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

Directional derivatives. Consider a scalar field $f : \mathbb{R}^n \to \mathbb{R}$ on $\mathbb{R}^n$. So far we have only considered the partial derivatives in the directions of the axes. For instance $\frac{\partial f}{\partial x}$ 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 $u$ be such a unit vector, $\|u\| = 1$. Then we define the directional derivative of $f$ in the direction $u$ as being the limit

$$D_u f(a) = \lim_{h \to 0} \frac{f(a + hu) - f(a)}{h}.$$ 

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

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

$$u = u_1 e_1 + u_2 e_2 + \cdots + u_n e_n.$$ 

And, when $f$ is differentiable, it is well-approximated by the linear function $g$ that describes the tangent plane, that is, by $g(x) = f(a) + f_{x_1}(a)(x_1 - a_1) + \cdots + f_{x_n}(a)(x_n - a_n)$.

Therefore,

$$D_u f(a) = \lim_{h \to 0} \frac{f(a + hu) - f(a)}{h}$$

$$= \lim_{h \to 0} \frac{g(a + hu) - f(a)}{h}$$

$$= \lim_{h \to 0} \frac{f_x(a)hu_1 + f_{x_2}(a)hu_2 + \cdots + f_{x_n}(a)hu_n}{h}$$

$$= f_{x_1}(a)u_1 + f_{x_2}(a)u_2 + \cdots + f_{x_n}(a)u_n$$

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

$$D_u f(a) = \nabla f(a) \cdot u$$

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

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

$$D_u f(a) = \nabla f(a) \cdot u = \|\nabla f(a)\| \cos \theta$$

since, in general, the dot product of two vectors $b$ and $c$ is

$$b \cdot c = \|b\| \|c\| \cos \theta$$

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

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

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

$$4x^2 + y^2 + z^2 = 16 \text{ for } z \geq 0.$$
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 \(f : \mathbb{R}^2 \to \mathbb{R}\) given by
\[
f(x, y) = \frac{1}{2} \sqrt{16 - 4x^2 - y^2}.
\]
The gradient of this function is
\[
\nabla f = \left( \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right) = \left( \frac{-2x}{\sqrt{16 - 4x^2 - y^2}}, \frac{-y}{2\sqrt{16 - 4x^2 - y^2}} \right).
\]
The curve of steepest descent will be in the opposite direction, \(-\nabla f\).
So, we’re looking for a path \(x(t) = (x(t), y(t))\) whose derivative is \(-\nabla f\). In other words, we need two functions \(x(t)\) and \(y(t)\) such that
\[
x'(t) = \frac{2x}{\sqrt{16 - 4x^2 - y^2}},
\]
\[
y'(t) = \frac{y}{2\sqrt{16 - 4x^2 - y^2}}.
\]
Each is a differential equation with independent variable \(t\). We can eliminate \(t\) from the discussion since
\[
\frac{dy}{dx} = \frac{dy}{dt} / \frac{dx}{dt} = \frac{y}{2x}.
\]
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
\[
\ln |y| = \frac{1}{4} \ln |x| + C,
\]
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 \(\mathbb{R}^3\) 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 \(f : \mathbb{R}^3 \to \mathbb{R}\). More generally, a typical hypersurface in \(\mathbb{R}^{n+1}\) is a level set of a function \(f : \mathbb{R}^n \to \mathbb{R}\).

Now, the gradient \(\nabla f(a)\) of \(f\) points in the direction of the greatest change of \(f\), and vectors orthogonal to \(\nabla f(a)\) 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 \(\nabla f(a)\) is a vector normal to the surface. If \(x\) is any point in \(\mathbb{R}^3\), then
\[
\nabla f(a) \cdot (a - x) = 0
\]
says that the vector \(a - x\) is orthogonal to \(\nabla f(a)\), and therefore lies in the tangent plane, and so \(x\) 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

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

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 \(\frac{df}{dx}\) is 0; likewise along the \(y\)-axis, and \(\frac{df}{dy}\) is 0.

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

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