The Calculus of Functions

of

Several Variables

Section 3.2

Directional Derivatives and the Gradient

For a function $\varphi : \mathbb{R} \to \mathbb{R}$, the derivative at a point c, that is,

$$\varphi'(c) = \lim_{h \to 0} \frac{\varphi(c+h) - \varphi(c)}{h}, \tag{3.2.1}$$

is the slope of the best affine approximation to φ at c. We may also regard it as the slope of the graph of φ at $(c, \varphi(c))$, or as the instantaneous rate of change of $\varphi(x)$ with respect to x when $x = c$. As a prelude to finding the best affine approximations for a function $f : \mathbb{R}^n \to \mathbb{R}$, we will first discuss how to generalize (3.2.1) to this setting using the ideas of slopes and rates of change for our motivation.

Directional derivatives

Example Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ defined by

$$f(x, y) = 4 - 2x^2 - y^2,$$

the graph of which is pictured in Figure 3.2.1. If we imagine a bug moving along this surface, then the slope of the path encountered by the bug will depend both on the bug's position and the direction in which it is moving. For example, if the bug is above the point $(1, 1)$ in the xy-plane, moving in the direction of the vector $\mathbf{v} = (-1, -1)$ will cause it to head directly towards the top of the graph, and thus have a steep rate of ascent, whereas moving in the direction of $-\mathbf{v} = (1, 1)$ would cause it to descend at a fast rate. These two possibilities are illustrated by the red curve on the surface in Figure 3.2.1. For another example, heading around the surface above the ellipse

$$2x^2 + y^2 = 3$$

in the xy-plane, which from $(1, 1)$ means heading initially in the direction of the vector $\mathbf{w} = (-1, 2)$, would lead the bug around the side of the hill with no change in elevation, and hence a slope of 0. This possibility is illustrated by the green curve on the surface in Figure 3.2.1. Thus in order to talk about the slope of the graph of f at a point, we must specify a direction as well. For example, suppose the bug moves in the direction of $\mathbf{v}$. If we let

$$\mathbf{u} = -\frac{1}{\sqrt{2}}(1, 1),$$

the direction of $\mathbf{v}$, then, letting $\mathbf{c} = (1, 1)$,

$$\frac{f(\mathbf{c} + h\mathbf{u}) - f(\mathbf{c})}{h}$$

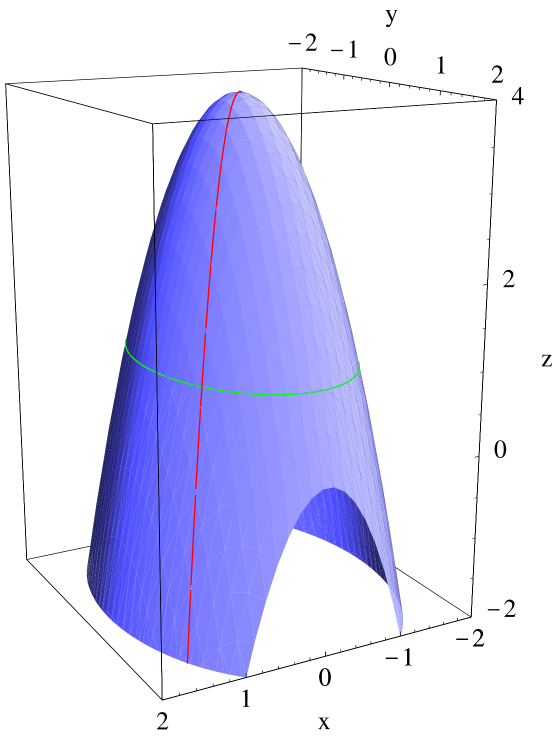

Figure 3.2.1 Graph of $f(x) = 4 - 2x^2 - y^2$

would, for any $h > 0$, represent an approximation to the slope of the graph of f at $(1,1)$ in the direction of $\mathbf{u}$. As in single-variable calculus, we should expect that taking the limit as h approaches 0 should give us the exact slope at $(1,1)$ in the direction of $\mathbf{u}$. Now

$$\begin{aligned} f(\mathbf{c}+h\mathbf{u}) - f(\mathbf{c}) &= f\left(1-\frac{h}{\sqrt{2}}, 1-\frac{h}{\sqrt{2}}\right) - f(1,1) \\ &= 4 - 2\left(1-\frac{h}{\sqrt{2}}\right)^2 - \left(1-\frac{h}{\sqrt{2}}\right)^2 - 1 \\ &= 3 - 3\left(1 - \sqrt{2}h + \frac{h^2}{2}\right) \\ &= 3\sqrt{2}h - \frac{3h^2}{2} \\ &= h\left(3\sqrt{2} - \frac{3h}{2}\right), \end{aligned}$$

so

$$\lim_{h\to 0} \frac{f(\mathbf{c}+h\mathbf{u}) - f(\mathbf{c})}{h} = \lim_{h\to 0}\left(3\sqrt{2} - \frac{3h}{2}\right) = 3\sqrt{2}.$$

Hence the graph of f has a slope of $3\sqrt{2}$ if we start above $(1,1)$ and head in the direction of $\mathbf{u}$; similar computations would show that the slope in the direction of $-\mathbf{u}$ is $-3\sqrt{2}$ and the slope in the direction of

$$\frac{\mathbf{w}}{\|\mathbf{w}\|} = \frac{1}{\sqrt{5}}\,(-1,2)$$

is 0.

Definition Suppose $f:\mathbb{R}^n \to \mathbb{R}$ is defined on an open ball about a point $\mathbf{c}$. Given a unit vector $\mathbf{u}$, we call

$$D_{\mathbf{u}}f(\mathbf{c}) = \lim_{h\to 0}\frac{f(\mathbf{c}+h\mathbf{u}) - f(\mathbf{c})}{h}, \tag{3.2.2}$$

provided the limit exists, the *directional derivative* of f in the direction of $\mathbf{u}$ at $\mathbf{c}$.

Example From our work above, if $f(x,y) = 4 - 2x^2 - y^2$ and

$$u = -\frac{1}{\sqrt{2}}(1,1),$$

then $D_{\mathbf{u}}f(1,1) = 3\sqrt{2}$.

Directional derivatives in the direction of the standard basis vectors will be of special importance.

Definition Suppose $f:\mathbb{R}^n \to \mathbb{R}$ is defined on an open ball about a point $\mathbf{c}$. If we consider f as a function of $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ and let $\mathbf{e}_k$ be the kth standard basis vector, $k = 1, 2, \ldots, n$, then we call $D_{\mathbf{e}_k}f(\mathbf{c})$, if it exists, the *partial derivative* of f with respect to x_k at $\mathbf{c}$.

Notations for the partial derivative of f with respect to x_k at an arbitrary point $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ include $D_{x_k}f(x_1, x_2, \ldots, x_n)$, $f_{x_k}(x_1, x_2, \ldots, x_n)$, and

$$\frac{\partial}{\partial x_k}f(x_1, x_2, \ldots, x_n).$$

Now suppose $f:\mathbb{R}^n \to \mathbb{R}$ and, for fixed $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, define $g:\mathbb{R}\to\mathbb{R}$ by

$$g(t) = f(t, x_2, \ldots, x_n).$$

Then

$$\begin{aligned}
f_{x_1}(x_1, x_2, \ldots, x_n) &= \lim_{h\to 0}\frac{f((x_1, x_2, \ldots, x_n) + h\mathbf{e}_1) - f(x_1, x_2, \ldots, x_n)}{h} \\
&= \lim_{h\to 0}\frac{f((x_1, x_2, \ldots, x_n) + (h, 0, \ldots, 0)) - f(x_1, x_2, \ldots, x_n)}{h} \\
&= \lim_{h\to 0}\frac{f(x_1 + h, x_2, \ldots, x_n) - f(x_1, x_2, \ldots, x_n)}{h} \\
&= \lim_{h\to 0}\frac{g(x_1 + h) - g(x_1)}{h} \\
&= g'(x_1).
\end{aligned} \tag{3.2.3}$$

In other words, we may compute the partial derivative $f_{x_1}(x_1, x_2, \ldots, x_n)$ by treating $x_2, x_3, \ldots, x_n$ as constants and differentiating with respect to x_1 as we would in single-variable calculus. The same statement holds for any coordinate: To find the partial derivative with respect to x_k, treat the other coordinates as constants and differentiate as if the function depended only on x_k.

Example If $f : \mathbb{R}^2 \to \mathbb{R}$ is defined by

$$f(x, y) = 3x^2 - 4xy^2,$$

then, treating y as a constant and differentiating with respect to x,

$$f_x(x, y) = 6x - 4y^2$$

and, treating x as a constant and differentiating with respect to y,

$$f_y(x, y) = -8xy.$$

Example If $f : \mathbb{R}^4 \to \mathbb{R}$ is defined by

$$f(w, x, y, z) = -\log(w^2 + x^2 + y^2 + z^2),$$

then

$$\frac{\partial}{\partial w} f(w, z, y, z) = -\frac{2w}{w^2 + x^2 + y^2 + z^2},$$

$$\frac{\partial}{\partial x} f(w, z, y, z) = -\frac{2x}{w^2 + x^2 + y^2 + z^2},$$

$$\frac{\partial}{\partial y} f(w, z, y, z) = -\frac{2y}{w^2 + x^2 + y^2 + z^2},$$

and

$$\frac{\partial}{\partial z} f(w, z, y, z) = -\frac{2z}{w^2 + x^2 + y^2 + z^2}.$$

Example Suppose $g : \mathbb{R}^2 \to \mathbb{R}$ is defined by

$$g(x, y) = \begin{cases} \dfrac{xy}{x^2 + y^2}, & \text{if } (x, y) \neq (0, 0), \\ 0, & \text{if } (x, y) = (0, 0). \end{cases}$$

We saw in Section 3.1 that $\lim\limits_{(x,y)\to(0,0)} g(x, y)$ does not exist; in particular, g is not continuous at $(0, 0)$. However,

$$\frac{\partial}{\partial x} g(0, 0) = \lim_{h \to 0} \frac{g((0, 0) + h(1, 0)) - g(0, 0)}{h} = \lim_{h \to 0} \frac{g(h, 0)}{h} = \lim_{h \to 0} \frac{0}{h} = 0$$

and

$$\frac{\partial}{\partial y}g(0,0) = \lim_{h\to 0}\frac{g((0,0)+h(0,1))-g(0,0)}{h} = \lim_{h\to 0}\frac{g(0,h)}{h} = \lim_{h\to 0}\frac{0}{h} = 0.$$

This shows that it is possible for a function to have partial derivatives at a point without being continuous at that point. However, we shall see in Section 3.3 that this function is not differentiable at $(0,0)$; that is, f does not have a best affine approximation at $(0,0)$.

The gradient

Definition Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is defined on an open ball containing the point $\mathbf{c}$ and $\frac{\partial}{\partial x_k}f(\mathbf{c})$ exists for $k = 1, 2, \ldots, n$. We call the vector

$$\nabla f(\mathbf{c}) = \left(\frac{\partial}{\partial x_1}f(\mathbf{c}), \frac{\partial}{\partial x_2}f(\mathbf{c}), \ldots, \frac{\partial}{\partial x_n}f(\mathbf{c})\right) \tag{3.2.4}$$

the *gradient* of f at $\mathbf{c}$.

Example If $f : \mathbb{R}^2 \to \mathbb{R}$ is defined by

$$f(x,y) = 3x^2 - 4xy^2,$$

then

$$\nabla f(x,y) = (6x - 4y^2, -8xy).$$

Thus, for example, $\nabla f(2,-1) = (8,16)$.

Example If $f : \mathbb{R}^4 \to \mathbb{R}$ is defined by

$$f(w,x,y,z) = -\log(w^2+x^2+y^2+z^2),$$

then

$$\nabla f(w,x,y,z) = -\frac{2}{w^2+x^2+y^2+z^2}(w,x,y,z).$$

Thus, for example,

$$\nabla f(1,2,2,1) = -\frac{1}{5}(1,2,2,1).$$

Notice that if $f : \mathbb{R}^n \to \mathbb{R}$, then $\nabla f : \mathbb{R}^n \to \mathbb{R}^n$; that is, we may view the gradient as a function which takes an n-dimensional vector for input and returns another n-dimensional vector. We call a function of this type a *vector field*.

Definition We say a function $f : \mathbb{R}^n \to \mathbb{R}$ is C^1 on an open set U if f is continuous on U and, for $k = 1, 2, \ldots, n$, $\frac{\partial f}{\partial x_k}$ is continuous on U.

Now suppose $f : \mathbb{R}^2 \to \mathbb{R}$ is C^1 on some open ball containing the point $\mathbf{c} = (c_1, c_2)$. Let $\mathbf{u} = (u_1, u_2)$ be a unit vector and suppose we wish to compute the directional derivative $D_{\mathbf{u}} f(\mathbf{c})$. From the definition, we have

$$\begin{aligned}
D_{\mathbf{u}} f(\mathbf{c}) &= \lim_{h \to 0} \frac{f(\mathbf{c} + h\mathbf{u}) - f(\mathbf{c})}{h} \\
&= \lim_{h \to 0} \frac{f(c_1 + hu_1, c_2 + hu_2) - f(c_1, c_2)}{h} \\
&= \lim_{h \to 0} \frac{f(c_1 + hu_1, c_2 + hu_2) - f(c_1 + hu_1, c_2) + f(c_1 + hu_1, c_2) - f(c_1, c_2)}{h} \\
&= \lim_{h \to 0} \left(\frac{f(c_1 + hu_1, c_2 + hu_2) - f(c_1 + hu_1, c_2)}{h} + \frac{f(c_1 + hu_1, c_2) - f(c_1, c_2)}{h} \right).
\end{aligned}$$

For a fixed value of $h \neq 0$, define $\varphi : \mathbb{R} \to \mathbb{R}$ by

$$\varphi(t) = f(c_1 + hu_1, c_2 + t). \tag{3.2.5}$$

Note that φ is differentiable with

$$\begin{aligned}
\varphi'(t) &= \lim_{s \to 0} \frac{\varphi(t + s) - \varphi(t)}{s} \\
&= \lim_{s \to 0} \frac{f(c_1 + hu_1, c_2 + t + s) - f(c_1 + hu_1, c_2 + t)}{s} \\
&= \frac{\partial}{\partial y} f(c_1 + hu_1, c_2 + t).
\end{aligned} \tag{3.2.6}$$

Hence if we define $\alpha : \mathbb{R} \to \mathbb{R}$ by

$$\alpha(t) = \varphi(u_2 t) = f(c_1 + hu_1, c_2 + tu_2), \tag{3.2.7}$$

then α is differentiable with

$$\alpha'(t) = u_2 \varphi'(u_2 t) = u_2 \frac{\partial}{\partial y} f(c_1 + hu_1, c_2 + tu_2). \tag{3.2.8}$$

By the Mean Value Theorem from single-variable calculus, there exists a number a between 0 and h such that

$$\frac{\alpha(h) - \alpha(0)}{h} = \alpha'(a). \tag{3.2.9}$$

Putting (3.2.7) and (3.2.8) into (3.2.9), we have

$$\frac{f(c_1 + hu_1, c_2 + hu_2) - f(c_1 + hu_2, c_2)}{h} = u_2 \frac{\partial}{\partial y} f(c_1 + hu_1, c_2 + au_2). \tag{3.2.10}$$

Similarly, if we define $\beta : \mathbb{R} \to \mathbb{R}$ by

$$\beta(t) = f(c_1 + tu_1, c_2), \tag{3.2.11}$$

then β is differentiable,

$$\beta'(t) = u_1 \frac{\partial}{\partial x} f(c_1 + tu_1, c_2), \tag{3.2.12}$$

and, using the Mean Value Theorem again, there exists a number b between 0 and h such that

$$\frac{f(c_1 + hu_1, c_2) - f(c_1, c_2)}{h} = \frac{\beta(h) - \beta(0)}{h} = \beta'(b) = u_1 \frac{\partial}{\partial x} f(c_1 + bu_1, c_2). \tag{3.2.13}$$

Putting (3.2.10) and (3.2.13) into our expression for $D_{\mathbf{u}} f(\mathbf{c})$ above, we have

$$D_{\mathbf{u}} f(\mathbf{c}) = \lim_{h \to 0} \left(u_2 \frac{\partial}{\partial y} f(c_1 + hu_1, c_2 + au_2) + u_1 \frac{\partial}{\partial x} f(c_1 + bu_1, c_2) \right). \tag{3.2.14}$$

Now both a and b approach 0 as h approaches 0 and both $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ are assumed to be continuous, so evaluating the limit in (3.2.14) gives us

$$D_{\mathbf{u}} f(\mathbf{c}) = u_2 \frac{\partial}{\partial y} f(c_1, c_2) + u_1 \frac{\partial}{\partial x} f(c_1, c_2) = \nabla f(\mathbf{c}) \cdot \mathbf{u}. \tag{3.2.15}$$

A straightforward generalization of (3.2.15) to the case of a function $f : \mathbb{R}^n \to \mathbb{R}$ gives us the following theorem.

Theorem Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is C^1 on an open ball containing the point $\mathbf{c}$. Then for any unit vector $\mathbf{u}$, $D_{\mathbf{u}} f(\mathbf{c})$ exists and

$$D_{\mathbf{u}} f(\mathbf{c}) = \nabla f(\mathbf{c}) \cdot \mathbf{u}. \tag{3.2.16}$$

Example If $f : \mathbb{R}^2 \to \mathbb{R}$ is defined by

$$f(x, y) = 4 - 2x^2 - y^2,$$

then

$$\nabla f(x, y) = (-4x, -2y).$$

If

$$\mathbf{u} = -\frac{1}{\sqrt{2}}(1, 1),$$

then

$$D_{\mathbf{u}} f(1, 1) = \nabla f(1, 1) \cdot \mathbf{u} = (-4, -2) \cdot \left(-\frac{1}{\sqrt{2}}(1, 1) \right) = \frac{6}{\sqrt{2}} = 3\sqrt{2},$$

as we saw in this first example of this section. Note also that

$$D_{-\mathbf{u}} f(1, 1) = \nabla f(1, 1) \cdot (-\mathbf{u}) = (-4, -2) \cdot \left(\frac{1}{\sqrt{2}}(1, 1) \right) = -\frac{6}{\sqrt{2}} = -3\sqrt{2}$$

and, if

$$\mathbf{w} = \frac{1}{\sqrt{5}}(-1, 2),$$

$$D_{\mathbf{w}}f(1,1) = \nabla f(1,1) \cdot (\mathbf{w}) = (-4, -2) \cdot \left(\frac{1}{\sqrt{5}}(-1, 2)\right) = 0,$$

as claimed earlier.

Example Suppose the temperature at a point in a metal cube is given by

$$T(x, y, z) = 80 - 20xe^{-\frac{1}{20}(x^2+y^2+z^2)},$$

where the center of the cube is taken to be at $(0, 0, 0)$. Then we have

$$\frac{\partial}{\partial x}T(x, y, z) = 2x^2e^{-\frac{1}{20}(x^2+y^2+z^2)} - 20e^{-\frac{1}{20}(x^2+y^2+z^2)},$$

$$\frac{\partial}{\partial y}T(x, y, z) = 2xye^{-\frac{1}{20}(x^2+y^2+z^2)},$$

and

$$\frac{\partial}{\partial z}T(x, y, z) = 2xze^{-\frac{1}{20}(x^2+y^2+z^2)},$$

so

$$\nabla T(x, y, z) = e^{-\frac{1}{20}(x^2+y^2+z^2)}(2x^2 - 20, 2xy, 2xz).$$

Hence, for example, the rate of change of temperature at the origin in the direction of the unit vector

$$\mathbf{u} = \frac{1}{\sqrt{3}}(1, -1, 1)$$

is

$$D_{\mathbf{u}}T(0, 0, 0) = \nabla T(0, 0, 0) \cdot \mathbf{u} = (-20, 0, 0) \cdot \left(\frac{1}{\sqrt{3}}(1, -1, 1)\right) = -\frac{20}{\sqrt{3}}.$$

An application of the Cauchy-Schwarz inequality to (3.2.16) shows us that

$$|D_{\mathbf{u}}f(\mathbf{c})| = |\nabla f(\mathbf{c}) \cdot \mathbf{u}| \leq \|\nabla f(\mathbf{c})\|\|\mathbf{u}\| = \|\nabla f(\mathbf{c})\|. \tag{3.2.17}$$

Thus the magnitude of the rate of change of f in any direction at a given point never exceeds the length of the gradient vector at that point. Moreover, in our discussion of the Cauchy-Schwarz inequality we saw that we have equality in (3.2.17) if and only if $\mathbf{u}$ is parallel to $\nabla f(\mathbf{c})$. Indeed, supposing $\nabla f(\mathbf{c}) \neq \mathbf{0}$, when

$$\mathbf{u} = \frac{\nabla f(\mathbf{c})}{\|\nabla f(\mathbf{c})\|},$$

we have

$$D_{\mathbf{u}}f(\mathbf{c}) = \nabla f(\mathbf{c}) \cdot \mathbf{u} = \frac{\nabla f(\mathbf{c}) \cdot \nabla f(\mathbf{c})}{\|\nabla f(\mathbf{c})\|} = \frac{\|\nabla f(\mathbf{c})\|^2}{\|\nabla f(\mathbf{c})\|} = \|\nabla f(\mathbf{c})\| \tag{3.2.18}$$

and

$$D_{-\mathbf{u}}f(\mathbf{c}) = -\|\nabla f(\mathbf{c})\|. \tag{3.2.19}$$

Hence we have the following result.

Proposition Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is C^1 on an open ball containing the point $\mathbf{c}$. Then $D_{\mathbf{u}}f(\mathbf{c})$ has a maximum value of $\|\nabla f(\mathbf{c})\|$ when $\mathbf{u}$ is the direction of $\nabla f(\mathbf{c})$ and a minimum value of $-\|\nabla f(\mathbf{c})\|$ when $\mathbf{u}$ is the direction of $-\nabla f(\mathbf{c})$.

In other words, the gradient vector points in the direction of the maximum rate of increase of the function and the negative of the gradient vector points in the direction of the maximum rate of decrease of the function. Moreover, the length of the gradient vector tells us the rate of increase in the direction of maximum increase and its negative tells us the rate of decrease in the direction of maximum decrease.

Example As we saw above, if $f : \mathbb{R}^2 \to \mathbb{R}$ is defined by

$$f(x, y) = 4 - 2x^2 - y^2,$$

then

$$\nabla f(x, y,) = (-4x, -2y).$$

Thus $\nabla f(1, 1) = (-4, -2)$. Hence if a bug standing above $(1, 1)$ on the graph of f wants to head in the direction of most rapid ascent, it should move in the direction of the unit vector

$$\mathbf{u} = \frac{\nabla f(1,1)}{\|\nabla f(1,1)\|} = -\frac{1}{\sqrt{5}}(2, 1).$$

If the bug wants to head in the direction of most rapid descent, it should move in the direction of the unit vector

$$-\mathbf{u} = \frac{1}{\sqrt{5}}(2, 1).$$

Moreover,

$$D_{\mathbf{u}}f(1, 1) = \|\nabla f(1, 1)\| = \sqrt{20}$$

and

$$D_{-\mathbf{u}}f(1, 1) = -\|\nabla f(1, 1)\| = -\sqrt{20}.$$

Figure 3.2.2 shows scaled values of $\nabla f(x, y)$ plotted for a grid of points (x, y). The vectors are scaled so that they fit in the plot, without overlap, yet still show their relative magnitudes. This is another good geometric way to view the behavior of the function. Supposing our bug were placed on the side of the graph above $(1, 1)$ and that it headed up the hill in such a manner that it always chose the direction of steepest ascent, we can see that it would head more quickly toward the y-axis than toward the x-axis. More explicitly,

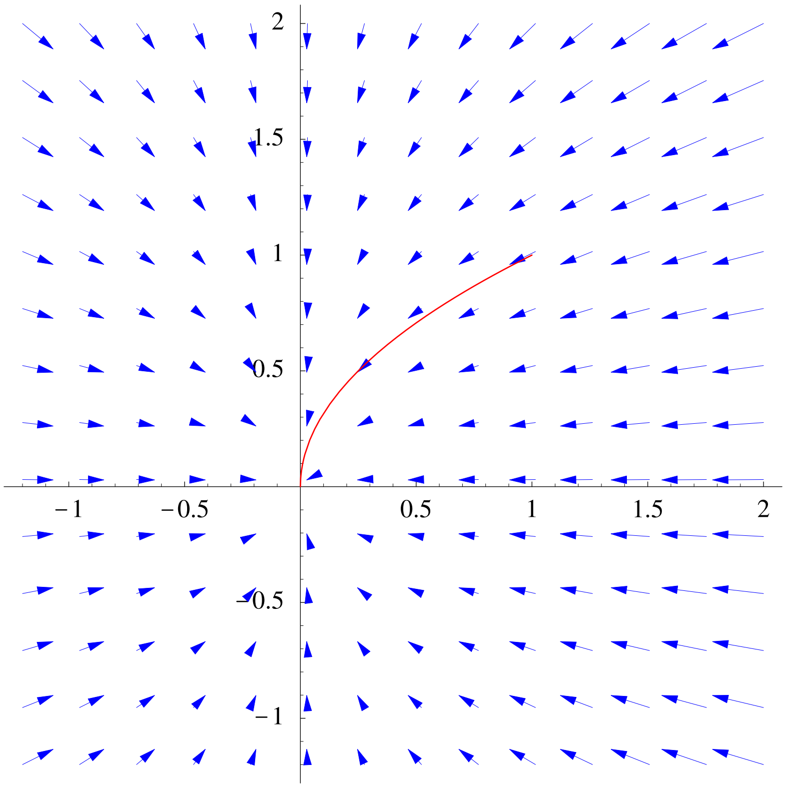

Figure 3.2.2 Scaled gradient vectors for $f(x,y) = 4 - 2x^2 - y^2$

if C is the shadow of the path of the bug in the xy-plane, then the slope of C at any point (x,y) would be

$$\frac{dy}{dx} = \frac{-2y}{-4x} = \frac{y}{2x}.$$

Hence

$$\frac{1}{y}\frac{dy}{dx} = \frac{1}{2x}.$$

If we integrate both sides of this equality, we have

$$\int \frac{1}{y}\frac{dy}{dx}\,dx = \int \frac{1}{2x}\,dx.$$

Thus

$$\log|y| = \frac{1}{2}\log|x| + c$$

for some constant c, from which we have

$$e^{\log|y|} = e^{\frac{1}{2}\log|x|+c}.$$

It follows that

$$y = k\sqrt{|x|},$$

where $k = \pm e^c$. Since $y = 1$ when $x = 1$, $k = 1$ and we see that C is the graph of $y = \sqrt{x}$. Figure 3.2.2 shows C along with the plot of the gradient vectors of f, while Figure 3.2.3 shows the actual path of the bug on the graph of f.

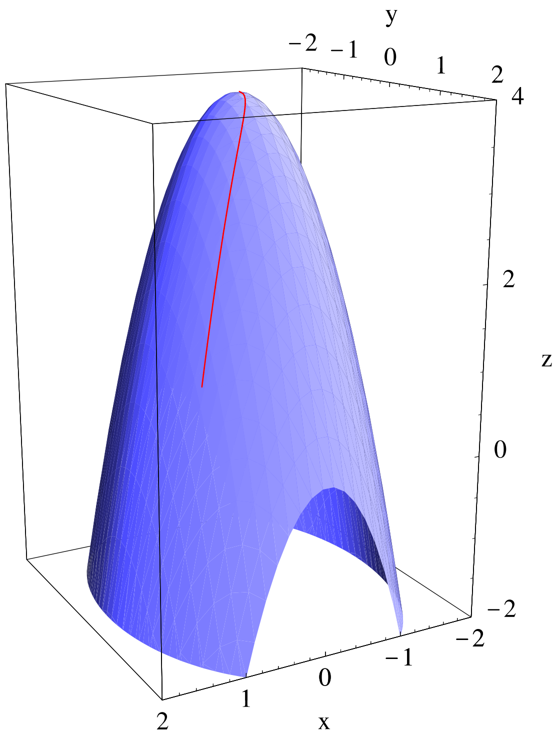

Figure 3.2.3 Graph $f(x,y) = 4 - 2x^2 - y^2$ with path of most rapid ascent from $(1,1,1)$

Example For a two-dimensional version of the temperature example discussed above, consider a metal plate heated so that its temperature at (x,y) is given by

$$T(x,y) = 80 - 20xe^{-\frac{1}{20}(x^2+y^2)}.$$

Then

$$\nabla T(x,y) = e^{-\frac{1}{20}(x^2+y^2)}(2x^2 - 20, 2xy),$$

so, for example,

$$\nabla T(0,0) = (-20, 0).$$

Thus at the origin the temperature is increasing most rapidly in the direction of $u = (-1,0)$ and decreasing most rapidly in the direction of $(1,0)$. Moreover,

$$D_{\mathbf{u}}T(0,0) = \|\nabla f(0,0)\| = 20$$

and

$$D_{-\mathbf{u}}T(0,0) = -\|\nabla f(0,0)\| = 20.$$

Note that

$$D_{-\mathbf{u}}T(0,0) = \frac{\partial}{\partial x}T(0,0)$$

and

$$D_{\mathbf{u}}T(0,0) = -\frac{\partial}{\partial x}T(0,0).$$

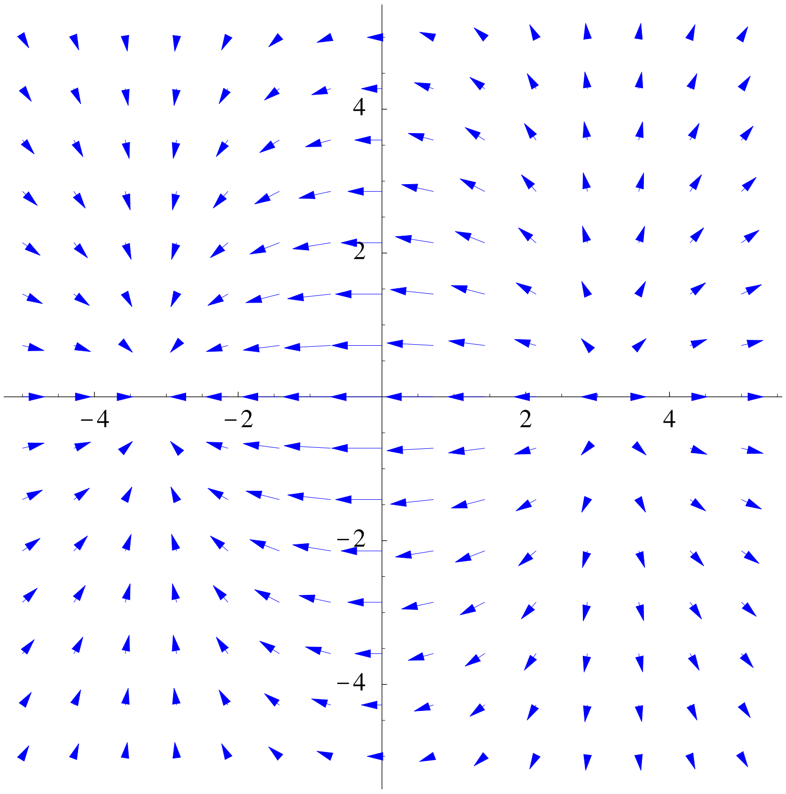

Figure 3.2.4 Scaled gradient vectors for $T(x,y) = 80 - 20xe^{-\frac{1}{20}(x^2+y^2)}$

Figure 3.2.4 is a plot of scaled gradient vectors for this temperature function. From the plot it is easy to see which direction a bug placed on this metal plate would have to choose in order to warm up as rapidly as possible. It should also be clear that the temperature has a relative maximum around $(-3, 0)$ and a relative minimum around $(3, 0)$; these points are, in fact, exactly $(-\sqrt{10}, 0)$ and $(\sqrt{10}, 0)$, the points where $\nabla T(x, y) = (0, 0)$. We will consider the problem of finding maximum and minimum values of functions of more than one variable in Section 3.5.

Problems

1. Suppose $f : \mathbb{R}^2 \to \mathbb{R}$ is defined by
$$f(x, y) = 3x^2 + 2y^2.$$
Let
$$\mathbf{u} = \frac{1}{\sqrt{5}}(1, 2).$$
Find $D_{\mathbf{u}}f(3, 1)$ directly from the definition (3.2.2).

2. For each of the following functions, find the partial derivatives with respect to each variable.

(a) $f(x,y) = \dfrac{4x}{x^2+y^2}$ (b) $g(x,y) = 4xy^2e^{-y^2}$

(c) $f(x,y,z) = 3x^2y^3z^4 - 13x^2y$ (d) $h(x,y,z) = 4xze^{-\frac{1}{x^2+y^2+z^2}}$

(e) $g(w,x,y,z) = \sin(\sqrt{w^2+x^2+2y^2+3z^2})$

3. Find the gradient of each of the following functions.

(a) $f(x,y,z) = \sqrt{x^2+y^2+z^2}$ (b) $g(x,y,z) = \dfrac{1}{\sqrt{x^2+y^2+z^2}}$

(c) $f(w,x,y,z) = \tan^{-1}(4w+3x+5y+z)$

4. Find $D_{\mathbf{u}}f(\mathbf{c})$ for each of the following.

(a) $f(x,y) = 3x^2+5y^2$, $\mathbf{u} = \dfrac{1}{\sqrt{13}}(3,-2)$, $\mathbf{c} = (-2,1)$

(b) $f(x,y) = x^2-2y^2$, $\mathbf{u} = \dfrac{1}{\sqrt{5}}(-1,2)$, $\mathbf{c} = (-2,3)$

(c) $f(x,y,z) = \dfrac{1}{\sqrt{x^2+y^2+z^2}}$, $\mathbf{u} = \dfrac{1}{\sqrt{6}}(1,2,1)$, $\mathbf{c} = (-2,2,1)$

5. For each of the following, find the directional derivative of f at the point $\mathbf{c}$ in the direction of the specified vector $\mathbf{w}$.

(a) $f(x,y) = 3x^2y$, $\mathbf{w} = (2,3)$, $\mathbf{c} = (-2,1)$

(b) $f(x,y,z) = \log(x^2+2y^2+z^2)$, $\mathbf{w} = (-1,2,3)$, $\mathbf{c} = (2,1,1)$

(c) $f(t,x,y,z) = tx^2yz^2$, $\mathbf{w} = (1,-1,2,3)$, $\mathbf{c} = (2,1,-1,2)$

6. A metal plate is heated so that its temperature at a point (x,y) is

$$T(x,y) = 50y^2e^{-\frac{1}{5}(x^2+y^2)}.$$

A bug is placed at the point $(2,1)$.

(a) The bug heads toward the point $(1,-2)$. What is the rate of change of temperature in this direction?

(b) In what direction should the bug head in order to warm up at the fastest rate? What is the rate of change of temperature in this direction?

(c) In what direction should the bug head in order to cool off at the fastest rate? What is the rate of change of temperature in this direction?

(d) Make a plot of the gradient vectors and discuss what it tells you about the temperatures on the plate.

7. A heat-seeking bug is a bug that always moves in the direction of the greatest increase in heat. Discuss the behavior of a heat seeking bug placed on a metal plate heated so that the temperature at (x,y) is given by

$$T(x,y) = 100 - 40xye^{-\frac{1}{10}(x^2+y^2)}.$$

8. Suppose $g : \mathbb{R}^2 \to \mathbb{R}$ is defined by

$$g(x,y) = \begin{cases} \dfrac{xy}{x^2+y^2}, & \text{if } (x,y) \neq (0,0), \\ 0, & \text{if } (x,y) = (0,0). \end{cases}$$

We saw above that both partial derivatives of g exist at $(0,0)$, although g is not continuous at $(0,0)$.

(a) Show that neither $\frac{\partial g}{\partial x}$ nor $\frac{\partial g}{\partial y}$ is continuous at $(0,0)$.

(b) Let

$$\mathbf{u} = \frac{1}{\sqrt{2}}(1,1).$$

Show that $D_{\mathbf{u}}g(0,0)$ does not exist. In particular, $D_{\mathbf{u}}g(0,0) \neq \nabla g(0,0) \cdot \mathbf{u}$.

9. Suppose the price of a certain commodity, call it commodity A, is x dollars per unit and the price of another commodity, B, is y dollars per unit. Moreover, suppose that $d_A(x,y)$ represents the number of units of A that will be sold at these prices and $d_B(x,y)$ represents the number of units of B that will be sold at these prices. These functions are known as the *demand functions* for A and B.

(a) Explain why it is reasonable to assume that

$$\frac{\partial}{\partial x} d_A(x,y) < 0$$

and

$$\frac{\partial}{\partial y} d_B(x,y) < 0$$

for all (x,y).

(b) Suppose the two commodities are competitive. For example, they might be two different brands of the same product. In this case, what would be reasonable assumptions for the signs of

$$\frac{\partial}{\partial y} d_A(x,y)$$

and

$$\frac{\partial}{\partial x} d_B(x,y)?$$

(c) Suppose the two commodities complement each other. For example, commodity A might be a computer and commodity B a type of software. In this case, what would be reasonable assumptions for the signs of

$$\frac{\partial}{\partial y} d_A(x,y)$$

and

$$\frac{\partial}{\partial x} d_B(x,y)?$$

10. Suppose $P(x_1, x_2, \ldots, x_n)$ represents the total production per week of a certain factory as a function of x_1, the number of workers, and other variables, such as the size of the supply inventory, the number of hours the assembly lines run per week, and so on. Show that average productivity

$$\frac{P(x_1, x_2, \ldots, x_n)}{x_1}$$

increases as x_1 increases if and only if

$$\frac{\partial}{\partial x_1} P(x_1, x_2, \ldots, x_n) > \frac{P(x_1, x_2, \ldots, x_n)}{x_1}.$$

11. Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is C^1 on an open ball about the point $\mathbf{c}$.
    (a) Given a unit vector $\mathbf{u}$, what is the relationship between $D_{\mathbf{u}} f(\mathbf{c})$ and $D_{-\mathbf{u}} f(\mathbf{c})$?
    (b) Is it possible that $D_{\mathbf{u}} f(\mathbf{c}) > 0$ for every unit vector $\mathbf{u}$?