Directional derivatives, steepest ascent, tangent planes
Math 131 Multivariate Calculus
D Joyce, Spring 2014

Directional derivatives. Consider a scalar field $f : \mathbf{R}^n \to \mathbf{R}$ on R^n. So far we have only considered the partial derivatives in the directions of the axes. For instance $\dfrac{\partial f}{\partial x}$ gives the rate of change along a line parallel to the x-axis. What if we want the rate of change in a direction which is not parallel to an axis?

First, we can identify directions as unit vectors, those vectors whose lengths equal 1. Let $\mathbf{u}$ be such a unit vector, $\|\mathbf{u}\| = 1$. Then we define the *directional derivative* of f in the direction $\mathbf{u}$ as being the limit

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

This is the rate of change as $\mathbf{x} \to \mathbf{a}$ in the direction $\mathbf{u}$. When $\mathbf{u}$ is the standard unit vector $\mathbf{e}_i$, then, as expected, this directional derivative is the i^{th} partial derivative, that is, $D_{\mathbf{e}_i} f(\mathbf{a}) = f_{x_i}(\mathbf{a})$.

These directional derivatives are linear combinations of the partial derivatives, at least when f is differentiable. Note that the direction $\mathbf{u} = (u_1, u_2, \ldots, u_n)$ is a linear combination of the standard unit vectors:

$$\mathbf{u} = u_1\mathbf{e}_1 + u_2\mathbf{e}_2 + \cdots + u_n\mathbf{e}_n.$$

And, when f is differentiable, it is well-approximated by the linear function g that describes the tangent plane, that is, by $g(\mathbf{x}) =$

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

Therefore,

$$\begin{aligned} & D_{\mathbf{u}} f(\mathbf{a}) \\ = & \lim_{h\to 0} \frac{f(\mathbf{a}+h\mathbf{u}) - f(\mathbf{a})}{h} \\ = & \lim_{h\to 0} \frac{g(\mathbf{a}+h\mathbf{u}) - f(\mathbf{a})}{h} \\ = & \lim_{h\to 0} \frac{f_{x_1}(\mathbf{a})hu_1 + f_{x_2}(\mathbf{a})hu_2 + \cdots + f_{x_n}(\mathbf{a})hu_n}{h} \\ = & f_{x_1}(\mathbf{a})u_1 + f_{x_2}(\mathbf{a})u_2 + \cdots + f_{x_n}(\mathbf{a})u_n \end{aligned}$$

In other notation, the directional derivative is the dot product of the gradient and the direction

$$\boxed{D_{\mathbf{u}} f(\mathbf{a}) = \nabla f(\mathbf{a}) \cdot \mathbf{u}}$$

We can interpret this as saying that the gradient, $\nabla f(\mathbf{a})$, has enough information to find the derivative in any direction.

Steepest ascent. The gradient $\nabla f(\mathbf{a})$ is a vector in a certain direction. Let $\mathbf{u}$ be any direction, that is, any unit vector, and let θ be the angle between the vectors $\nabla f(\mathbf{a})$ and $\mathbf{u}$. Now, we may conclude that the directional derivative

$$D_{\mathbf{u}} f(\mathbf{a}) = \nabla f(\mathbf{a}) \cdot \mathbf{u} = \|\nabla f(\mathbf{a})\| \cos\theta$$

since, in general, the dot product of two vectors $\mathbf{b}$ and $\mathbf{c}$ is

$$\mathbf{b} \cdot \mathbf{c} = \|\mathbf{b}\|\,\|\mathbf{c}\| \cos\theta$$

but in our case, $\mathbf{u}$ is a unit vector. But $\cos\theta$ is between -1 and 1, so the largest the directional derivative $D_{\mathbf{u}} f(\mathbf{a})$ can be is when θ is 0, that is when $\mathbf{u}$ is the direction of the gradient $\nabla f(\mathbf{a})$.

In other words, the gradient $\nabla f(\mathbf{a})$ points in the direction of the greatest increase of f, that is, the direction of steepest ascent. Of course, the opposite direction, $-\nabla f(\mathbf{a})$, is the direction of steepest descent.

Example 1. Find the curves of steepest descent for the ellipsoid

$$4x^2 + y^2 + 4z^2 = 16 \text{ for } z \geq 0.$$

If we can describe the projections of the curves in the (x, y)-plane, that's enough. This ellipsoid is the graph of a function $f : \mathbf{R}^2 \to \mathbf{R}$ given by

$$f(x, y) = \tfrac{1}{2}\sqrt{16 - 4x^2 - y^2}.$$

The gradient of this function is

$$\begin{aligned} \nabla f &= \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right) \\ &= \left(\frac{-2x}{\sqrt{16 - 4x^2 - y^2}}, \frac{-y}{2\sqrt{16 - 4x^2 - y^2}}\right) \end{aligned}$$

The curve of steepest descent will be in the opposite direction, $-\nabla f$.

So, we're looking for a path $\mathbf{x}(t) = (x(t), y(t))$ whose derivative is $-\nabla f$. In other words, we need two functions $x(t)$ and $y(t)$ such that

$$\begin{aligned} x'(t) &= \frac{2x}{\sqrt{16 - 4x^2 - y^2}}, \\ y'(t) &= \frac{y}{2\sqrt{16 - 4x^2 - y^2}}. \end{aligned}$$

Each is a differential equation with independent variable t. We can eliminate t from the discussion since

$$\frac{dy}{dx} = \frac{dy}{dt}\bigg/\frac{dx}{dt} = \frac{y}{2x}.$$

A common method to solve differential equations is separation of variables, which we can use here. From the last equation, we get

$$\frac{dy}{y} = \frac{dx}{4x}$$

and, then integrating,

$$\int \frac{dy}{y} = \int \frac{dx}{4x},$$

so

$$\ln|y| = \tfrac{1}{4}\ln|x| + C,$$

which gives us, writing A for e^C,

$$|y| = A\sqrt{|x|}.$$

That describes the curves of steepest descent as a family of curves parameterized by the real constant A (different from the last constant A)

$$x = Ay^4.$$

Tangent planes. We can, of course, use gradients to find equations for planes tangent to surfaces. A typical surface in $\mathbf{R}^3$ is given by an equation

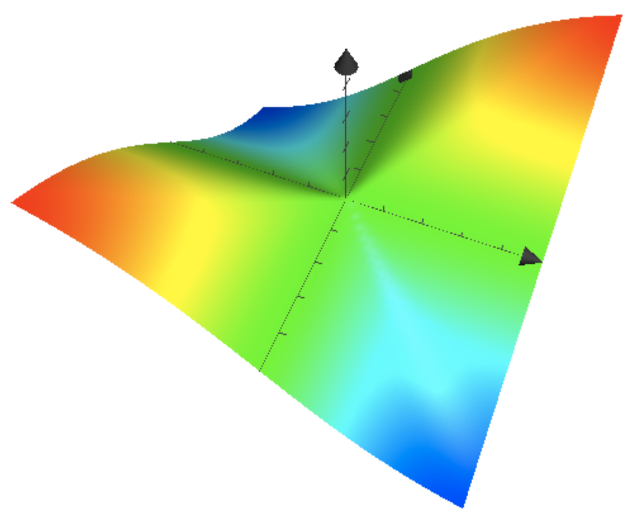
$$f(x, y, z) = c.$$

That is to say, a surface is a level set of a scalar-valued function $f : \mathbf{R}^3 \to \mathbf{R}$. More generally, a typical hypersurface in $\mathbf{R}^{n+1}$ is a level set of a function $f : \mathbf{R}^n \to \mathbf{R}$.

Now, the gradient $\nabla f(\mathbf{a})$ of f points in the direction of the greatest change of f, and vectors orthogonal to $\nabla f(\mathbf{a})$ point in directions of 0 change of f, that is to say, they lie on the tangent plane. Another way of saying that is that $\nabla f(\mathbf{a})$ is a vector normal to the surface. If $\mathbf{x}$ is any point in $\mathbf{R}^3$, then

$$\nabla f(\mathbf{a}) \cdot (\mathbf{a} - \mathbf{x}) = 0$$

says that the vector $\mathbf{a} - \mathbf{x}$ is orthogonal to $\nabla f(\mathbf{a})$, and therefore lies in the tangent plane, and so $\mathbf{x}$ is a point on that plane.



Example 2 (Continuous, nondifferentiable function)**.** You're familiar with functions of one variable that not continuous everywhere. For example, $f(x) = |x|$ is continuous, and it's differentiable everywhere except at $x = 0$. The left derivative is -1 there, but the right derivative is 1.

Things like that can happen for functions of more than one variable. Consider the function

$$f(x) = \begin{cases} 0 & \text{if } x = y = 0 \\ \dfrac{xy}{\sqrt{x^2 + y^2}} & \text{otherwise} \end{cases}$$

This function is continuous everywhere, but it's not differentiable at $(x, y) = (0, 0)$. The graph $z = f(x, y)$ has no tangent plane there. There are directional derivatives in two directions, namely, along the x-axis the function is constantly 0, so the partial derivative $\dfrac{df}{dx}$ is 0; likewise along the y-axis, and $\dfrac{df}{dy}$ is 0.

But in all other directions, the directional derivative does not exist. For instance, along the line $y = x$ the function is $f(x, x) = |x|/\sqrt{2}$, which has no derivative at $x = 0$.

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