Uses of the gradient

John P. D'Angelo
Dept. of Mathematics, Univ. of Illinois, 1409 W. Green St., Urbana IL 61801
jpda@math.uiuc.edu

definition

Let $f : \mathbf{R}^3 \to \mathbf{R}$ be a function. The partial derivatives of f are defined as usual. For example

$$\frac{\partial f}{\partial x}(x, y, z) = \lim_{h \to 0} \frac{f(x + h, y, z) - f(x, y, z)}{h}.$$

and similar equations hold for partials with respect to y and z.

We also have the notion of directional derivative:

$$\frac{\partial f}{\partial v}(p) = \lim_{t \to 0} \frac{f(p + tv) - f(p)}{t}.$$

Notice here that $p = (x, y, z)$. Also note that the partial $\frac{\partial f}{\partial x}$ is the directional derivative in the $\mathbf{i}$ direction; in other words, put $v = (1, 0, 0)$. Then $\frac{\partial f}{\partial x} = \frac{\partial f}{\partial v}$. Similar statements hold for the other coordinate directions.

The gradient ∇f is extremely important in this course and in science in general. When the partial derivatives exist, we put them in a vector:

$$\nabla f(p) = (\frac{\partial f}{\partial x}(p), \frac{\partial f}{\partial y}(p), \frac{\partial f}{\partial z}(p)). \tag{1}$$

Note that the gradient vector generally depends on the base point. Make sure you realize what that means geometrically!

We have the following uses of the gradient.

1) If $p \to \nabla f(p)$ is continuous, then f is differentiable. As a consequence we have the crucial approximate formula

$$f(p+h) \approx f(p) + \nabla f(p) \cdot h. \tag{2}$$

Recall that f is differentiable at p if

$$\lim_{h \to 0} \frac{|f(p+h) - f(p) - \nabla f(p) \cdot h|}{||h||} = 0. \tag{2.5}$$

The numerator in (2.5) is the error term. Thus (2.5) implies (2). We often use (2) to find approximate values of $f(p+h)$ when $f(p)$ is known, $\nabla f(p)$ is known, and $||h||$ is small.

2) For all directions v we have

$$\frac{\partial f}{\partial v}(p) = \nabla f(p) \cdot v. \tag{3}$$

Equation (3) is a *must-know* method for finding directional derivatives.

3) The gradient at p points in the direction in which f is increasing most rapidly. In particular $\nabla f(p)$ is orthogonal at p to the level set $\{f = f(p)\}$.

4) Tangent planes. Because of item 3), and the point-normal form of the equation of a plane, we can

use the gradient to find tangent planes to level surfaces. Obviously this technique arises throughout science. (level surfaces can be isotherms, equipotential surfaces, indifference curves, and so on.) Consider the level set $\{f = f(p)\}$. The tangent plane to this set at p is given by the equation

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

In coordinates, if $p = (x_0, y_0, z_0)$ and $\nabla f(p) = (a, b, c)$, then the equation is

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0. \qquad (3)$$

This formula must be understood.

5) Recall that the projection of v onto $\mathbf{n}$ is given by

$$\frac{v \cdot n}{||\mathbf{n}||^2}\mathbf{n},$$

and that the length of the projection is

$$\frac{|v \cdot n|}{||\mathbf{n}||}.$$

When $\mathbf{n} = \nabla f(p)$ is the gradient of f at p, then this projection becomes the projection of v onto the direction in which f is increasing most rapidly

6) We say more about items 2), 3), and 5). By formula (3), we can compute the directional derivative by taking a dot product with the gradient. By the Cauchy-Schwarz inequality we have

$$|\frac{\partial f}{\partial v}(p)| = |\nabla f(p) \cdot v| \leq ||\nabla f(p)||\;||v||. \qquad (4).$$

Furthermore, equality holds only if v and $\nabla f(p)$ point in the same direction. Therefore 3) and 5) hold. In order to see that 2) holds, we must use the definition of the directional derivative and the chain rule. The important thing to observe is that the definition of directional derivative asks us to restrict the function to a line, and take the one-dimensional derivative along the line. I will elaborate in class.

min-max problems

Remark 0.1. One often need to find the minimum and maximum values of a function. More generally, one often needs to find the minimumand maximum values of a function subject to constraints. Both these problems use gradients.

We will add information to this section as we continue.