Section 15.6
Directional Derivatives and the Gradient Vector

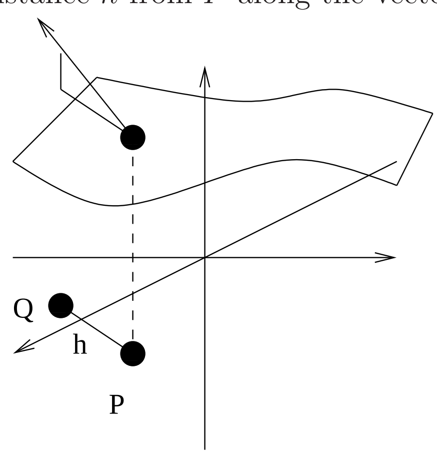
"Finding rates of change in different directions"

Recall that when we first started considering derivatives of functions of more than one variable, we had to introduce the concept of a partial derivative since there is more than one variable we can differentiate with respect to. For a function of two variables, the partial derivatives f_x and f_y had a geometric interpretation - specifically, they represented the rate of change of f in the $\vec{i}$ direction and $\vec{j}$ directions respectively. In this section we consider the more general case of determining the rate of change of a function f in the direction of any arbitrary vector $\vec{v}$ by using partial derivaitives.

1. Directional Derivatives

As with single variable calculus, by "the rate of change of f in the direction of the vector $\vec{v}$", we mean the slope of a tangent line to f at the point (a, b) in the direction of $\vec{v}$. In order to calculate this, we simply generalize the ideas from single variable calculus.

- Suppose that we want to find the rate of change of $f(x, y)$ in the direction of some unit vector $\vec{u} = u_1\vec{i} + u_2\vec{j}$ at the point $P(a, b)$ (it is important that $\vec{u}$ is a unit vector or this will not work).
- Let $Q = (a + u_1h, b + u_2h)$ (so the point in the xy-plane a distance h from P along the vector $\vec{u}$).



- The average rate of change from P to Q can be calculated using a difference quotient. Specifically, the average rate of change will equal

$$\frac{f(Q) - f(P)}{|\vec{PQ}|} = \frac{f(a + u_1h, b + u_2h) - f(a, b)}{h}$$

(observe that since $\vec{u}$ is a unit vector, the length of $\vec{PQ}$ will be h (WHY?)).

- As with single variable, we can take the limit $h \to 0$, and as h gets smaller, the difference quotient approximates the slope of f at P in the direction of $\vec{u}$ more accurately.

This motivates the following definition.

Definition 1.1. The directional derivative of $f(x, y)$ at (a, b) in the direction of a unit vector $\vec{u}$ is

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

if this limit exists.

Example 1.2. For an arbitrary function $f(x, y)$, determine the directional derivatives $D_{\vec{i}}f(x, y)$ and $D_{\vec{j}}f(x, y)$.

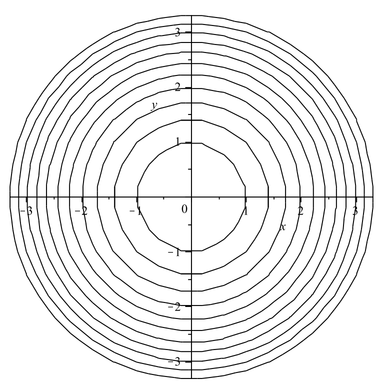
By definition,

$$D_{\vec{i}}f(x, y) = \lim_{h\to 0} \frac{f(x + h, y) - f(x, y)}{h} = f_x$$

and

$$D_{\vec{j}}f(x, y) = \lim_{h\to 0} \frac{f(x + u_1h, y) - f(x, y)}{h} = f_y.$$

Example 1.3. If the following is a contour diagram for $f(x, y)$ with the $z = 0$ contour at the origin, going up by 1 for each concentric circle, approximate the rate of change of $f(x, y)$ at $(1, 1)$ in the direction of $\vec{u} = \vec{i} + \vec{j}$.



Drawing a vector out from the point $(1, 1)$ in the direction of $\vec{u} = \vec{i}+\vec{j}$, we can use a difference quotient to approximate the rate of change. Specifically, the vector $\vec{u}$ with its tail at $(1, 1)$ has its head at $(2, 2)$ and at this point, it looks like $f(2, 2) \simeq 8$. Then we have

$$D_{\vec{u}}f(2, 2) \simeq \frac{f(2, 2) - f(1, 1)}{|\vec{u}|} = \frac{8 - 1}{\sqrt{2}} = \frac{7}{\sqrt{2}}$$

In general, we do not want to have to calculate limits when determining directional derivatives, or approximate directional derivatives using difference quotients, so we need to find a way to avoid doing this. From the previous examples, it is clear that directional derivatives are closely related to partial derivatives. In fact, we can use partial derivatives to determine directional derivatives as the following suggests.

- By definition, the directional derivative of f at (a, b) in the direction of $\vec{u}$ is
$$D_{\vec{u}}f(a,b) = \lim_{h\to 0} \frac{f(a+u_1h, b+u_2h) - f(a,b)}{h}.$$
- The numerator of this fraction measures the rate of change of f of Δf. Using linear approximations, we know $\Delta f \sim f_x(a,b)\Delta x + f_y(a,b)\Delta y$ where Δx and Δy are the change in x and y respectively.
- Observe however that we know the change in x and y. Specifically, we have $\Delta x = u_1h$ and $\Delta y = u_2h$.
- Substituting in, we have
$$\frac{f(a+u_1h, b+u_2h) - f(a,b)}{h} \sim \frac{f_x(a,b)u_1h + f_y(a,b)u_2h}{h}$$
$$= f_x(a,b)u_1 + f_y(a,b)u_2 = (f_x(a,b)\vec{i} + f_y(a,b)\vec{j}) \cdot (u_1\vec{i} + u_2\vec{j})$$

Thus we have proved the following very important result on how to calculate directional derivatives.

Result 1.4. If f is a differentiable function of x and y, then f has a directional derivative in the direction of any unit vector $\vec{u} = u_1\vec{i} + u_2\vec{j}$ at any point (a, b) which can be calculated using the formula
$$D_{\vec{u}}f(a,b) = f_x(a,b)u_1 + f_y(a,b)u_2 = (f_x(a,b)\vec{i} + f_y(a,b)\vec{j}) \cdot (u_1\vec{i} + u_2\vec{j}).$$

Note that $\vec{u}$ **must be a unit vector** to apply this formula. We illustrate with some examples.

Example 1.5. (*i*) Find the directional derivative of $f(x, y) = x^2e^y$ at $(2, 0)$ in the direction of $\vec{u} = \vec{i} + \vec{j}$.

First, we need to make $\vec{v}$ into a unit vector: $\vec{u} = \vec{i}/\sqrt{2} + \vec{j}/\sqrt{2}$. Then we calculate partial derivatives and evaluate: $f_x = 2xe^y$ so $f_x(2, 0) = 4$, and $f_y = x^2e^y$, so $f_y(a, b) = 4$. Then applying the formula, we have $D_{\vec{u}}f(2,0) = (4\vec{i} + 4\vec{j}) \cdot (\vec{i}/\sqrt{2} + \vec{j}/\sqrt{2}) = 8/\sqrt{2}$.

(*ii*) Find the direction(s) in which the directional derivative of $f(x, y) = x^2 + \sin(xy)$ at the point $(1, 0)$ has value 1.

Let $\vec{u} = u_1\vec{i} + u_2\vec{j}$ denote an arbitrary unit vector. We know $f_x = 2x + y\cos(xy)$ so $f_x(1, 0) = 2$ and $f_y = x\cos(xy)$ so

$f_y(1,0) = 1$. Thus, the directional derivative in the direction of $\vec{u}$ at $(1,0)$ will be

$$D_{\vec{u}}f(1,0) = (2\vec{i} + \vec{j}) \cdot (u_1\vec{i} + u_2\vec{j}) = 2u_1 + u_2$$

and we want this to equal 1. This means that $u_2 = 1 - 2u_1$. However, we also want $\vec{u}$ to be a unit vector, so

$$u_1^2 + u_2^2 = u_1^2 + 1 - 4u_1 + 4u_1^2 = 5u_1^2 - 4u_1 + 1 = 1.$$

This implies $u_1(5u_1 - 4) = 0$, so $u_1 = 0$ and $u_2 = 1$ or $u_1 = 4/5$ and $u_2 = -3/5$. So the directions are $\vec{j}$ and $4\vec{i}/5 - 3\vec{j}/5$.

Example 1.6. Determine a formula for the tangent line to $f(x,y) = xe^y$ at the point $(1,0)$ in the direction of $\vec{u} = 3\vec{i} - 4\vec{j}$.

In order to determine the equation for the tangent line, we need a direction and a point. We know point will be $(1,0,1)$ since $f(1,0) = 1$, so we just need the direction. However, $D_{\vec{u}}f(1,0)$ will tell use the rate of change of f in the direction of $\vec{u}$. Specifically, if we move a distance 1 in the direction of $\vec{u}$, then the change along the tangent line will be $D_{\vec{u}}f(1,0)$. We have $f_x(1,0) = 1$ and $f_y(1,0) = 1$, so

$$D_{\vec{u}}f(1,0) = (\vec{i} + \vec{j}) \cdot (\frac{3}{5}\vec{i} - \frac{4}{5}\vec{j}) = -\frac{1}{5}.$$

Therefore, the tangent line will point in the direction of $\frac{3}{5}\vec{i} - \frac{4}{5}\vec{j} - \frac{1}{5}\vec{k}$ and thus a vector equation for the tangent line will be

$$\vec{i} + \vec{k} + t\left(\frac{3}{5}\vec{i} - \frac{4}{5}\vec{j} - \frac{1}{5}\vec{k}\right).$$

Remark 1.7. Note that we can consider the directional derivative $D_{\vec{u}}f(x,y)$ as an operator on two different objects - both vectors and points. Specifically, to obtain a value, we need to input both a point and a vector i.e. a point and a direction.

2. The Gradient Vector

In the last section, the vector $f_x\vec{i} + f_y\vec{j}$ played a very important role when finding directional derivatives. For this reason, we give it a special name.

Definition 2.1. If f is a function of two variables, then the gradient vector function of f denoted by ∇f is defined by

$$\nabla f = f_x\vec{i} + f_y\vec{j}$$

This means that to find a directional derivative, we can use the following formula:

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

In addition to telling us about directional derivatives, the gradient vector also tells us certain information about the function f.

Result 2.2. If $f(x,y)$ is differentiable at (a,b) and $\nabla f(a,b) \neq \vec{0}$ then the direction of ∇f is

(i) in the direction in which f is increasing the most
(ii) perpendicular to the contours at (a,b)

and the magnitude $||\nabla f||$ is

(i) the greatest rate of change of f at (a,b)
(ii) large when contours are close together and small when contours are far apart.

Proof. Observe that for any unit vector $\vec{u}$, we have

$$D_{\vec{u}} f(a,b) = \nabla f(a,b) \cdot \vec{u} = ||\nabla f|| ||\vec{u}|| \cos(\vartheta) = ||\nabla f|| \cos(\vartheta)$$

where ϑ is the angle between them. This is largest when $\vartheta = 0$, so $\vec{u}$ is pointing in the direction of ∇f and in this case, we have $D_{\vec{u}} f(a,b) = ||\nabla f(a,b)||$.

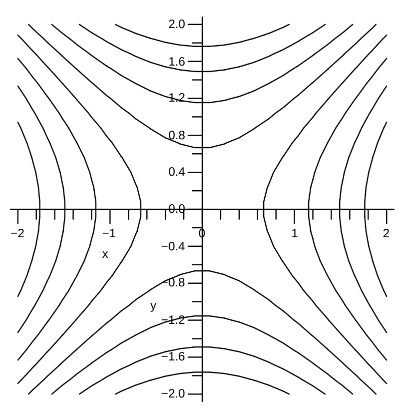
□

This means that the gradient can be used to find the greatest rate of change at a point. It can also be used to find the direction of greatest decrease $-\nabla f$ and the directions where the directional derivative is 0 (at points perpendicular to ∇f). We illustrate with some examples.

Example 2.3. (i) Find the direction of greatest increase of $f(x,y) = y^2/x$ at the point $(2,4)$ and find the value of greatest increase.

The direction of greatest increase is $\nabla f(2,4)$. Calculating, we have $f_x = -y^2/x^2$, so $f_x(2,4) = -16/4 = -4$, and $f_y = 2y/x$, so $f_y(2,4) = 8/2 = 4$. So the direction of greatest increase is $\nabla f = -4\vec{i} + 4\vec{j}$. The greatest value of increase if $||\nabla f(2,4)|| = \sqrt{16+16} = 4\sqrt{2}$.

(ii) Sketch the gradient vector on the following contour diagram at the point $(1,2)$ (contours are getting larger as we move out from the origin).



3. Functions of Three Variables

For a function of three variables, we define the directional derivative in exactly the same way as a directional derivative for a function of two variables. Moreover, all the same results hold:

(i) $\nabla f = f_x\vec{i} + f_y\vec{j} + f_z\vec{k}$
(ii) $D_{\vec{u}} = \nabla f \cdot \vec{u}$ if $\vec{u}$ is a unit vector in 3-space
(iii) The rate of change of f is most in the direction of ∇f
(iv) ∇f increases most in that direction.
(v) ∇f points in the direction perpendicular to the contours

We illustrate with an example.

Example 3.1. Find the maximum rate of increase and its direction of $f(x, y, z) = \ln(xy^2z^3)$ at the point $(1, -1, 1)$.

First we simplify: $f(x, y, z) = \ln(x) + 2\ln(y) + 3\ln(z)$, so $f_x = 1/x$, $f_y = 2/y$ and $f_z = 3/z$ giving $f_x(1, -1, 1) = 1$, $f_y(1, -1, 1) = -2$ and $f_z(1, -1, 1) = 3$. Thus we have $\nabla f = \vec{i} - 2\vec{j} + 3\vec{k}$ as the direction of greatest increase and $||\nabla f|| = \sqrt{1 + 4 + 9} = \sqrt{14}$.

Recall that a contour of a function of three variables is a surface in 3-space. This means that the gradient vector gives us a way to determine a normal vector to a surface at a point and in particular, it will allow us to determine an equation for a tangent plane at any point of a contour. Specifically, we have the following:

- Suppose $f(x, y, z)$ is a function of three variables and S is the surface which is the contour $f(x, y, z) = c$. Let (x_0, y_0, z_0) be a point on the contour (so $f(x_0, y_0, z_0) = c$).
- Then, $\nabla f(x_0, y_0, z_0)$ will be a vector pointing perpendicular to S at the point (x_0, y_0, z_0). This means we have a point and a vector, so we can form the equation of the tangent plane to S at the point (x_0, y_0, z_0).
- Specifically, $\nabla f(x_0, y_0, z_0) = f_x(x_0, y_0, z_0)\vec{i} + f_y(x_0, y_0, z_0)\vec{j} + f_z(x_0, y_0, z_0)\vec{k}$ is the vector, (x_0, y_0, z_0) is the point, so the plane will have equation

$$f_x(x_0, y_0, z_0)(x - x_0) + f_y(x_0, y_0, z_0)(y - y_0) + f_z(x_0, y_0, z_0)(z - z_0) = 0.$$

Definition 3.2. The normal line to a surface S at a point P is defined to be a line that passes through the point P perpendicular to the tangent plane at P to S.

Observe that calculation of the normal line and tangent planes is straight forward.

Example 3.3. Find the equation for the tangent plane and normal line to the sphere centered at the origin with radius 5 at the point $(4, -3, 0)$.

This surface is the contour to the function $f(x, y, z) = x^2 + y^2 + z^2$ with $f = 25$. The gradient vector will be $\nabla f = 2x\vec{i} + 2y\vec{j} + 2z\vec{k}$, so at this point on the contour will be $\nabla f(4, -3, 0) = 8\vec{i} - 6\vec{j}$. Thus the tangent plane to this contour at this point will have equation

$$8(x - 4) - 6(y + 3) = 0$$

or

$$8x - 6y = 50.$$

The parametric equations for the normal line are

$$\vec{r}(t) = (4 + 8t)\vec{i} + (-3 - 6t)\vec{j}.$$