CLASS 4, 4/4/2011, FOR MATH 13, SPRING 2011

Partial derivatives not only tell us the rate of change of a function $f(x, y)$ in either the x or y direction, they can also help us calculate objects of interest such as tangent planes and normal lines. Furthermore, they let us calculate the gradient of f, which allows us to easily compute directional derivatives, among other things.

1. Directional derivatives and the gradient

Let $f(x, y)$ be a function of two variables. Then the partial derivatives f_x, f_y can be interpreted as the rate of change of a function in either the x or y direction. However, there is nothing intrinsically special about these two directions: we may just as well ask what the rate of change of $f(x, y)$ in some other direction is.

More specifically, suppose we want to study the rate of change of $f(x, y)$ at a point (a, b). A direction can be specified by giving a unit vector; this vector points in some direction, and is uniquely determined by a direction. Let $\mathbf{u}$ be such a vector; we sometimes call unit vectors in this context a *direction vector*. Then we can ask how $f(x, y)$ changes at (a, b) as we go in the direction of $\mathbf{u}$. We define the *directional derivative* of $f(x, y)$ at $\mathbf{v} = (a, b)$ in the direction of the unit vector $\mathbf{u}$ to be the value of the limit

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

Partial derivatives are given by either letting $\mathbf{u} = \langle 1, 0\rangle$ or $\langle 0, 1\rangle$. Intuitively, if we think of $f(x, y)$ as the height of a hill, then the directional derivative at (a, b) in the direction of $\mathbf{u}$ is the rate at which the height increases or decreases if we walk in the direction of $\mathbf{u}$ at (a, b).

How can we quickly calculate directional derivatives? To answer this question, we introduce the *gradient* of a function $f(x, y)$. The gradient of $f(x, y)$, written $\nabla f(x, y)$, is defined to be the function

$$\nabla f(x, y) = \langle f_x(x, y), f_y(x, y)\rangle.$$

This is a function which has domain $\mathbb{R}^2$, and takes values in $\mathbb{R}^2$: that is, the gradient of f is a vector-valued function defined on $\mathbb{R}^2$.

It turns out that directional derivatives can easily be calculated in terms of $\nabla f(x, y)$:

$$D_{\mathbf{u}} f(a, b) = \nabla f(a, b) \cdot \mathbf{u}.$$

When using this formula to calculate partial derivatives, be absolutely sure that you are using a <u>unit</u> vector for $\mathbf{u}$.

Example. Calculate the directional derivative of $f(x, y) = x^2 + y^2$ at $(4, 7)$ in the direction $\langle 1, 2\rangle$. Remember that when calculating directional derivatives, our directions need to be specified by a unit vector. The unit vector that points in the same direction as $\langle 1, 2\rangle$ is $\langle 1/\sqrt{5}, 2/\sqrt{5}\rangle$. The gradient of $f(x, y)$ is $\nabla f = \langle 2x, 2y\rangle$. In particular, $\nabla f(4, 7) = \langle 8, 14\rangle$. Then the directional derivative in question is

$$\langle 8, 14\rangle \cdot \frac{1}{\sqrt{5}}\langle 1, 2\rangle = \frac{36}{\sqrt{5}}.$$

This formula also allows us to easily see two properties of the gradient vector. Since $|\mathbf{u}| = 1$ regardless of the choice of $\mathbf{u}$, the directional derivative $D_{\mathbf{u}} f(a,b)$ is evidently maximized when $\mathbf{u}$ points in the same direction as $\nabla f(a,b)$, because

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

which is clearly maximized when $\theta = 0$. Furthermore, the rate of change in the direction of maximum increase is given by $|\nabla f(a,b)|$. Therefore, we see that the gradient vector (1) points in the direction in which a function is increasing most rapidly, and (2) the magnitude of the gradient vector tells us the rate of this increase.

Example. Consider $z = x^2 + y^2$. At the point $(2,3)$, in what direction is z increasing most rapidly? How rapidly is z increasing in that direction? We begin by calculating $\nabla z = \langle 2x, 2y \rangle$. Therefore, $\nabla z(2,3) = \langle 4, 6 \rangle$. This is the direction in which z is increasing most rapidly. Furthermore, z is increasing at a rate of $|\nabla z(2,3)| = \sqrt{4^2+6^2} = 2\sqrt{13}$ in this direction.

2. Tangent planes and normal lines

Recall that the derivative of a single variable function $f(x)$ can be interpreted as the slope of the tangent line to the graph $y = f(x)$. In particular, at x_0, this tangent line has equation

$$y - f(x_0) = f'(x_0)(x - x_0).$$

We seek a similar formula for the tangent plane to a graph of a function of two variables. Intuitively, it is somewhat clear that a surface should usually have many lines tangent to it, and perhaps less obvious that these lines will form a plane. It turns out that if a function $f(x,y)$ is differentiable at a point (x_0, y_0), then the tangent plane is given by the equation

$$z - f(x_0, y_0) = f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0).$$

This formula is very similar to the equation for a tangent line.

Example. Find the tangent plane to $f(x,y) = xy + y^2$ at $(1,2)$. We find $f_x = y$, $f_y = x + 2y$, so $f_x(1,2) = 2$, $f_y(1,2) = 5$. The equation for the tangent plane is thus

$$z - 6 = 2(x-1) + 5(y-2).$$

There is actually a more general equation for the tangent plane to a surface given by an equation of the form $F(x,y,z) = C$, for some constant C. (In particular, the case of a graph of a function $f(x,y)$ is the special case $F(x,y,z) = z - f(x,y) = 0$.) The tangent plane to $F(x,y,z) = C$ at (x_0, y_0, z_0) is given by the formula

$$F_x(x_0,y_0,z_0)(x - x_0) + F_y(x_0,y_0,z_0)(y - y_0) + F_z(x_0,y_0,z_0)(z - z_0) = 0.$$

In particular, notice that $\nabla F(x_0, y_0, z_0)$ is a normal vector for this plane. This tells us that another interpretation of ∇F is as a vector which is orthogonal to level curves/surfaces of F.

Finally, this formula also provides us with a convenient way to calculate the *normal line* to a surface, which are the lines which are orthogonal to the tangent planes of a surface. Since $\nabla F(x_0, y_0, z_0)$ is normal to a tangent plane, a vector equation for a normal line is given by

$$\langle x_0, y_0, z_0 \rangle + t \langle F_x(x_0,y_0,z_0) + F_y(x_0,y_0,z_0) + F_z(x_0,y_0,z_0) \rangle.$$

Of course, this discussion of normal lines is valid not only for functions $F(x, y, z) = C$, but also $F(x, y) = C$.

Examples.

- Determine the equations for the normal lines to the graph of $x^2 - y^2 = 1$ for a general point on this graph. In this situation, $F(x, y) = x^2 - y^2$, so $\nabla F(x, y) = \langle 2x, -2y \rangle$. Therefore, the normal line at x_0, y_0 is given by
$$\langle x_0, y_0 \rangle + t\langle 2x_0, -2y_0 \rangle.$$
- Consider the sphere $x^2 + y^2 + z^2 = 9$. Calculate the equation for the tangent plane and normal line to the sphere at $(2, 1, 2)$.
  We begin by calculating the gradient of $f(x, y, z) = x^2 + y^2 + z^2$. We see that $\nabla f = \langle 2x, 2y, 2z \rangle$. Therefore, the gradient at $(2, 1, 2)$ is equal to $\nabla f(2, 1, 2) = \langle 4, 2, 4 \rangle$. Therefore, the tangent plane to $x^2 + y^2 + z^2 = 9$ at $(2, 1, 2)$ has normal vector $\langle 4, 2, 4 \rangle$. The equation of this plane must then be
$$4x + 2y + 4z = 18, \text{or } 2x + y + 2z = 9.$$
  The normal line has direction vector $\langle 4, 2, 4 \rangle$ and passes through $(2, 1, 2)$. Therefore, the normal line is given by parametric equations $x = 2 + 4t, y = 1 + 2t, z = 2 + 4t$. Notice that this line passes through the origin.

3. Integration

Now that the quick review of Math 8 is finished, let's start with the new material in this class. We are interested in developing a theory of integration for functions of several variables. Let us begin with the case of functions $f(x, y)$ of two real variables. We want to define an integral for this function. How should we proceed?

Notice that there is no obvious candidate for an antiderivative, so the strategy of trying to define an integral as an antiderivative is not going to work. We should therefore emulate the definition of a definite integral for functions of a single variable.

How is a definite integral
$$\int_a^b f(x)\, dx$$
defined? Recall that the definition of this integral is as the value of a limit
$$\lim_{||P|| \to 0} \sum f(x_i^*) \Delta x_i$$
where the x_i are a partition of $[a, b]$, with largest Δx_i equal to $||P||$, and x_i^* a number between x_i and x_{i+1}. Each of these sums is called a Riemann sum, and can be thought of as an approximation to the area under the graph of $y = f(x)$ on the interval $[a, b]$. This limit is just a fancy way of saying that we are interested in the limit of better and better approximations to this area, which we obtain by making the rectangles thinner and thinner. Presumably, our intuition tells us that this should make the error in approximation smaller and smaller.

What is the correct analogue of the notion of a Riemann sum for functions of two variables? Such a sum should estimate the volume under the graph of $z = f(x, y)$ on some region. The closest analogue to an interval $[a, b]$ in $\mathbb{R}$ might be a rectangle R in $\mathbb{R}^2$, perhaps of the form $[a, b] \times [c, d]$. We want a sum which estimates the volume of the region under $z = f(x, y)$ over the rectangle R.

Just like how we use rectangles to approximate the area under $y = f(x)$, we can use rectangular prisms to approximate the area under $z = f(x, y)$. In particular, over a rectangle $[x_i, x_{i+1}] \times [y_i, y_{i+1}]$, we can use a rectangular prism with that rectangle as base, and height given by $f(x_i^*, y_i^*)$ as an approximation to the volume over this smaller rectangle. This rectangular prism has volume

$$f(x_i^*, y_i^*)\Delta x_i \Delta y_i.$$

($\Delta x_i = x_{i+1} - x_i$, and similarly for Δy_i.) Therefore, the sum of these rectangular prisms, which is approximating the volume under $z = f(x, y)$ over the rectangle R, is

$$\sum f(x_i^*, y_i^*)\Delta x_i \Delta y_i$$

where the summation is over all small rectangles making up the large rectangle R. This sum is called a *Riemann sum* for $z = f(x, y)$ over the rectangle R. If we make these small rectangles smaller and smaller, then the limit of these sums (if the limit exists) will be called the *definite integral* of $f(x, y)$ over R and will be denoted

$$\iint_R f(x, y)\, dA.$$

The two integral signs remind us that we are integrating over a two-dimensional region; namely, the rectangle R which is written underneath the two integral signs. The differential dA tells us that we are integrating with respect to an *area element*. It is mostly to be thought of as notation further reminding us that we are integrating over a two-dimensional region.

Of course, in practice, we never use this definition of a definite integral to actually evaluate definite integrals

$$\int_a^b f(x)\, dx.$$

We sometimes see how to evaluate these integrals directly using the limit definition for certain special functions, like $f(x) = x^2$, but these calculations are long and tedious. Instead, we use the fundamental theorem of Calculus, which tells us that to evaluate a definite integral we need only find an antiderivative for $f(x)$, and then evaluate it at a, b.

What we seek now is a convenient way to evaluate double integrals. Even though we should not expect a 'fundamental theorem of Calculus' for double integrals yet, we can still make use of the FTC for functions of a single variable, as we shall see next class.