Derivatives Part IV: Tangent Planes, Gradients and Directional Derivatives

Math 232 Section 2

Suppose you have a real-valued function of two variables, $f(x, y)$. If f is differentiable at (x_0, y_0), then the total derivative is the 1×2 matrix (or row vector)

$$\begin{pmatrix} \frac{\partial f}{\partial x}(x_0, y_0) & \frac{\partial f}{\partial y}(x_0, y_0) \end{pmatrix}.$$

In this special case, we call the (row) vector $\langle \partial f/\partial x, \partial f/\partial y \rangle$ the *gradient* of f. This is typically written $\mathbf{grad} f$ or ∇f. We define the gradient for a real-valued function of three variables similarly. Since $\nabla f(x_0, y_0)$ is the total derivative of f at the point (x_0, y_0), this means that the linearization of f at the point (x_0, y_0) is

$$\begin{aligned} L(x, y) &= \nabla f(x_0, y_0) \Delta \begin{pmatrix} x \\ y \end{pmatrix} + f(x_0, y_0) \\ &= \begin{pmatrix} \frac{\partial f}{\partial x}(x_0, y_0) & \frac{\partial f}{\partial y}(x_0, y_0) \end{pmatrix} \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix} + f(x_0, y_0) \\ &= f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) + f(x_0, y_0). \end{aligned}$$

The graph of the linearization here will be a plane, called the *tangent plane* to the graph of f at the point (x_0, y_0). This plane has normal vector $\langle f_x(x_0, y_0), f_y(x_0, y_0), -1 \rangle$.

Example 8: Let f be the function from example 7. Then the gradient of f is

$$\nabla f = (f_x, f_y) = \begin{pmatrix} e^y - ye^x & xe^y + e^{-x} \end{pmatrix}.$$

The gradient at the point $x = 0,\ y = 1$ is $\nabla f(0, 1) = \begin{pmatrix} e - 1 & 1 \end{pmatrix}$. An equation for the tangent plane to the graph of $f(x, y)$ at the point $(0, 1)$ is

$$z = (e - 1)(x - 0) + 1(y - 1) + f(0, 1) = (e - 1)x + y.$$

Exercises 20-24. Do exercises 2, 5, 6, 12 and 15 from section 15.4 of the textbook.

Often we need to find how fast a real-valued function of two (or three) variables changes as we move in a certain direction. Specifically, suppose we have a function $f(x, y)$ and a point (x_0, y_0) in the domain of f. What is the rate of change of f if, starting at (x_0, y_0), we move in the direction $\mathbf{u}$? (Recall that a direction is just a unit vector.) We can express this algebraically as: What is

$$\lim_{t \to 0} \frac{f\big(t\mathbf{u} + (x_0, y_0)\big) - f(x_0, y_0)}{t} \ ?$$

One way to view this is as a composition. We have our function f already. We compose this with the (linear) vector-valued function

$$\mathbf{r}(t) = \langle x_0, y_0 \rangle + t\mathbf{u} = \begin{pmatrix} x_0 + tu_1 \\ y_0 + tu_2 \end{pmatrix}.$$

The total derivative of f is the gradient ∇f, a 1×2 matrix (row vector). The total derivative of $\mathbf{r}$ is the 2×1 matrix

$$\begin{pmatrix} \frac{\partial x}{\partial t} \\ \frac{\partial y}{\partial t} \end{pmatrix} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \mathbf{u}$$

(when we view $\mathbf{u}$ as a column vector). By the chain rule, the total derivative of the composition $f \circ \mathbf{r}(t)$ is the 1×1 matrix which is simply the dot product $\nabla f \cdot \mathbf{u}$.

We call the dot product $\nabla f(x_0, y_0) \cdot \mathbf{u}$ the *directional derivative* of f at (x_0, y_0) in the direction $\mathbf{u}$. It is sometimes written $D_{\mathbf{u}} f(x_0, y_0)$. For functions $f(x, y, z)$ we define the directional derivative the same way (with the obvious changes to three-dimensional vectors).

Example 9: Let f be the same function as in example 7 again. Then $\nabla f(0, 1) = \langle e - 1, 1 \rangle$. The directional derivative of f at $(0, 1)$ in the direction $\langle \sqrt{3}/2, 1/2 \rangle$ is

$$\langle e - 1, 1 \rangle \cdot \langle \sqrt{3}/2, 1/2 \rangle = \frac{e\sqrt{3} - \sqrt{3} + 1}{2}.$$

Exercises 25-28. Do exercises 6, 9, 15 and 20 from section 15.6 of the textbook.

Since the directional derivative is a dot product of two vectors, we know that

$$D_{\mathbf{u}} f(x_0, y_0) = |\nabla f(x_0, y_0)||\mathbf{u}| \cos\theta = |\nabla f(x_0, y_0)| \cos\theta,$$

where θ is the angle between the gradient and the direction $\mathbf{u}$. In particular, notice that the directional derivative is largest, with value $|\nabla f(x_0, y_0)|$, when $\theta = 0$, i.e., when you go in the direction of the gradient. It's smallest, with exactly the opposite value, when you go in the opposite direction. It's zero when the direction is orthogonal to the gradient.

Let's expand on that last sentence. Suppose you have a real-valued function $f(x, y)$ and you're looking at a level curve: $f(x, y) = k$ for some fixed number k. Obviously, as you move along this curve the function value doesn't change (it's always just k.) If you take a point (x_0, y_0) on this level curve and move in a direction tangent to the curve, the directional derivative is zero. So the gradient $\nabla f(x_0, y_0)$ is orthogonal to the level curve at (x_0, y_0). The same reasoning works for level surfaces of a function $f(x,y,z)$: the gradient $\nabla f(x_0,y_0,z_0)$ is orthogonal to the level surface at (x_0,y_0,z_0).

We can use this information to find "formulas" for implicit differentiation. Suppose you have a function $f(x,y)$ and you're interested in a level curve $f(x,y)=k$. Take a point (x_0,y_0) on this curve. Since the gradient $\nabla f(x_0,y_0)$ is orthogonal to the curve, the equation

$$f_x(x_0,y_0)(x-x_0)+f_y(x_0,y_0)(y-y_0)=0$$

gives the tangent line to the level curve at the point (x_0,y_0). In other words, the slope of the level curve at the point (x_0,y_0) is $\frac{-f_x(x_0,y_0)}{f_y(x_0,y_0)}$ (assuming $f_y \neq 0$). Now think of the equation for the level curve $f(x,y)=k$ as defining y implicitly as a function of x (at least when we stay near the point (x_0,y_0)). Then the slope is given by the derivative $\frac{dy}{dx}$. So

$$\frac{dy}{dx}=\frac{-f_x}{f_y}.$$

In a similar vein, if (x_0,y_0,z_0) is a point on a level surface $f(x,y,z)=k$, then the gradient $\nabla f(x_0,y_0,z_0)$ is a normal vector for the tangent plane to the level surface at the point (x_0,y_0,z_0). Assuming $f_z \neq 0$,

$$\left\langle \frac{-f_x(x_0,y_0,z_0)}{f_z(x_0,y_0,z_0)}, \frac{-f_y(x_0,y_0,z_0)}{f_z(x_0,y_0,z_0)}, -1 \right\rangle$$

is also a normal vector. Now think of the equation for the level surface $f(x,y,z)=k$ as defining z implicitly as a function of x and y (at least when we stay near the point (x_0,y_0,z_0)). Then we said on page one that

$$\left\langle \frac{\partial z}{\partial x}(x_0,y_0), \frac{\partial z}{\partial y}(x_0,y_0), -1 \right\rangle$$

is a normal vector to the tangent plane. So

$$\frac{\partial z}{\partial x}=\frac{-f_x}{f_z}, \qquad \frac{\partial z}{\partial y}=\frac{-f_y}{f_z}.$$

Exercises 29-33. Do exercises 21, 28, 39 and 50 from section 15.6 and exercise 29 from section 15.5 of the textbook.