Section Summary: 14.6
Directional Derivatives and the Gradient Vector

a. Definitions

The **directional derivative** of f at (x_0, y_0) in the direction of unit vector $\mathbf{u} = \langle a, b \rangle$ is

$$D_{\mathbf{u}} f(x_0, y_0) = \lim_{h \to 0} \frac{f(x_0 + ha, y_0 + hb) - f(x_0, y_0)}{h}$$

if this limit exists.

If f is a function of two variables x and y, then the **gradient** of f is the vector function ∇f defined by

$$\nabla f(x, y) = \langle f_x(x, y), f_y(x, y) \rangle = \frac{\partial f}{\partial x} \hat{i} + \frac{\partial f}{\partial y} \hat{j}$$

Hence,

$$D_{\mathbf{u}} f(x, y) = \nabla f(x, y) \cdot \mathbf{u}$$

More generally (in higher dimensions),

$$D_{\mathbf{u}} f(\mathbf{x}_0) = \lim_{h \to 0} \frac{f(\mathbf{x}_0 + h\mathbf{u}) - f(\mathbf{x}_0)}{h}$$

(Isn't that a beautiful analogy with the univariate derivative?)

b. Theorems

If f is a differentiable function of x and y, then f has a directional derivative in the direction of any unit vector $\mathbf{u} = \langle a, b \rangle$ and

$$D_{\mathbf{u}} f(x, y) = f_x(x, y)a + f_y(x, y)b$$

Suppose f is a differentiable multivariate function. The maximum value of the directional derivative $D_{\mathbf{u}} f(\mathbf{x})$ is $|\nabla f(\mathbf{x})|$ and it occurs when $\mathbf{u}$ has the same direction as the gradient vector $\nabla f(\mathbf{x})$.

c. Properties/Tricks/Hints/Etc.

The gradient vector will be very important in optimization problems.

d. Summary

We shouldn't be a stick in the mud about the orientation of our axes: why x and y, and not some other pair of directions which are mutually perpendicular? Perhaps we are interested in the slope of the surface along some direction other than x or y: hence the idea behind directional derivatives. At a given point at which a function is differentiable, one natural choice for two directions might be the direction in which the function is increasing fastest, and the direction perpendicular to this.

Because the function is essentially planer at a differentiable point, it might not be surprising that the partial derivative of f wrt (with respect to) a given direction *other than* x and y is some linear combination of the partial derivatives of f wrt x and y.

The gradient vector is the vector with these two partials (wrt x and y) weighting the respective axes unit vectors $\hat{\imath}$ and $\hat{\jmath}$, and is used to give a convenient inner product definition of the directional derivative. It points in the direction of fastest increase of the function, which simultaneously means that it is perpendicular to level curves.

The importance of this vector cannot be overestimated: it is used for minimization, as the derivative is used in the univariate case. It is via the gradient that we are able to minimize multivariate functions.