Chain Rule and the Gradient

November 19, 2007

The gradient

Definition: Let f be a function of n variables: $x_1, x_2, \ldots, x_n$, then the **gradient** is

$$\boxed{\nabla f(x_1, \ldots, x_n) = \langle \frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2} \ldots, \frac{\partial f}{\partial x_n} \rangle}$$

Linear approximations for functions of n variables

For a function of n variables we can use this notation to write the linear approximation at $\mathbf{a} = (a_1, a_2, \ldots, a_n)$

$$\boxed{L(x_1, \ldots, x_n) = f(\mathbf{a}) + \nabla f(\mathbf{a}) \cdot \langle x_1 - a_1, \ldots, x_n - a_n \rangle}$$

Definition of differentiable

Definition: We say that a function of n variables, f is **differentiable** at a point $\mathbf{a} = (a_1, a_2, \ldots, a_n)$ if all the partial derivatives $\frac{\partial f}{\partial x_i}(\mathbf{a})$ exist and the linear approximation $L(x_1, \ldots, x_n)$ is a **good approximation**. I.e.

$$\lim_{(x_1,\ldots,x_n)\to\mathbf{a}} \frac{f(x_1, \ldots, x_n) - L(x_1, \ldots, x_n)}{\|(x_1, \ldots, x_n) - (a_1, \ldots, a_n)\|} = 0$$

Chain Rule - Case 1

Suppose that $z = f(x, y)$ is a differentiable function of two variables and $x = g(t)$ and $y = h(t)$ are both differentiable functions of t. Then the derivative of f with respect to t is:

$$\frac{df}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}$$

$$\boxed{\frac{df}{dt} = \nabla f \cdot \langle \frac{dx}{dt}, \frac{dy}{dt} \rangle}$$

Chain Rule - Case 2

Suppose that $z = f(x, y)$ is a differentiable function of x and y and $x = g(s, t)$ and $y = h(s, t)$ are differentiable functions of s and t. Then

$$\begin{aligned} \frac{\partial f}{\partial s} &= \frac{\partial f}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial s} \\ \frac{\partial f}{\partial t} &= \frac{\partial f}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial t} \end{aligned}$$

Or in vector notation

$$\frac{\partial f}{\partial s} = \nabla f \cdot \langle \frac{\partial x}{\partial s}, \frac{\partial y}{\partial s} \rangle \quad \text{and} \quad \frac{\partial f}{\partial t} = \nabla f \cdot \langle \frac{\partial x}{\partial t}, \frac{\partial y}{\partial t} \rangle$$

Chain Rule in general

Suppose f is a differentiable function in n variables: $x_1, x_2, \ldots, x_n$ and each x_i is a differentiable function in m variables $t_1, t_2, \ldots, t_m$. Then f is a function of $t_1, t_2, \ldots, t_m$ and for each $i = 1, \cdots, m$:

$$\frac{\partial f}{\partial t_i} = \frac{\partial f}{\partial x_1}\frac{\partial x_1}{\partial t_i} + \frac{\partial f}{\partial x_2}\frac{\partial x_2}{\partial t_i} + \cdots + \frac{\partial f}{\partial x_n}\frac{\partial x_n}{\partial t_i}$$

In vector notation

$$\frac{\partial f}{\partial t_i} = \nabla f \cdot \langle \frac{\partial x_1}{\partial t_i}, \frac{\partial x_2}{\partial t_i}, \ldots, \frac{\partial x_n}{\partial t_i} \rangle$$

Implicit Differentiation

Suppose that y is defined implicitly by the equation

$$F(x, y) = 0$$

Then

$$\boxed{\frac{dy}{dx} = -\frac{F_x}{F_y}}$$