14.6 Directional Derivatives and the Gradient Vector

It is natural to ask about the rate of change of a function $f(x, y)$ as the arguments change in any direction around a point (x_0, y_0) not just along the coordinate axes, and to ask questions like in which direction is change the fastest.

The cautionary example in Section 14.5 shows that the partial derivatives do not always answer this question, but when a function is *differentiable* at the point, it is well approximated by the tangent plane there, and the linear approximation becomes a good candidate for giving information about the rate of change in various directions, from the partial derivatives alone.

Directional Derivatives

A direction of change in the plane can be specified by a unit vector $\vec{u} = \langle c, d\rangle$, and we can consider how $f(x, y)$ changes in this direction near (x_0, y_0) looking at a "slice" of the function, along the line $\langle x_0, y_0\rangle + t\vec{u}$. The value of the function along this line is $f(x_0 + ct, y_0 + dt)$ and its rate of change is given by the Chain Rule as

$$\frac{df}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} = f_x(x_0, y_0)c + f_y(x_0, y_0)d$$

This is the **directional derivative of** f **at** (x_0, y_0), denoted $D_{\vec{u}}f(x_0, y_0)$, and it does indeed depend only on the partial derivatives at the point. Note that its value is the same as if one used the linear approximation T at that point in place of f.

Directional Derivatives Defined With Limits

Equivalently, the directional derivative can be defined in terms of limits as

$$D_{\vec{u}}f(x_0, y_0) = \lim_{h\to 0} \frac{f(x_0 + ch, y_0 + dh) - f(x_0, y_0)}{h}$$

This exists when f is differentiable at the point (x_0, y_0).

To summarize, for any unit vector $\vec{u} = \langle u_1, u_2\rangle$ (or indeed any non-zero vector) and any function f differentiable at (x_0, y_0), the directional derivative of f at (x_0, y_0) in direction $\vec{u}$ is

$$D_{\vec{u}}f(x_0, y_0) = f_x(x_0, y_0)u_1 + f_y(x_0, y_0)u_2 = \vec{u} \cdot \langle f_x(x_0, y_0), f_y(x_0, y_0)\rangle$$

See Example 2.

The Gradient Vector

The vector function $\langle f_x(x_0, y_0), f_y(x_0, y_0)\rangle$ appearing in the formula for the directional derivative encapsulates all information about directional and partial derivatives of f at (x_0, y_0) in away that has a nice geometrical meaning.

It is called **the gradient vector of** f **at** (x_0, y_0), and denoted **grad** f or ∇f, the latter sometimes pronounced "del f". Considered as a function of position

$$\textbf{grad}\ f(x, y) = \nabla f(x, y) = \langle f_x(x, y), f_y(x, y)\rangle = \frac{\partial f}{\partial x}\vec{\imath} + \frac{\partial f}{\partial y}\vec{\jmath},$$

our first case of a vector valued function of several variables.

See Example 3.

Directional Derivatives in terms of The Gradient Vector

The directional derivative above can be written as $D_{\vec{u}}f(x,y) = \vec{u}\cdot\nabla f(x,y),$ so the gradient contains all information about directional derivatives.

See Example 4.

Tangent Planes to Level Curves

If $\dfrac{\partial F}{\partial y} \neq 0$ at point (x_0, y_0), the implicit function theorem says that the level curve of $F(x,y)$ for value $k = F(x_0, y_0)$ is described nearby by a function $y = f(x)$. Then Chain Rule differentiation of $k = F(x, f(x))$ gives

$$0 = \frac{\partial F}{\partial x}\frac{dx}{dx} + \frac{\partial F}{\partial y}\frac{dy}{dx}$$

so

$$\frac{dy}{dx} = -\frac{\partial F/\partial x}{\partial F/\partial y}$$

Thus the tangent line to the level curve through this point has this slope, so a tangent vector is $\left\langle -\frac{\partial F}{\partial y}, \frac{\partial F}{\partial x} \right\rangle$. This is perpendicular to the gradient vector $\nabla F(x_0, y_0)$, so the gradient at such a point on the curve is normal to the tangent line to the level curve at that point, and can be considered normal to the level curve.

Similarly, if $\partial F/\partial x \neq 0$ at any point, swapping x and y above shows that part of the described by an implicit function $x = g(y)$, and again the gradient vector is normal to the level curve. Hence

Theorem (The Gradient is Normal to Level Curves)**.** *At any point (x_0, y_0) where $F(x,y)$ is differentiable and its gradient $\nabla F(x_0, y_0)$ is non-zero, this gradient vector is normal to the level curve of F through that point, in that it is normal to the line tangent to the level curve at that point.*

Maximizing the Directional Derivative

The gradient gives the direction in which the directional derivative is greatest, and is thus the direction of most rapid increase not value of the function. One physical interpretation is that if the function value is altitude, the gradient vector indicates the direction "straight up-hill".

To see this, recall that if the angle between ∇F at (x_0, y_0) and a unit vector $\vec{u}$ is θ, $D_{\vec{u}}F(x_0, y_0) = \nabla F \cdot \vec{u} = |\nabla F||\vec{u}|\cos\theta$, and this has maximum value when $\cos\theta = 1$, which is when $\theta = 0$, so $\vec{u}$ is a positive multiple of ∇F. Thus $\vec{u} = \nabla F/|\nabla F|$ gives the direction in which $D_{\vec{u}}F$ is greatest and so F is increasing the fastest.

See Examples 6 and 7.

Functions of Three Variables

For a differentiable function f of three variables one can likewise define the directional derivative of in direction $\vec{u} = \langle u_1, u_2, u_3 \rangle$

$$D_{\vec{u}}f(x,y,z) = \lim_{h \to 0} \frac{f(x + u_1 h, y + u_2 h, z + u_3 h) - f(x,y,z)}{h}$$

and the gradient

$$\textbf{grad}\ f(x,y,z) = \nabla f(x,y,z) = \langle f_x(x,y,z), f_y(x,y,z), f_z(x,y,z) \rangle$$

Again

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

and again the directional derivative is greatest in the direction of the gradient.

See Example 5.

Tangent Planes to Level Surfaces

For a function F, consider the level surface S given by $F(x,y,z) = k$ through a point $P(x_0, y_0, z_0)$ and any curve $\vec{r}(t) = \langle x(t), y(t), z(t)\rangle$ that lies in this level surface and passes through this point when $t = t_0$. The composition of F with these three component functions give the constant function $F(x(t), y(t), z(t)) = k$, and by the Chain Rule, its zero derivative is also

$$0 = \frac{\partial F}{\partial x}\frac{dx}{dt} + \frac{\partial F}{\partial y}\frac{dy}{dt} + \frac{\partial F}{\partial z}\frac{dz}{dt} = \nabla F \cdot \vec{r}'(t)$$

In particular $\nabla F(x_0, y_0, z_0) \cdot \vec{r}'(t_0) = 0$.

The possible directions $\vec{r}'(t_0)$ for a curve passing through this point and staying in the level curve are thus all orthogonal to $\nabla F(x_0, y_0, z_0)$ and thus are directions in the plane

$$\begin{aligned}&\nabla F(x_0, y_0, z_0) \cdot \langle x - x_0, y - y_0, z - z_0\rangle \\ &\qquad = F_x(x_0, y_0, z_0)(x - x_0) + F_y(x_0, y_0, z_0)(y - y_0) + F_z(x_0, y_0, z_0)(z - z_0) = 0\end{aligned}$$

Thus it is natural to call this plane **the tangent plane to this level surface S at point P** and to say that the direction of $\nabla F(x_0, y_0, z_0)$ is normal to S at P.

Normal Lines to Level Surfaces

The line through this point normal to the surface is called the **normal line to S through P**, and has symmetric equations

$$\frac{x - x_0}{F_x(x_0, y_0, z_0)} = \frac{y - y_0}{F_y(x_0, y_0, z_0)} = \frac{z - z_0}{F_z(x_0, y_0, z_0)}.$$

See Example 8.

Tangent Planes to Graphs

An important special case is the tangent plane to the graph of a function of two variables, $z = f(x,y)$. This is given as a level surface by $F(x,y,z) = f(x,y) - z = 0$, and then $F_z(x,y,z) = -1$, leading to the equation

$$f_x(x_0, y_0)(x - x_0) + f_x(x_0, y_0)(y - y_0) - (z - z_0) = 0.$$

Using $z_0 = f(x_0, y_0)$, this is

$$z = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_x(x_0, y_0)(y - y_0),$$

the equation of the tangent plane to $z = f(x,y)$ at (x_0, y_0) seen in Section 14.4.

Homework

Exercises 7-13, 14*, 19-26, 28, 30*, 31, 32, 50-56.