Directional Derivatives

Up to this point we have discussed how to take the derivative in the x-direction or in the y-direction. These were the partial derivatives with respect to x and y, respectively. But what happens if we want to take the derivative in a direction that is not parallel to the coordinate axes?

It seems reasonable to suspect that if we were to go halfway in between the x and y axes, the directional derivative would be the average of the partial derivatives with respect to x and y. But how do we combine them?

Recall that when we discussed the interpretation of the derivative, it was the change in the function if the input were to change by one unit. So, we want to define the directional derivative as the change in the function for a unit step in that particular direction.

We shall specify the direction in which we wish to take the derivative as $\mathbf{u} = u_1\mathbf{i} + u_2\mathbf{j}$. Because we want to take a unit step in this direction, we require that $\mathbf{u}$ be a unit vector. If we are given a vector that is not a unit vector, we need to convert it to a unit vector first by using the fact that $\mathbf{u} = \dfrac{\mathbf{v}}{\|\mathbf{v}\|}$. With that, we are ready to define the direction derivative.

Directional Derivative of $f(x, y)$ at (a, b) in the Direction of a Unit Vector u

If $\mathbf{u} = u_1\mathbf{i} + u_2\mathbf{j}$ is a unit vector, we define the direction derivative $f_{\mathbf{u}}$ at the point (a, b) by

$$f_{\mathbf{u}}(a,b) = \begin{matrix}\text{Rate of change of } f(x,y) \text{ in the} \\ \text{direction of } \mathbf{u} \text{ at the point } (a,b)\end{matrix} = \lim_{h\to 0}\frac{f(a+hu_1, b+hu_2) - f(a,b)}{h}$$

provided that the limit exists.

Notice that if $\mathbf{u} = \mathbf{i}$, so $u_1 = 1$ and $u_2 = 0$, then the directional derivative is $f_x(a, b)$ since $f_{\mathbf{i}}(a,b) = \lim_{h\to 0}\dfrac{f(a+h,b) - f(a,b)}{h} = f_x(a,b)$. Similarly, if $\mathbf{u} = \mathbf{j}$, then the direction derivative is equal to $f_y(a, b)$.

Example 1:

Compute the directional derivative of $f(x, y) = x + y^2$ at the point (4, 0) in the direction $\mathbf{u} = \frac{1}{2}\mathbf{i} + \frac{\sqrt{3}}{2}\mathbf{j}$.

Solution:

We notice that **u** is already a unit vector since $\|\mathbf{u}\| = \sqrt{\left(\frac{1}{2}\right)^2 + \left(\frac{\sqrt{3}}{\sqrt{2}}\right)^2} = 1$. Thus, we have that

$$f_{\mathbf{u}}(a,b) = \lim_{h\to 0} \frac{f(4 + h\frac{1}{2}, 0 + h\frac{\sqrt{3}}{2}) - f(4,0)}{h} = \lim_{h\to 0} \frac{\left(4 + h\frac{1}{2}\right) + \left(h\frac{\sqrt{3}}{2}\right)^2 - 4}{h}$$

$$= \lim_{h\to 0} \frac{4 + \frac{h}{2} + \frac{3h^2}{4} - 4}{h} = \lim_{h\to 0}\left(\frac{1}{2} + \frac{3}{4}h\right) = \frac{1}{2}.$$

We can use the partial derivatives to compute the directional derivative, as we alluded to in the opening of this section.

Recall, the directional derivative is $f_{\mathbf{u}}(a,b) = \lim_{h\to 0} \frac{f(a + hu_1, b + hu_2) - f(a,b)}{h} = \lim_{h\to 0} \frac{\Delta f}{h}$, where $\Delta f = f(a + hu_1, b + hu_2) - f(a, b)$.

Using the differential notation and the local linearity which was discussed in the previous section, we have that $\Delta f \approx f_x(a, b)\Delta x + f_y(a, b)\Delta y$, where $\Delta x = a + hu_1 - a = hu_1$ and $\Delta y = b + hu_2 - b = hu_2$.

Thus, $\frac{\Delta f}{h} \approx \frac{f_x(a,b)\Delta x + f_y(a,b)\Delta y}{h} = \frac{f_x(a,b)hu_1 + f_y(a,b)hu_2}{h} = f_x(a,b)u_1 + f_y(a,b)u_2$.

As $h \longrightarrow 0$, this approximation becomes exact. Thus, we have a formula for expressing the directional derivative in terms of a combination of the partial derivatives.

Directional Derivative of $f(x, y)$ at (a, b) in terms of Partial Derivatives

If $\mathbf{u} = u_1\mathbf{i} + u_2\mathbf{j}$ is a unit vector, we define the direction derivative $f_{\mathbf{u}}$ at the point (a, b) by $f_{\mathbf{u}}(a, b) = f_x(a, b)u_1 + f_y(a, b)u_2$.

Our expression for the directional derivative can be written as a dot product between two vectors. Namely, the vectors $f_x(a, b)\mathbf{i} + f_y(a, b)\mathbf{j}$ and $\mathbf{u} = u_1\mathbf{i} + u_2\mathbf{j}$. This is because $\left(f_x(a,b)\mathbf{i} + f_y(a,b)\mathbf{j}\right)\bullet\left(u_1\mathbf{i} + u_2\mathbf{j}\right) = f_x(a,b)u_1 + f_y(a,b)u_2$.

This may seem a bit arbitrary at first, but the reason for doing this will soon become apparent. The vector $f_x(a, b)\mathbf{i} + f_y(a, b)\mathbf{j}$ will have a nice geometric interpretation later on. Because of this, we give it a special name.

The Gradient Vector of $f(x, y)$

The **gradient vector** of a differentiable function $f(x, y)$ at the point (a, b) is

$$\nabla f(a, b) = f_x(a, b)\mathbf{i} + f_y(a, b)\mathbf{j}.$$

Using this notation, we can define the directional derivative in terms of a unit vector in the desired direction and the gradient.

Directional Derivative of $f(x, y)$ at (a, b) in terms of the Gradient Vector

If $\mathbf{u} = u_1\mathbf{i} + u_2\mathbf{j}$ is a unit vector and $f(x, y)$ is differentiable at the point (a, b), then we have that

$$f_{\mathbf{u}}(a, b) = f_x(a, b)u_1 + f_y(a, b)u_2 = \nabla f(a,b)\bullet\mathbf{u}.$$

Example 2:

Find the directional derivative of $f(x,y) = \dfrac{1}{1+x^2+y^2}$ at the point (1, 0) in the direction of the vector $\mathbf{v} = 4\mathbf{i} + 3\mathbf{j}$.

Solution:

First off, notice that the direction vector, $\mathbf{v}$, is not a unit vector. We can convert into a unit vector by using the fact that $\mathbf{u} = \dfrac{\mathbf{v}}{\|\mathbf{v}\|}$. Here, $\|\mathbf{v}\| = \sqrt{3^2 + 4^2} = 5$. Thus, we have that $\mathbf{u} = \dfrac{4}{5}\mathbf{i} + \dfrac{3}{5}\mathbf{j}$.

Notice that $f_x(x,y) = \dfrac{-2x}{(1+x^2+y^2)^2}$ and $f_y(x,y) = \dfrac{-2y}{(1+x^2+y^2)^2}$. Thus, we have that $f_x(1,0) = \dfrac{-2(1)}{(1+1^2+0^2)^2} = -\dfrac{1}{2}$ and $f_y(1,0) = \dfrac{-2(0)}{(1+1^2+0^2)^2} = 0$. Putting this all together, we have that $f_{\mathbf{u}}(1,0) = -\dfrac{1}{2}\left(\dfrac{4}{5}\right) + 0\left(\dfrac{3}{5}\right) = -\dfrac{2}{5}$.

Recall that we have a formula for expressing the dot product of two vectors in terms of their lengths and the angle in between them. In particular, we have that $\mathbf{v} \bullet \mathbf{w} = \|\mathbf{v}\|\|\mathbf{w}\|\cos\theta$.

Because we can express the directional derivative as the dot product of the gradient vector and a unit vector in the desired direction, we have another formula for the directional derivative, namely $\nabla f(a,b) \bullet \mathbf{u} = \|\nabla f(a,b)\|\|\mathbf{u}\|\cos\theta = \|\nabla f(a,b)\|\cos\theta$, where θ is the angle between the gradient and the desired direction and we used the fact that the length of a unit vector is 1, i.e. $\|\mathbf{u}\| = 1$.

This representation allows us to investigate how large the directional derivative can be and under what conditions this happens.

We see that the maximum value of $f_{\mathbf{u}}(a, b)$ occurs when $\cos\theta = 1$. That is, when $\theta = 0$ and **u** is pointing in the direction of $\nabla f(a, b)$. Similarly, we have that the minimum value of $f_{\mathbf{u}}(a, b)$ occurs when $\cos\theta = -1$. That is, when $\theta = \pi$ and **u** is pointing in the opposite direction of $\nabla f(a, b)$.

Also, when $\theta = \pi/2$ or $3\pi/2$, $\cos\theta = 0$, and so the directional derivative is zero. This corresponds to remaining on a particular contour. We can summarize these results in terms of the following.

Geometric Properties of the Gradient Vector

If $f(x, y)$ is differentiable at the point (a, b) and $\nabla f(a, b) \neq \mathbf{0}$, then

The direction of $\nabla f(a, b)$
- Is perpendicular to the contour of $f(x, y)$ at the point (a, b)
- Points in the direction of the maximum increase of $f(x, y)$

The magnitude of the gradient vector, $\|\nabla f(a, b)\|$, is
- The maximum rate of change of $f(x, y)$ at the point (a, b)
- Large when the contours are close and small when they are far apart

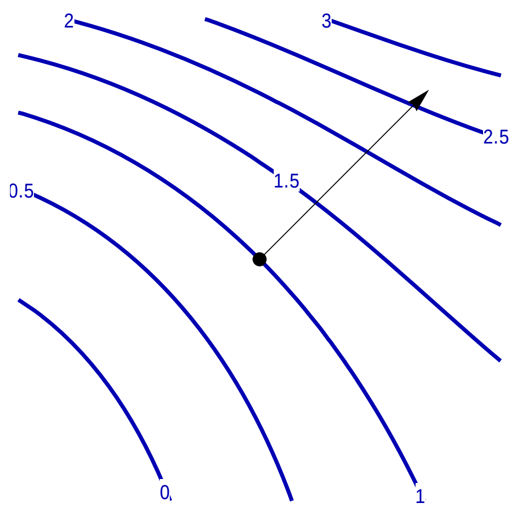
Example 3:

Find the direction in which the function $f(x, y) = \sin(x) + e^{y-1}$ has the greatest rate of change at the point (0, 1). Plot some contours and label this direction. What is the rate of change in that direction?

Solution:

The gradient points in the direction of the maximum increase. Since $f_x(x, y) = \cos(x)$ and $f_y(x, y) = e^{y-1}$, we have that $f_x(0, 1) = \cos(0) = 1$ and $f_y(0, 1) = e^{1-1} = 1$. Thus, the gradient vector at (0, 1) is equal to $\nabla f(0, 1) = \mathbf{i} + \mathbf{j}$.

A contour plot appears below. The directional derivative has been included as well.



We can determine the rate of change either by computing using the directional derivative formula or recognizing that the magnitude of gradient vector is the maximum rate of change. Using either way, we determine that the rate of change is $\sqrt{2}$.