MATH 8 FALL 2010 CLASS 27, 11/19/2010

1. Directional derivatives

Recall that the definitions of partial derivatives of $f(x, y)$ involved limits

$$\lim_{h\to 0} \frac{f(a+h,b)-f(a,b)}{h}, \lim_{h\to 0} \frac{f(a,b+h)-f(a,b)}{h}.$$

In these two limits, we only need to know the values of $f(x, y)$ when $y = b$, $x = a$ respectively: that is, we only need to know the values of $f(x, y)$ on some line in the xy-plane which passes through (a, b). Furthermore, these lines are either horizontal or vertical. If we think of the functions $f(x, b)$ or $f(a, y)$ as functions of the single variable x, y, respectively, then the partial derivatives of $f(x, y)$ correspond to the usual derivatives of these single variable functions.

We want to now describe a generalization of this idea, called a directional derivative. Suppose that we consider $f(x, y)$ on some line passing through (a, b) which is not necessarily vertical or horizontal anymore. Then how can we calculate the rate of change of this function on this line?

Before tackling this problem, we consider the question of how we can describe the various lines which pass through $f(x, y)$. We can, for example, use parametric equations given by using a direction vector for a line and the point (a, b). However, for our purposes, we will want to parameterize these lines using the parameter t in such a way that we travel a distance of one unit on the line when t changes by 1.

Example. Consider the line given by $x = 1+3t, y = 3+4t$. What is the distance between two points $(1+3t, 3+4t)$ and $(1+3(t+1), 3+4(t+1))$? That is, when we increment t by 1, how far does the point (x, y) travel? A quick calculation shows that the answer is 5; and in general, that if a line has direction vector $\mathbf{v}$, then two points whose parameter t are separated by 1 will be distance $|\mathbf{v}|$ apart.

If we want to re-parameterize this line in such a way so that we travel unit distance when t changes by 1, we will need to multiply $\mathbf{v}$ by a scalar to make this vector unit length: that is, we need to find a unit vector which points in the same direction as $\mathbf{v}$. We know that to do this we should divide $\mathbf{v}$ by $|\mathbf{v}| = 5$; therefore, $x = 1 + 3t/5, y = 3 + 4t/5$ gives another parameterization of this line which satisfies the property that when t increases by 1, the point (x, y) moves by distance 1.

Now suppose we are given a function $f(x, y)$, a point (a, b), and a unit vector $\mathbf{u} = \langle u_1, u_2\rangle$, which we think of as describing a line passing through (a, b). This line is given by the parametric equations $x = a + u_1t, y = b + u_2t$. Then the directional derivative of $f(x, y)$ at (a, b) in the direction (or with respect to the direction) $\mathbf{u}$ is written as $D_{\mathbf{u}}f$ and defined to be the limit

$$D_{\mathbf{u}}f(a,b) = \lim_{h\to 0} \frac{f(a+u_1h, b+u_2h)-f(a,b)}{h}.$$

This can be thought of as the usual derivative of the single variable function $f(a+u_1t, b+u_2t)$ with respect to the variable t.

Remarks.

- This definition of directional derivative is a genuine generalization of partial derivatives: $f_x(a,b)$ corresponds to the directional derivative of f in the direction of the unit vector $\mathbf{i} = \langle 1, 0\rangle : D_{\mathbf{i}}f(a,b)$, and $f_y(a,b)$ corresponds to the directional derivative of f in the direction of $\mathbf{j} = \langle 0, 1\rangle$.
- When calculating directional derivatives, be absolutely sure that you are using a unit vector in the definition. For example, if you use $\langle 3, 4\rangle$ instead of $\langle 3/5, 4/5\rangle$, your answer will be off by a factor of 5.
- There are two unit vectors which can be used as direction vectors for a line, and they are negatives of each other. When calculating directional derivatives, the choice of unit vector does matter, since replacing a direction vector with its negative will flip the sign of the corresponding directional derivative. This corresponds to the intuitive fact that if you go uphill when walking in some direction, if you go in the opposite direction you will go downhill at an equal rate.
- The value of the directional derivative can be interpreted as the rate of change of the function $f(x,y)$ in the direction $\mathbf{u}$. If you imagine $f(x,y)$ as describing the height of a hill over (x,y), then the directional derivative is a measure of how steep the hill is in the direction of $\mathbf{u}$.
- Related to the above remark, if we look at the graph of the function $f(a+u_1t, b+u_2t)$, this gives a curve in three dimensions. Then the tangent line to this curve at $t = 0$, or $(a, b, f(a,b))$ can be thought of as having slope equal to the directional derivative of $f(x,y)$ at (a,b), in the direction of $\langle u_1, u_2\rangle$.

2. The gradient

We now try to find a quick way of calculating directional derivatives. The definition of a directional derivative is

$$D_{\mathbf{u}}f(a,b) = \lim_{h\to 0} \frac{f(a+u_1h, b+u_2h) - f(a,b)}{h}.$$

We can rewrite the numerator of this expression as

$$f(a+u_1h, b+u_2h) - f(a+u_1h, b) + (f(a+u_1h, b) - f(a,b)).$$

In the first term, only the y variable changes, while in the second term, only the x variable changes. Therefore, it should not be too surprising (and one can easily check once one knows to perform this trick) that the directional derivative in the direction of $\mathbf{u} = \langle u_1, u_2\rangle$ is equal to

$$u_1 f_x(a,b) + u_2 f_y(a,b),$$

at least when the function $f(x,y)$ is differentiable at (a,b). That is, computing a directional derivative boils down to calculating some expression involving partial derivatives. A useful way of remembering this formula involves the following function:

Definition. Let $f(x,y)$ be a function of two variables. Then the gradient of f is the vector-valued function written ∇f and defined by

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

∇f is only defined whenever both f_x, f_y are both defined.

We commonly call ∇ 'del' and may call the gradient of f 'del' f. Sometimes you may see grad$f$ instead of ∇f. In certain English-speaking countries (India in particular), 'del' is instead called 'nabla'.

With this definition in hand, the directional derivative of $f(x, y)$ at (a, b) in the direction of the unit vector $\mathbf{u} = \langle u_1, u_2$ is given by the dot product

$$\nabla f(a, b) \cdot \langle u_1, u_2\rangle = u_1 f_x(a, b) + u_2 f_y(a, b).$$

Examples.

- Compute the gradient of $f(x, y) = x^2y+\sin(xy)$. Calculating a gradient is equivalent to calculating partial derivatives: $f_x = 2xy + y\cos(xy)$, $f_y = x^2 + x\cos(xy)$, so the gradient of f is given by $\nabla f = \langle 2xy + y \cos(xy), x^2 + x \cos(xy)\rangle$.
- 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
$$\langle 8, 14\rangle \cdot \frac{1}{\sqrt{5}}\langle 1, 2\rangle = \frac{36}{\sqrt{5}}.$$
- Calculate the directional derivative of $f(x, y) = e^{xy}$ at $(0, 1)$ in the direction of $\langle 1, 1\rangle$. The gradient of $f(x, y)$ is $\nabla f(x, y) = \langle ye^{xy}, xe^{xy}\rangle$. At $(0, 1)$ this is equal to $\nabla f(0, 1) = \langle 1, 0\rangle$. Then the directional derivative in question is equal to
$$\langle 1, 0\rangle \cdot \frac{1}{\sqrt{2}}\langle 1, 1\rangle = \frac{1}{\sqrt{2}}.$$
- Sometimes we can specify a direction not using a vector, but instead using an angle. For example, we may ask for the directional derivative of a function $f(x, y)$ at (a, b) in the direction of an angle $\theta = \pi/3$, say. By this, we mean the unit vector which forms an angle of θ in the counterclockwise direction with $\langle 1, 0\rangle = \mathbf{i}$. To calculate a directional derivative given this description for a direction, we need to find the unit vector which forms an angle θ (in the counterclockwise direction) with $\langle 1, 0\rangle$. This is evidently the vector $\langle \cos\theta, \sin\theta\rangle$, which in the case of $\theta = \pi/3$ is $\langle 1/2, \sqrt{3}/2\rangle$. Then the directional derivative of $f(x, y)$ at (a, b) in the direction $\theta = \pi/3$ is given by the dot product
$$\nabla f(a, b) \cdot \langle \cos\pi/3, \sin\pi/3\rangle = \nabla f(a, b) \cdot \frac{1}{2}\langle 1, \sqrt{3}\rangle.$$
- You may sometimes be asked for the angle above the horizontal that the tangent line (or tangent vector) to $f(x, y)$ at (a, b) in the direction of $\mathbf{u}$ is. If you think of $z = f(x, y)$ as describing the surface of a hill, for instance, then this angle is the angle of ascent as you move in the direction of $\mathbf{u}$. For example, consider the function $f(x, y) = \ln x + \ln y$. At the point $(1, 1)$, what is the angle above the horizontal of the tangent lines to $z = f(x, y)$ in the directions of $\langle 1, 0\rangle, \langle 1/\sqrt{2}, 1/\sqrt{2}\rangle$?

  We begin by calculating $\nabla f(1, 1) = \langle 1, 1\rangle$. Then the directional derivatives in the direction of $\langle 1, 0\rangle, \langle 1/\sqrt{2}, 1/\sqrt{2}\rangle$ are $1, \sqrt{2}$, respectively. The angles of the tangent lines above the horizontals are then $\arctan 1, \arctan\sqrt{2}$. Indeed, if we travel unit distance along each of the lines given by these direction vectors, then the z-value of the corresponding tangent line increases by $1, \sqrt{2}$. These angles are then given by $\arctan 1/1, \arctan\sqrt{2}/1$.

We conclude by remarking that all of the above ideas can be generalized to functions of n variables, not just 2 variables. For example, given a function $f(x_1, \dots, x_n)$, the gradient of f is the vector-valued function

$$\nabla f = \left\langle \frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} \right\rangle.$$

Given a unit vector $\mathbf{u} = \langle u_1, \dots, u_n \rangle$, the partial derivative of f at $(a_1, \dots, a_n)$ in the direction of $\mathbf{u}$ is defined by the limit

$$D_{\mathbf{u}} f(a_1, \dots, a_n) = \lim_{h \to 0} \frac{f(a_1 + u_1 h, \dots, a_n + u_n h) - f(a_1, \dots, a_n)}{h}.$$

In practice, we calculate this directional derivative by taking the dot product

$$\nabla f(a_1, \dots, a_n) \cdot \mathbf{u}.$$

Example. Calculate the directional derivative of $f(x, y, z) = xy + y \ln z$ at $(1, 2, 1)$ in the direction of $\langle 2, 2, 1 \rangle$. We begin by calculating the gradient of $f(x, y, z)$:

$$\nabla f = \langle y, x + \ln z, y/z \rangle.$$

At the point $(1, 2, 1)$, this is equal to $\nabla f(1, 2, 1) = \langle 2, 1, 2 \rangle$. The unit vector that points in the same direction as $\langle 2, 2, 1 \rangle$ is $\langle 2/3, 2/3, 1/3 \rangle$, and so the directional derivative is

$$\langle 2, 1, 2 \rangle \cdot \frac{2}{3} \langle 2, 2, 1 \rangle = \frac{8}{3}.$$

3. The direction of maximum increase

Suppose we have a function $f(x, y)$ and a point (a, b) we are interested in. If we think about the surface $z = f(x, y)$ and the point above (a, b), then there should be a direction in which $f(x, y)$ increases most rapidly. How can we find this direction?

The directional derivative of $f(x, y)$ at (a, b) in the direction of a unit vector $\langle u_1, u_2 = \mathbf{u}$ is given by the dot product $\nabla f(a, b) \cdot \mathbf{u}$. We want to maximize this number amongst all possible unit vectors u. Recall, however, that

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

where θ is the angle between $\nabla f(a, b)$ and $\mathbf{u}$. Since the lengths of these two vectors do not depend on $\mathbf{u}$, this expression is maximal when $\theta = 1$: that is, when $\mathbf{u}, \nabla f(a, b)$ point in the same direction. Therefore, we see that the gradient vector points in the direction in which f is increasing most rapidly, and is increasing at a rate of $|\nabla f(a, b)||\mathbf{u}| = |\nabla f(a, b)|$.

Examples.

- 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.
- Suppose a hill has height given by $f(x, y) = \dfrac{100}{1 + x^2 + y}$. If we are at the point $(2, 5)$, in what direction should we go if we want to go as downhill as possible? What is the

rate of descent of the hill in that direction? Again, we start by calculating $\nabla f(x,y)$. In this case, we have

$$\nabla f(x,y) = \left\langle \frac{100(-2x)}{(1+x^2+y)^2}, \frac{-100}{(1+x^2+y)^2} \right\rangle.$$

At the point $(2,5)$, the gradient is equal to $\nabla f(2,5) = \langle -4,-1\rangle$. Therefore, the height increases fastest in the direction of $\langle -4,-1\rangle$. However, we want the direction in which the height decreases the fastest. A bit of thought suggests that this should simply be the opposite direction from $\langle -4,-1\rangle$; namely, the direction $\langle 4,1\rangle$. In this direction, the rate of descent is equal to $\sqrt{17} = |-\nabla f(2,5)|$.

4. The gradient and level curves/surfaces, tangent planes

Consider the function $f(x,y) = x^2+y^2$. Recall that the level curves of this function are (unevenly) spaced concentric circles. On the other hand, the gradient is equal to $\nabla f(x,y) = \langle 2x, 2y\rangle$. If we sketch the gradient and level curves on the same graph, we quickly see that the gradient vectors all seem to be orthogonal to the level curves of $f(x,y)$. This turns out to be true in general.

Consider a general level curve $f(x,y) = k$. Suppose we parameterize this level curve by a parameter t, so that $\langle x(t), y(t)\rangle$ describes this level curve. (It doesn't matter what the exact parameterization is.) Then we have $f(x(t),y(t)) = k$. Suppose we differentiate this equation with respect to the variable t, using the chain rule:

$$\frac{df}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} = 0.$$

Recall that $dx/dt = x'(t), dy/dt = y'(t)$ are the components to tangent vectors of the vector-valued function $\langle x(t), y(t)\rangle$; that is, $\langle x'(t), y'(t)\rangle$ is the tangent vector to $\langle x(t), y(t)\rangle$. In particular, this tangent vector is a direction vector for the tangent line to $f(x,y) = k$ at the point given by the parameter t. On the other hand, $\nabla f(x,y) = \left\langle \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right\rangle$. Therefore, the previous equation can be rewritten as

$$\frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} = \nabla f(x,y)\cdot\langle x'(t), y'(t)\rangle = 0.$$

That is, $\nabla f(x,y)$ is orthogonal to the tangent line to $f(x,y) = k$, which is equivalent to saying that $\nabla f(x,y)$ is orthogonal to the curve $f(x,y) = k$.

There was nothing special about the situation of two variables. In particular, if instead we have a function $f(x,y,z)$ of three variables, and consider the level surface $f(x,y,z) = k$, then $\nabla f(a,b,c)$, which is now a vector in $\mathbb{R}^3$, will be orthogonal to the tangent line of any curve on $f(x,y,z) = k$ passing through (a,b,c). It is not hard to show that if f is differentiable, these tangent lines actually form a plane, which is the tangent plane to $f(x,y,z) = k$ at (a,b,c). Then what we have shown is that the gradient vector $\nabla f(a,b,c)$ is a normal vector for the tangent plane to $f(x,y,z) = k$ at (a,b,c). Furthermore, if we think of the line passing through (a,b,c) with direction vector $\nabla f(a,b,c)$, then this line is normal to the tangent plane, and we sometimes call this the normal line to $f(x,y,z) = k$ at (a,b,c).

In particular, the tangent plane to $f(x,y,z) = k$ at (a,b,c) has equation

$$f_x(a,b,c)(x-a) + f_y(a,b,c)(y-b) + f_z(a,b,c)(z-c) = 0.$$

Example. 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

$$4x + 2y + 4y = 18, \text{or } 2x + y + 2y = 9.$$

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.

We see that the gradient vector provides a means of calculating not only directional derivatives, but also provides information on the direction of greatest increase or decrease, and also provides a convenient way of calculating the equation of tangent lines or tangent planes to level curves.