This week we’ll explore some of the ideas from chapter 15, focusing mostly on the gradient. We’ll motivate this exploration with an example that we carry through the discussion, but we’ll see notably fewer worked examples this week than usual. The examples should become substantially easier once we have a mastery of the ideas involved.

Throughout this discussion, let’s suppose we have a function \( f(x, y) \) which gives the elevation (in whatever units you like) at the point \((x, y)\). For instance, Figure 1 shows a possible graph of an elevation function, along with some level curves of that function. As you’ve learned, the partial derivatives \( f_x \) and \( f_y \) are meant to give the slopes of \( z = f(x, y) \) in the \( x \)- and \( y \)-directions, respectively. That is, if we’re standing at the point \((a, b)\), our elevation is \( f(a, b) \) and we expect this elevation to increase by approximately \( f_x(a, b) \) units if we move to \((a + 1, b)\). But what if we move in some direction other than the \( x \)-direction or \( y \)-direction? You’ve investigated this question in the context of local linear approximations; today we want to think about it in a slightly different way.

(a) The graph \( z = f(x, y) \).
(b) Some level curves.

Figure 1: An elevation function.

Suppose we drive around the region for which \( f(x, y) \) provides the elevation function and record our elevation as we drive. For instance, Figure 2 shows a path our car might follow, as well as the elevation we would experience along this path. Notice that we begin with a relatively high elevation, descend to a saddle point, continue descending into the valley surrounding the point \((1,1)\), and then ascend a bit; all of this elevation change is reflected in the elevation profile on the right. If we write our path as \( \mathbf{r}(t) = (x(t), y(t)) \), then the elevation we experience at time \( t \) should be given by

\[
 f(\mathbf{r}(t)) = f(x(t), y(t)).
\]

How quickly does this elevation change with respect to time? This should be given by the derivative:

\[
 \frac{d}{dt} f(\mathbf{r}(t)) = \frac{d}{dt} f(x(t), y(t)).
\]
To evaluate this derivative it will help to remember the chain rule from single-variable calculus. Remember that if \( g \) and \( h \) are single-variable (differentiable) functions, then

\[
\frac{d}{dx} (g(h(x))) = g'(h(x))h'(x).
\]

We can visualize the chain rule as follows:

\[
\begin{array}{c}
 x \\
 \xrightarrow{h'} \hspace{1cm} \xrightarrow{h} \hspace{1cm} g' \hspace{1cm} g \\
 \end{array}
\]

We want to know how a small change in the value of \( x \) affects the value of \( g(h(x)) \). When we change \( x \) we change \( h(x) \), and this change is measured by \( h'(x) \). Changing \( h(x) \) then affects the value of \( g(h(x)) \), and this change is measured by \( g'(h(x)) \). The total change is then the product \( g'(h(x))h'(x) \).

Differentiating our elevation function will follow a similar pattern. We can visualize the function \( f(x(t), y(t)) \) as

\[
\begin{array}{c}
 x \\
 \xrightarrow{x'} \hspace{1cm} f_x \\
 \xrightarrow{t} \hspace{1cm} f \\
 \xrightarrow{y'} \hspace{1cm} f_y \\
 y \\
 \end{array}
\]

The elevation \( f \) depends on the coordinates \( x \) and \( y \), which in turn both depend on \( t \). Changing the value of \( t \) will change the value of \( x(t) \), which will then affect the value of \( f(x(t), y(t)) \). But changing \( t \) will also change \( y(t) \) and this change will also affect the value of \( f(x(t), y(t)) \). Summing
these effects tells us that
\[
\frac{d}{dt}(f(x(t), y(t))) = f_x(x(t), y(t))x'(t) + f_y(x(t), y(t))y'(t).
\]
Notice that the terms in this sum come in two types: there are derivatives of the elevation function \(f\) and derivatives of our path \(r(t)\). Moreover, we can separate these terms by writing the expression as a dot product:
\[
\frac{d}{dt}(f(x(t), y(t))) = \langle f_x(x(t), y(t)), f_y(x(t), y(t)) \rangle \cdot \langle x'(t), y'(t) \rangle.
\]
The first vector captures everything we need to know (for this particular application) about derivatives of \(f\), so we give it a name: the gradient of \(f\) is a vector-valued function of two variables defined by
\[
\nabla f(x, y) = \langle f_x(x, y), f_y(x, y) \rangle.
\]
Finally we have a very succinct way of expressing the time derivative of our elevation function:
\[
\frac{d}{dt}(f(r(t))) = \nabla f(r(t)) \cdot r'(t).
\]
To recap: we started with an elevation function \(f(x, y)\) and a parametrization \(r(t)\). There are then two vectors based at \(r(t)\) that represent some sort of derivative. The gradient \(\nabla f(r(t))\) represents a derivative of \(f\) at \(r(t)\), and \(r'(t)\), our velocity vector, is a derivative of \(r(t)\) at \(r'(t)\). Dotting these vectors together gives the derivative of the function \(f(r(t))\).

**Example 1.** Consider the elevation function \(f(x, y) = 150 - 2x^2 - 8y^2\) and the path \(r(t) = \langle \cos t, 2 \sin t \rangle\), \(t \in [0, 2\pi]\). Some level curves of \(f\) and the path \(r(t)\) are shown in Figure 3. How quickly is our elevation changing at time \(t = 3\pi/4\)? Identify the critical points of the elevation function along the path.

![Figure 3: Some level curves of \(f(x, y) = 150 - 2x^2 - 8y^2\) and the path \(r(t) = \langle \cos t, 2 \sin t \rangle\).](image.png)
Our elevation at time $t$ is given by $f(r(t))$. Because we have expressions for $f(x, y)$ and $r(t)$ we could make appropriate substitutions and use the tools of single-variable calculus, but instead let’s use the ideas developed above. We have

$$\nabla f(x, y) = \langle -4x, -16y \rangle,$$

and $r(3\pi/4) = \langle 1/\sqrt{2}, 2/\sqrt{2} \rangle$, so

$$\nabla f(r(3\pi/4)) = \left\langle -\frac{4}{\sqrt{2}}, -\frac{32}{\sqrt{2}} \right\rangle = \langle -2\sqrt{2}, -16\sqrt{2} \rangle.$$

We also have $r'(t) = \langle -\sin t, 2\cos t \rangle$, so

$$r'(3\pi/4) = \langle -1/\sqrt{2}, -\sqrt{2} \rangle.$$

Finally,

$$\left.\frac{d}{dt}(f(r(t)))\right|_{t=3\pi/4} = \nabla f(r(3\pi/4)) \cdot r'(3\pi/4) = (-2\sqrt{2})(-1/\sqrt{2}) + (-16\sqrt{2})(-\sqrt{2}) = 34.$$

Now the critical points of $f(r(t))$ will of course be the points where $\frac{d}{dt}(f(r(t)))$ vanishes. We have

$$\frac{d}{dt}(f(r(t))) = \nabla f(r(t)) \cdot r'(t) = \langle -4\cos t, -16(2\sin t) \rangle \cdot \langle -\sin t, 2\cos t \rangle$$

$$= 4\cos t \sin t - 64\sin t \cos t = -60\cos t \sin t.$$

So $f(r(t))$ will have a critical point any time either $\cos t$ or $\sin t$ is zero — that is, at $0, \pi/2, \pi, 3\pi/2$, and $2\pi$. Notice that these values correspond to the points where our path is tangent to a level curve of $f$. \hfill \Diamond

Next, let’s consider the following problem: we still have the elevation function $f(x, y)$, and we’re standing at the point $(a, b)$ (which we assume is not a critical point for $f$). We plan to move away from $(a, b)$ at unit speed and want to know in which direction we should travel so that

(a) our elevation increases as quickly as possible;

(b) our elevation decreases as quickly as possible;

(c) our elevation does not change.

If we move away from $(a, b)$ along a straight line, then our path can be parametrized as

$$r(t) = \langle a, b \rangle + tu, \quad t \geq 0,$$

where $u$ is a unit vector. In this case $r'(0) = u$ and the conditions we consider above involve

$$\left.\frac{d}{dt}(f(r(t)))\right|_{t=0} = \nabla f(r(0)) \cdot r'(0) = \nabla f(a, b) \cdot u.$$

This is the slope\(^1\) along our path, which we’re hoping to either maximize (in the case of (a)), minimize (in the case of (b)), or set to zero (in the case of (c)). (We assume that $\nabla f(a, b)$ is not

---

\(^1\)Indeed, since we’re moving at unit speed this is truly the slope of the tangent line to the curve $(x(t), y(t), f(x(t), y(t)))$ in 3-space.
the zero vector.) If \( \theta \) is the smaller angle between \( \mathbf{u} \) (so that \( 0 \leq \theta \leq \pi \)) and the vector \( \nabla f(a, b) \) then we can write

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

since \( \mathbf{u} \) is a unit vector. This expression is maximized when \( \cos \theta = 1 \), which occurs when \( \theta = 0 \). So if we want our elevation to increase as quickly as possible, we should move in the direction of \( \nabla f(a, b) \). Similarly, our elevation will decrease as quickly as possible if we move in the direction of \(-\nabla f(a, b)\), where we have \( \theta = \pi \) and \( \cos \theta = -1 \). Finally, our elevation will not change (instantaneously) if we have \( \theta = \pi/2 \), for in this case \( \cos \theta = 0 \) and \( \frac{d}{dt}(f(\mathbf{r}(t))) = 0 \). In particular, a line perpendicular to the gradient vector is tangent to a level curve. So we have the following observations:

1. At \((a, b)\) the direction of greatest increase for \( f \) is \( \nabla f(a, b) \).
2. The direction of greatest decrease is \(-\nabla f(a, b)\).
3. The vector \( \nabla f(a, b) \) is perpendicular to the level curve passing through \((a, b)\).

Remember that in the previous example the critical points of our elevation function (where our slope was zero) were the points where our velocity vector was tangent to a level curve of the elevation function. This agrees with our newer observations.

We actually have a name for the derivatives considered above. Say we have a function \( f(x, y) \), a point \( P = (a, b) \) and a vector \( \mathbf{v} \) based at \( P \). We define the **directional derivative** of \( f \) at \( P \) in the direction of \( \mathbf{v} \) to be

\[
D_{\mathbf{v}} f = \nabla f(P) \cdot \mathbf{u},
\]

where \( \mathbf{u} = \mathbf{v}/\|\mathbf{v}\| \) is the unit vector in the direction of \( \mathbf{v} \). As we learned in our above discussion, this number is the slope of the line tangent to \( z = f(x, y) \) passing through \((a, b, f(a, b))\) and moving in the direction of \( \mathbf{v} \).

**Example 2.** Find the directional derivative of \( f(x, y, z) = x^3 + yz \) at the point \( P = (-1, 3, -2) \) in the direction pointing to the origin.

**(Solution)** The vector pointing from \( P \) to the origin is given by \( \mathbf{v} = (1, -3, 2) \). The unit vector pointing in this direction is then

\[
\mathbf{u} = \frac{1}{\sqrt{1 + 9 + 4}} \mathbf{v} = \left( \frac{1}{\sqrt{14}}, -\frac{3}{\sqrt{14}}, \frac{2}{\sqrt{14}} \right).
\]

We also have

\[
\nabla f = (3x^2, z, y),
\]

so \( \nabla f(P) = (3, -2, 3) \). Then

\[
D_{\mathbf{u}} f = \nabla f(P) \cdot \mathbf{u} = (3, -2, 3) \cdot \left( \frac{1}{\sqrt{14}}, -\frac{3}{\sqrt{14}}, \frac{2}{\sqrt{14}} \right) = \frac{15}{\sqrt{14}}.
\]