Gradients
Math 131 Multivariate Calculus
D Joyce, Spring 2014

Last time. Introduced partial derivatives like $\dfrac{\partial f}{\partial x}$ of scalar-valued functions $\mathbf{R}^n \to \mathbf{R}$, also called *scalar fields* on $\mathbf{R}^n$.

Total derivatives. We've seen what partial derivatives of scalar-valued functions $f : \mathbf{R}^n \to \mathbf{R}$ are and what they mean geometrically. But if these are only partial derivatives, then what is the 'total' derivative? The answer will be, more or less, that the partial derivatives, taken together, form the total derivative.

First, we'll develop the concept of total derivative for a scalar-valued function $f : \mathbf{R}^2 \to \mathbf{R}$, that is, a scalar field on the plane $\mathbf{R}^2$. We'll find that that total derivative is what we'll call the *gradient* of f, denoted ∇f. Next, we'll slightly generalize that to a scalar-valued function $f : \mathbf{R}^n \to \mathbf{R}$ defined on n-space $\mathbf{R}^n$. Finally we'll generalize that to a vector-valued function $\mathbf{f} : \mathbf{R}^n \to \mathbf{R}^m$.

The gradient of a function $\mathbf{R}^2 \to \mathbf{R}$. Let f be a function $\mathbf{R}^2 \to \mathbf{R}$. The graph of this function, $z = f(x, y)$, is a surface in $\mathbf{R}^3$. We would like the derivative of f to be the 'slope' of the tangent plane. But a plane doesn't have a single slope; it slopes differently in different directions. The plane tangent to this surface and passing through the point $(a, b, f(a, b))$ has the equation

$$z = f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b).$$

Here, $f_x(a, b)$ is the partial derivative of f evaluated at (a, b), and it's the slope in the x-direction. Likewise $f_y(a, b)$ is the slope in the y-direction. These two slopes determine the plane, and we'll see soon how to compute slopes in other directions from them. In that sense, the pair of these two slopes will do what we want the 'slope' of the plane to be.

Figure 1: A Lyre of Ur

We'll call the vector whose coordinates are these partial derivatives the *gradient* of f, denoted ∇f, or grad$f$.

The symbol ∇ is a nabla, and is pronounced "del" even though it's an upside down delta. The word nabla is a variant of nevel or nebel, an ancient harp or lyre. One is shown in figure 1. See the wiki article on the Lyres of Ur, the oldest surviving stringed instruments.

When $n = 2$, the gradient of the scalar field $f : \mathbf{R}^2 \to \mathbf{R}$ is

$$\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right).$$

Note that the partial derivatives $\dfrac{\partial f}{\partial x}$ and $\dfrac{\partial f}{\partial y}$ are actually functions of two variables, $\mathbf{R}^2 \to \mathbf{R}$, therefore, the gradient is also a function of two variables, but it's a vector-valued function, $\nabla f : \mathbf{R}^2 \to \mathbf{R}^2$,

$$\nabla f(x, y) = (f_x, f_y) = \left(\frac{\partial f}{\partial x}(x, y), \frac{\partial f}{\partial y}(x, y) \right).$$

Thus, whereas f was a scaler field, its gradient ∇f is a vector field.

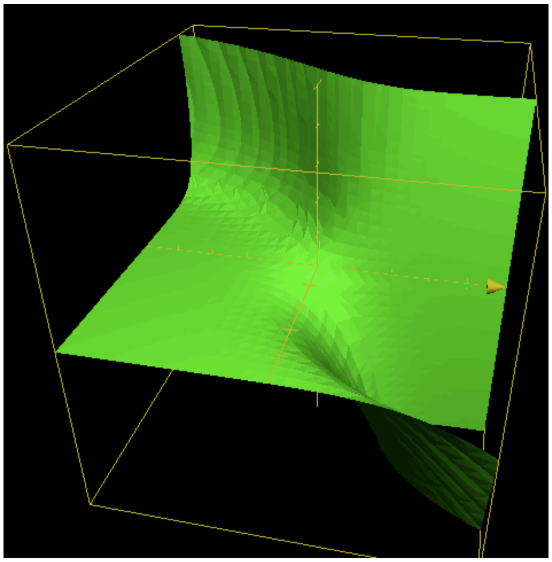

Figure 2: $z = ye^{3x+4y}$

Example. Let $f(x, y) = ye^{3x+4y}$. Its graph is shown in figure 2. Note that there's a deep valley in the 4th quadrant of the xy-plane. Also, the surface rises very quickly for positive values of y.

First, compute the partial derivatives of f. You'll find that $\frac{\partial f}{\partial x} = 3ye^{3x+4y}$, and, by the product rule, $\frac{\partial f}{\partial y} = e^{3x+4y} + 4ye^{3x+4y}$. Therefore, the gradient of f is

$$\begin{aligned} \nabla f(x, y) &= \left(\frac{\partial f}{\partial x}(x, y), \frac{\partial f}{\partial y}(x, y)\right) \\ &= (3ye^{3x+4y}, e^{3x+4y} + 4ye^{3x+4y}). \end{aligned}$$

The gradient if f is illustrated in figure 3. The vectors in this vector field point in the direction of fastest ascent. In the 4th quadrant, they point left meaning that the quickest way up out of that deep valley is to decrease x. In the 1st and 2nd quadrants, the up the surface is to increase y.

Tangent planes and differentiability. Sometimes, the partial derivatives don't tell the whole story. There may be derivatives in both the x- and y-directions, that is, there may be tangent lines in

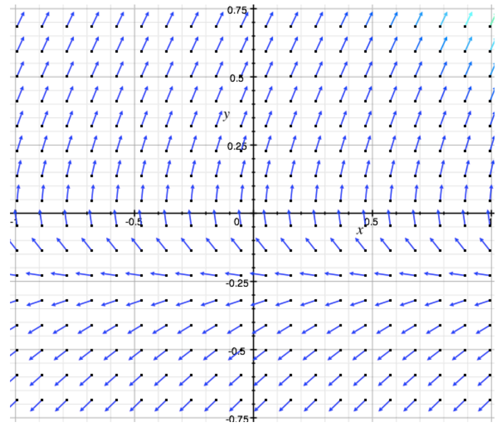

Figure 3: Vector field $\nabla f(x, y)$

both these directions, but there might not be a tangent plane. For the ordinary functions we encounter, there is always a tangent plane, but you can construct weird functions that don't have them. The question becomes: how can you tell if a function has a tangent plane?

We know, if there is a tangent plane, then its equation will be

$$z = f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b).$$

So the question becomes: how can you tell if this plane actually is a tangent plane? The answer is that the linear function

$$h(x, y) = f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b)$$

is a good approximation to the function $f(x, y)$ near (a, b). More precisely, the vertical distance between $f(x, y)$ and $h(x, y)$ approaches 0 much faster than the horizontal distance between (x, y) and (a, b). That, in turn, can be described in terms of limits of ratios as follows:

$$\lim_{(x,y)\to(a,b)} \frac{f(x, y) - h(x, y)}{\|(x, y) - (a, b)\|} = 0.$$

(Note that this limit is a vector limit.)

That leads us to our definition of differentiability for a function $f : \mathbf{R}^2 \to \mathbf{R}$. We'll say f is *differentiable* at (a, b) if, first, the partial derivatives $f_x$and f_y exist, and, second, the linear function

$$h(x, y) = f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b)$$

defined by the partial derivatives is a good approximation of f at (a,b) in the sense that

$$\lim_{(x,y)\to(a,b)} \frac{f(x,y)-h(x,y)}{\|(x,y)-(a,b)\|} = 0.$$

Then, when a function is differentiable, we'll take the gradient, ∇f, which is the vector of partial derivatives, to be the derivative.

Most common functions are differentiable. Partial derivatives are easy to compute, so it's easy to show the first condition of differentiability is satisfied just by exhibiting the derivatives. We expect the second condition to usually hold, and it does. A useful theorem (which we won't prove), says that if the partial derivatives are continuous, then the second condition holds, so the function is differentiable. That means, in practice, that the partial derivatives are enough.

Generalize to scalar-valued functions $\mathbf{R}^n \to \mathbf{R}$ with $n > 2$. Given a scalar-valued function $f : \mathbf{R}^n \to \mathbf{R}$, we'll say f is *differentiable* at $\mathbf{a}$ when it's partial derivatives $f_{x_1}(\mathbf{a}), f_{x_2}(\mathbf{a}), \ldots, f_{x_n}(\mathbf{a})$ exist and the linear function $h(\mathbf{x}) =$

$$f(\mathbf{a})+f_{x_1}(\mathbf{a})(x_1-a_1)+f_{x_2}(\mathbf{a})(x_2-a_2)+\cdots+f_{x_n}(\mathbf{a})(x_n-a_n)$$

is a good approximation of f near $\mathbf{a}$ in the sense that

$$\lim_{\mathbf{x}\to\mathbf{a}} \frac{f(\mathbf{x})-h(\mathbf{x})}{\|\mathbf{x}-\mathbf{a}\|} = 0.$$

Since $h(\mathbf{x})$ can be rewritten in terms of dot products and the gradient as

$$h(\mathbf{x}) = f(\mathbf{a}) + \nabla f(\mathbf{a}) \cdot (\mathbf{x}-\mathbf{a}),$$

that last limit can be also be written as

$$\lim_{\mathbf{x}\to\mathbf{a}} \frac{f(\mathbf{x})-[f(\mathbf{a})+\nabla f(\mathbf{a})\cdot(\mathbf{x}-\mathbf{a})]}{\|\mathbf{x}-\mathbf{a}\|} = 0.$$

This last condition is analogous to the condition of differentiability for ordinary functions $f : \mathbf{R} \to \mathbf{R}$ that looks like

$$\lim_{x\to a} \frac{f(x)-[f(a)+f'(a)(x-a)]}{(x-a)} = 0.$$

In summary, for scalar-valued functions $f : \mathbf{R}^n \to \mathbf{R}$, the gradient

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

is the derivative.

Math 131 Home Page at
`http://math.clarku.edu/~djoyce/ma131/`