f(\mathbf{a}) \cdot \mathbf{v}}$$

where $\mathbf{v}$ is a unit vector.

Maximum and minimum values of $D_{\mathbf{v}}f(\mathbf{a})$

- $D_{\mathbf{v}}f(\mathbf{a})$ is maximized when $\mathbf{v}$ points in the **same direction** of the gradient, $\nabla f(\mathbf{a})$.

- $D_{\mathbf{v}}f(\mathbf{a})$ is minimized when $\mathbf{v}$ points in the **opposite direction** of the gradient, $-\nabla f(\mathbf{a})$.

- Furthermore, the maximum and minimum values of $D_{\mathbf{v}}f(\mathbf{a})$ are $\|\nabla f(\mathbf{a})\|$ and $-\|\nabla f(\mathbf{a})\|$, respectively.

Tangent planes to level surfaces: $f(\mathbf{x}) = c$

Let c be any constant and $f : X \subseteq \mathbb{R}^3 \to \mathbb{R}$

If $\mathbf{x}_0$ is a point on the level surface

$$f(\mathbf{x}) = c,$$

then the vector $\nabla f(\mathbf{x}_0)$ is perpendicular to the level surface at $\mathbf{x}_0$.

Computing Tangent plane for level surfaces

Given the equation of a level surface

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

and a point $\mathbf{x}_0 = (x_0, y_0, z_0)$, then the equation of the tangent plane is

$$\boxed{\nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) = 0}$$

or if $\mathbf{x}_0 = (x_0, y_0, z_0)$ then

$$f_x(\mathbf{x}_0)(x - x_0) + f_y(\mathbf{x}_0)(y - y_0) + f_z(\mathbf{x}_0)(z - z_0) = 0.$$