Chapter 10

Derivatives of Multivariable Functions

10.1 Limits

Motivating Questions

In this section, we strive to understand the ideas generated by the following important questions:

• What do we mean by the limit of a function \( f \) of two variables at a point \( (a, b) \)?

• What techniques can we use to show that a function of two variables does not have a limit at a point \( (a, b) \)?

• What does it mean for a function \( f \) of two variables to be continuous at a point \( (a, b) \)?

Introduction

In this section, we will study limits of functions of several variables, with a focus on limits of functions of two variables. In single variable calculus, we studied the notion of limit, which turned out to be a critical concept that formed the basis for the derivative and the definite integral. In this section we will begin to understand how the concept of limit for functions of two variables is similar to what we encountered for functions of a single variable. The limit will again be the fundamental idea in multivariable calculus, and we will use this notion of the limit of a function of several variables to define the important concept of differentiability later in this chapter. We have already seen its use in the derivatives of vector-valued functions in Section 9.7.

Let’s begin by reviewing what we mean by the limit of a function of one variable. We say that a function \( f \) has a limit \( L \) as \( x \) approaches \( a \) provided that we can make the values \( f(x) \) as close to \( L \) as we like by taking \( x \) sufficiently close (but not equal) to \( a \). We denote this behavior by writing

\[
\lim_{x \to a} f(x) = L.
\]
Preview Activity 10.1. We investigate the limits of several different functions by working with tables and graphs.

(a) Consider the function $f$ defined by

$$f(x) = 3 - x.$$ 

Complete the following table of values.

<table>
<thead>
<tr>
<th>$x$</th>
<th>$f(x)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>−0.2</td>
<td></td>
</tr>
<tr>
<td>−0.1</td>
<td></td>
</tr>
<tr>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td>0.2</td>
<td></td>
</tr>
</tbody>
</table>

What does the table suggest regarding $\lim_{x \to 0} f(x)$?

(b) Explain how your results in (a) are reflected in Figure 10.1.

(c) Next, consider

$$g(x) = \frac{x}{|x|}. $$

Complete the following table of values near $x = 0$, the point at which $g$ is not defined.

<table>
<thead>
<tr>
<th>$x$</th>
<th>$g(x)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>−0.1</td>
<td></td>
</tr>
<tr>
<td>−0.01</td>
<td></td>
</tr>
<tr>
<td>−0.001</td>
<td></td>
</tr>
<tr>
<td>0.001</td>
<td></td>
</tr>
<tr>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>0.1</td>
<td></td>
</tr>
</tbody>
</table>

What does this suggest about $\lim_{x \to 0} g(x)$?
10.1. LIMITS

-1

1

-1

1

Figure 10.2: The graph of $g(x) = \frac{x}{|x|}$.

(d) Explain how your results in (c) are reflected in Figure 10.2.

(e) Now, let’s examine a function of two variables. Let

$$f(x, y) = 3 - x - 2y$$

and complete the following table of values.

<table>
<thead>
<tr>
<th>$x$</th>
<th>$y$</th>
<th>-1</th>
<th>-0.1</th>
<th>0</th>
<th>0.1</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-0.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

What does the table suggest about $\lim_{(x,y) \to (0,0)} f(x, y)$?

(f) Explain how your results in (e) are reflected in Figure 10.3. Compare this limit to the limit in part (a). How are the limits similar and how are they different?

(g) Finally, consider

$$g(x, y) = \frac{2xy}{x^2 + y^2},$$

which is not defined at $(0, 0)$, and complete the following table of values of $g(x, y)$.

<table>
<thead>
<tr>
<th>$x$</th>
<th>$y$</th>
<th>-1</th>
<th>-0.1</th>
<th>0</th>
<th>0.1</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-0.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

What does this suggest about $\lim_{(x,y) \to (0,0)} g(x, y)$?
Figure 10.3: At left, the graph of \( f(x, y) = 3 - x - 2y \); at right, its contour plot.

(h) Explain how your results are reflected in Figure 10.4. Compare this limit to the limit in part (b). How are the results similar and how are they different?

Figure 10.4: At left, the graph of \( g(x, y) = \frac{2xy}{x^2 + y^2} \); at right, its contour plot.

In Preview Activity 10.1, we recalled the notion of limit from single variable calculus and saw that a similar concept applies to functions of two variables. Though we will focus on functions of two variables, for the sake of discussion, all the ideas we establish here are valid for functions of any number of variables. In a natural followup to our work in Preview Activity 10.1, we now formally
define what it means for a function of two variables to have a limit at a point.

**Definition 10.1.** Given a function \( f = f(x, y) \), we say that \( f \) has **limit** \( L \) as \( (x, y) \) approaches \( (a, b) \) provided that we can make \( f(x, y) \) as close to \( L \) as we like by taking \( (x, y) \) sufficiently close (but not equal) to \( (a, b) \). We write

\[
\lim_{(x,y) \to (a,b)} f(x, y) = L.
\]

To investigate the limit of a single variable function, \( \lim_{x \to a} f(x) \), we often consider the behavior of \( f \) as \( x \) approaches \( a \) from the right and from the left. Similarly, we may investigate limits of two-variable functions, \( \lim_{(x,y) \to (a,b)} f(x, y) \) by considering the behavior of \( f \) as \( (x, y) \) approaches \( (a, b) \) from various directions. This situation is more complicated because there are infinitely many ways in which \( (x, y) \) may approach \( (a, b) \). In the next activity, we see how it is important to consider a variety of those paths in investigating whether or not a limit exists.

**Activity 10.1.**

Consider the function \( f \), defined by

\[
f(x, y) = \frac{y}{\sqrt{x^2 + y^2}},
\]

whose graph is shown below in Figure 10.5

![Figure 10.5: The graph of \( f(x, y) = \frac{y}{\sqrt{x^2 + y^2}} \).](image)

(a) Is \( f \) defined at the point \((0, 0)\)? What, if anything, does this say about whether \( f \) has a limit at the point \((0, 0)\)?

(b) Values of \( f \) (to three decimal places) at several points close to \((0, 0)\) are shown in the table below.
Based on these calculations, state whether $f$ has a limit at $(0, 0)$ and give an argument supporting your statement. (Hint: The blank spaces in the table are there to help you see the patterns.)

(c) Now let’s consider what happens if we restrict our attention to the $x$-axis; that is, consider what happens when $y = 0$. What is the behavior of $f(x, 0)$ as $x \to 0$? If we approach $(0, 0)$ by moving along the $x$-axis, what value do we find as the limit?

(d) What is the behavior of $f$ along the line $y = x$ when $x > 0$; that is, what is the value of $f(x, x)$ when $x > 0$? If we approach $(0, 0)$ by moving along the line $y = x$ in the first quadrant (thus considering $f(x, x)$ as $x \to 0$, what value do we find as the limit?

(e) In general, if $\lim_{(x, y) \to (0,0)} f(x, y) = L$, then $f(x, y)$ approaches $L$ as $(x, y)$ approaches $(0, 0)$, regardless of the path we take in letting $(x, y) \to (0, 0)$. Explain what the last two parts of this activity imply about the existence of $\lim_{(x, y) \to (0,0)} f(x, y)$.

(f) Shown below in Figure 10.6 is a set of contour lines of the function $f$. What is the behavior of $f(x, y)$ as $(x, y)$ approaches $(0, 0)$ along any straight line? How does this observation reinforce your conclusion about the existence of $\lim_{(x, y) \to (0,0)} f(x, y)$ from the previous part of this activity? (Hint: Use the fact that a non-vertical line has equation $y = mx$ for some constant $m$.)

<table>
<thead>
<tr>
<th>$x$ \ $y$</th>
<th>-1</th>
<th>-0.1</th>
<th>0</th>
<th>0.1</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1</td>
<td>-0.707</td>
<td>—</td>
<td>0</td>
<td>—</td>
<td>0.707</td>
</tr>
<tr>
<td>-0.1</td>
<td>—</td>
<td>-0.707</td>
<td>0</td>
<td>0.707</td>
<td>—</td>
</tr>
<tr>
<td>0</td>
<td>-1</td>
<td>-1</td>
<td>—</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>0.1</td>
<td>—</td>
<td>-0.707</td>
<td>0</td>
<td>0.707</td>
<td>—</td>
</tr>
<tr>
<td>1</td>
<td>-0.707</td>
<td>—</td>
<td>0</td>
<td>—</td>
<td>0.707</td>
</tr>
</tbody>
</table>

Figure 10.6: Contour lines of $f(x, y) = \frac{y}{\sqrt{x^2+y^2}}$. 
As we have seen in Activity 10.1, if we approach \((a, b)\) along two different paths and find that \(f(x, y)\) has two different limits, we can conclude that \(\lim_{(x,y)\to(a,b)} f(x,y)\) does not exist. This is similar to the one-variable example \(f(x) = x/|x|\) as shown in 10.7; \(\lim_{x\to0} f(x)\) does not exist because we see different limits as \(x\) approaches 0 from the left and the right.

![Figure 10.7: The graph of \(g(x) = \frac{x}{|x|}\).](image)

As a general rule, we have

If \(f(x, y)\) has two different limits as \((x, y)\) approaches \((a, b)\) along two different paths, then \(\lim_{(x,y)\to(a,b)} f(x,y)\) does not exist.

As the next activity shows, studying the limit of a two-variable function \(f\) by considering the behavior of \(f\) along various paths can require subtle insights.

**Activity 10.2.**

Let’s consider the function \(g\) defined by

\[
g(x, y) = \frac{x^2y}{x^4 + y^2}
\]

and investigate the limit \(\lim_{(x,y)\to(0,0)} g(x, y)\).

(a) What is the behavior of \(g\) on the \(x\)-axis? That is, what is \(g(x, 0)\) and what is the limit of \(g\) as \((x, y)\) approaches \((0, 0)\) along the \(x\)-axis?

(b) What is the behavior of \(g\) on the \(y\)-axis? That is, what is \(g(0, y)\) and what is the limit of \(g\) as \((x, y)\) approaches \((0, 0)\) along the \(y\)-axis?

(c) What is the behavior of \(g\) on the line \(y = mx\)? That is, what is \(g(x, mx)\) and what is the limit of \(g\) as \((x, y)\) approaches \((0, 0)\) along the line \(y = mx\)?

(d) Based on what you have seen so far, do you think \(\lim_{(x,y)\to(0,0)} g(x, y)\) exists? If so, what do you think its value is?

(e) Now consider the behavior of \(g\) on the parabola \(y = x^2\)? What is \(g(x, x^2)\) and what is the limit of \(g\) as \((x, y)\) approaches \((0, 0)\) along this parabola?
(f) State whether the limit \( \lim_{(x,y) \to (0,0)} g(x, y) \) exists or not and provide a justification of your statement.

This activity shows that we need to be careful when studying the limit of a two-variable function by considering its behavior along different paths. If we find two different paths that result in two different limits, then we may conclude that the limit does not exist. However, we can never conclude that the limit of a function exists only by considering its behavior along different paths.

Generally speaking, concluding that a limit \( \lim_{(x,y) \to (a,b)} f(x, y) \) exists requires a more careful argument. For example, let’s consider the function defined by

\[
f(x, y) = \frac{x^2 y^2}{x^2 + y^2}
\]

and ask whether \( \lim_{(x,y) \to (0,0)} f(x, y) \) exists.

Note that if either \( x \) or \( y \) is 0, then \( f(x, y) = 0 \). Therefore, if \( f \) has a limit at \((0, 0)\), it must be 0. We will therefore argue that

\[
\lim_{(x,y) \to (0,0)} f(x, y) = 0,
\]

by showing that we can make \( f(x, y) \) as close to 0 as we wish by taking \((x, y)\) sufficiently close (but not equal) to \((0, 0)\). In what follows, we view \( x \) and \( y \) as being real numbers that are close, but not equal, to 0.

Since \( 0 \leq x^2 \), we have

\[
y^2 \leq x^2 + y^2,
\]

which implies that

\[
\frac{y^2}{x^2 + y^2} \leq 1.
\]

Multiplying both sides by \( x^2 \) and observing that \( f(x, y) \geq 0 \) for all \((x, y)\) gives

\[
0 \leq f(x, y) = \frac{x^2 y^2}{x^2 + y^2} = x^2 \left( \frac{y^2}{x^2 + y^2} \right) \leq x^2.
\]

This shows that we can make \( f(x, y) \) as close to 0 as we like by taking \( x \) sufficiently close to 0 (for this example, it turns out that we don’t even need to worry about making \( y \) close to 0). Therefore,

\[
\lim_{(x,y) \to (0,0)} \frac{x^2 y^2}{x^2 + y^2} = 0.
\]

In spite of the fact that these two most recent examples illustrate some of the complications that arise when studying limits of two-variable functions, many of the properties that are familiar
Properties of Limits. Let \( f = f(x, y) \) and \( g = g(x, y) \) be functions so that \( \lim_{(x, y) \to (a, b)} f(x, y) \) and \( \lim_{(x, y) \to (a, b)} g(x, y) \) both exist. Then

1. \( \lim_{(x, y) \to (a, b)} x = a \) and \( \lim_{(x, y) \to (a, b)} y = b. \)

2. \( \lim_{(x, y) \to (a, b)} cf(x, y) = c \left( \lim_{(x, y) \to (a, b)} f(x, y) \right) \)
   for any scalar \( c, \)

3. \( \lim_{(x, y) \to (a, b)} [f(x, y) \pm g(x, y)] = \lim_{(x, y) \to (a, b)} f(x, y) \pm \lim_{(x, y) \to (a, b)} g(x, y), \)

4. \( \lim_{(x, y) \to (a, b)} [f(x, y) \cdot g(x, y)] = \left( \lim_{(x, y) \to (a, b)} f(x, y) \right) \cdot \left( \lim_{(x, y) \to (a, b)} g(x, y) \right), \)

5. \( \lim_{(x, y) \to (a, b)} \frac{f(x, y)}{g(x, y)} = \frac{\lim_{(x, y) \to (a, b)} f(x, y)}{\lim_{(x, y) \to (a, b)} g(x, y)} \)
   if \( \lim_{(x, y) \to (a, b)} g(x, y) \neq 0. \)

We can use these properties and results from single variable calculus to verify that many limits exist. For example, these properties show that the function \( f \) defined by

\[
 f(x, y) = 3x^2y^3 + 2xy^2 - 3x + 1
\]

has a limit at every point \((a, b)\) and, moreover,

\[
 \lim_{(x, y) \to (a, b)} f(x, y) = f(a, b).
\]

The reason for this is that polynomial functions of a single variable have limits at every point.

Continuity

Recall that a function \( f \) of a single variable \( x \) is said to be continuous at \( x = a \) provided that the following three conditions are satisfied:

1. \( f(a) \) exists,

2. \( \lim_{x \to a} f(x) \) exists, and

3. \( \lim_{x \to a} f(x) = f(a). \)

Using our understanding of limits of multivariable functions, we can define continuity in the same
Definition 10.2. A function $f = f(x, y)$ is **continuous** at the point $(a, b)$ provided that

1. $f$ is defined at the point $(a, b)$,
2. $\lim_{(x, y) \to (a, b)} f(x, y)$ exists, and
3. $\lim_{(x, y) \to (a, b)} f(x, y) = f(a, b)$.

For instance, we have seen that the function $f$ defined by $f(x, y) = 3x^2y^3 + 2xy^2 - 3x + 1$ is continuous at every point. And just as with single variable functions, continuity has certain properties that are based on the properties of limits.

**Properties of Continuity.** Let $f$ and $g$ be functions of two variables that are continuous at the point $(a, b)$. Then

1. $cf$ is continuous at $(a, b)$ for any scalar $c$
2. $f + g$ is continuous at $(a, b)$
3. $f - g$ is continuous at $(a, b)$
4. $fg$ is continuous at $(a, b)$
5. $\frac{f}{g}$ is continuous at $(a, b)$ if $g(a, b) \neq 0$

Using these properties, we can apply results from single variable calculus to decide about continuity of multivariable functions. For example, the coordinate functions $f$ and $g$ defined by $f(x, y) = x$ and $g(x, y) = y$ are continuous at every point. We can then use properties of continuity listed to conclude that every polynomial function in $x$ and $y$ is continuous at every point. For example, $g(x, y) = x^2$ and $h(x, y) = y^3$ are continuous functions, so their product $f(x, y) = x^2y^3$ is a continuous multivariable function.

**Summary**

- A function $f = f(x, y)$ has a limit $L$ at a point $(a, b)$ provided that we can make $f(x, y)$ as close to $L$ as we like by taking $(x, y)$ sufficiently close (but not equal) to $(a, b)$.
- If $(x, y)$ has two different limits as $(x, y)$ approaches $(a, b)$ along two different paths, we can conclude that $\lim_{(x, y) \to (a, b)} f(x, y)$ does not exist.
- Properties similar to those for one-variable functions allow us to conclude that many limits exist and to evaluate them.
10.1. LIMITS

- A function \( f = f(x, y) \) is continuous at a point \((a, b)\) in its domain if \( f \) has a limit at \((a, b)\) and

\[
f(a, b) = \lim_{(x,y) \to (a,b)} f(x,y).
\]

Exercises

1. Consider the function \( f \) defined by \( f(x, y) = \frac{xy}{x^2 + y^2 + 1} \).

   (a) What is the domain of \( f \)?
   (b) Evaluate limit of \( f \) at \((0, 0)\) along the following paths: \( x = 0, y = 0, y = x, \) and \( y = x^2 \).
   (c) What do you conjecture is the value of \( \lim_{(x,y) \to (0,0)} f(x,y) \)?
   (d) Is \( f \) continuous at \((0, 0)\)? Why or why not?
   (e) Use appropriate technology to sketch both surface and contour plots of \( f \) near \((0, 0)\). Write several sentences to say how your plots affirm your findings in (a) - (d).

2. Consider the function \( g \) defined by \( g(x, y) = \frac{xy}{x^2 + y^2} \).

   (a) What is the domain of \( g \)?
   (b) Evaluate limit of \( g \) at \((0, 0)\) along the following paths: \( x = 0, y = x, \) and \( y = 2x \).
   (c) What can you now say about the value of \( \lim_{(x,y) \to (0,0)} g(x,y) \)?
   (d) Is \( g \) continuous at \((0, 0)\)? Why or why not?
   (e) Use appropriate technology to sketch both surface and contour plots of \( g \) near \((0, 0)\). Write several sentences to say how your plots affirm your findings in (a) - (d).

3. For each of the following prompts, provide an example of a function of two variables with the desired properties (with justification), or explain why such a function does not exist.

   (a) A function \( p \) that is defined at \((0, 0)\), but \( \lim_{(x,y) \to (0,0)} p(x,y) \) does not exist.
   (b) A function \( q \) that does not have a limit at \((0, 0)\), but that has the same limiting value along any line \( y = mx \) as \( x \to 0 \).
   (c) A function \( r \) that is continuous at \((0, 0)\), but \( \lim_{(x,y) \to (0,0)} r(x,y) \) does not exist.
   (d) A function \( s \) such that

\[
\lim_{(x,x) \to (0,0)} s(x,x) = 3 \quad \text{and} \quad \lim_{(x,2x) \to (0,0)} s(x,2x) = 6,
\]

for which \( \lim_{(x,y) \to (0,0)} s(x,y) \) exists.
(e) A function \( t \) that is not defined at \((1,1)\) but \( \lim_{(x,y)\to(1,1)} t(x,y) \) does exist.
10.2 First-Order Partial Derivatives

Motivating Questions

In this section, we strive to understand the ideas generated by the following important questions:

- How are the first-order partial derivatives of a function \( f \) of the independent variables \( x \) and \( y \) defined?
- Given a function \( f \) of the independent variables \( x \) and \( y \), what do the first-order partial derivatives \( \frac{\partial f}{\partial x} \) and \( \frac{\partial f}{\partial y} \) tell us about \( f \)?

Introduction

The derivative plays a central role in first semester calculus because it provides important information about a function. Thinking graphically, for instance, the derivative at a point tells us the slope of the tangent line to the graph at that point. In addition, the derivative at a point also provides the instantaneous rate of change of the function with respect to changes in the independent variable.

Now that we are investigating functions of two or more variables, we can still ask how fast the function is changing, though we have to be careful about what we mean. Thinking graphically again, we can try to measure how steep the graph of the function is in a particular direction. Alternatively, we may want to know how fast a function’s output changes in response to a change in one of the inputs. Over the next few sections, we will develop tools for addressing issues such as these these. Preview Activity 10.2 explores some issues with what we will come to call partial derivatives.

Preview Activity 10.2. Let’s return to the function we considered in Preview Activity 9.1. Suppose we take out a $18,000 car loan at interest rate \( r \) and we agree to pay off the loan in \( t \) years. The monthly payment, in dollars, is

\[
M(r, t) = \frac{1500r}{1 - (1 + \frac{r}{12})^{-12t}}.
\]

(a) What is the monthly payment if the interest rate is 3\% so that \( r = 0.03 \), and we pay the loan off in \( t = 4 \) years?

(b) Suppose the interest rate is fixed at 3\%. Express \( M \) as a function \( f \) of \( t \) alone using \( r = 0.03 \). That is, let \( f(t) = M(0.03, t) \). Sketch the graph of \( f \) on the left of Figure 10.8. Explain the meaning of the function \( f \).

(c) Find the instantaneous rate of change \( f'(4) \) and state the units on this quantity. What information does \( f'(4) \) tell us about our car loan? What information does \( f'(4) \) tell us about the graph you sketched in (b)?
(d) Express \( M \) as a function of \( r \) alone, using a fixed time of \( t = 4 \). That is, let \( g(r) = M(r, 4) \). Sketch the graph of \( g \) on the right of Figure 10.8. Explain the meaning of the function \( g \).

(e) Find the instantaneous rate of change \( g'(0.03) \) and state the units on this quantity. What information does \( g'(0.03) \) tell us about our car loan? What information does \( g'(0.03) \) tell us about the graph you sketched in (d)?

First-Order Partial Derivatives

In Section 9.1, we studied the behavior of a function of two or more variables by considering the traces of the function. Recall that in one example, we considered the function \( f \) defined by

\[
f(x, y) = \frac{x^2 \sin(2y)}{32},
\]

which measures the range, or horizontal distance, in feet, traveled by a projectile launched with an initial speed of \( x \) feet per second at an angle \( y \) radians to the horizontal. The graph of this function is given again on the left in Figure 10.9. Moreover, if we fix the angle \( y = 0.6 \), we may view the trace \( f(x, 0.6) \) as a function of \( x \) alone, as seen at right in Figure 10.9.

Since the trace is a one-variable function, we may consider its derivative just as we did in the first semester of calculus. With \( y = 0.6 \), we have

\[
f(x, 0.6) = \frac{\sin(1.2)}{32} x^2,
\]

and therefore

\[
d \frac{d}{dx} [f(x, 0.6)] = \frac{\sin(1.2)}{16} x.
\]

When \( x = 150 \), this gives

\[
d \frac{d}{dx} [f(x, 0.6)]|_{x=150} = \frac{\sin(1.2)}{16} 150 \approx 8.74 \text{ feet per feet per second},
\]
which gives the slope of the tangent line shown on the right of Figure 10.9. Thinking of this derivative as an instantaneous rate of change implies that if we increase the initial speed of the projectile by one foot per second, we expect the horizontal distance traveled to increase by approximately 8.74 feet if we hold the launch angle constant at 0.6 radians.

By holding \( y \) fixed and differentiating with respect to \( x \), we obtain the first-order partial derivative of \( f \) with respect to \( x \). Denoting this partial derivative as \( f_x \), we have seen that

\[
f_x(150, 0.6) = \frac{d}{dx} f(x, 0.6)|_{x=150} = \lim_{h \to 0} \frac{f(150 + h, 0.6) - f(150, 0.6)}{h}.
\]

More generally, we have

\[
f_x(a, b) = \lim_{h \to 0} \frac{f(a + h, b) - f(a, b)}{h},
\]

provided this limit exists.

In the same way, we may obtain a trace by setting, say, \( x = 150 \) as shown in Figure 10.10. This gives

\[
f(150, y) = \frac{150^2}{32} \sin(2y),
\]

and therefore

\[
\frac{d}{dy} [f(150, y)] = \frac{150^2}{16} \cos(2y).
\]

If we evaluate this quantity at \( y = 0.6 \), we have

\[
\frac{d}{dy} [f(150, y)]|_{y=0.6} = \frac{150^2}{16} \cos(1.2) \approx 509.5 \text{ feet per radian}.
\]

Once again, the derivative gives the slope of the tangent line shown on the right in Figure 10.10. Thinking of the derivative as an instantaneous rate of change, we expect that the range of the
projectile increases by 509.5 feet for every radian we increase the launch angle \( y \) if we keep the initial speed of the projectile constant at 150 feet per second.

By holding \( x \) fixed and differentiating with respect to \( y \), we obtain the first-order partial derivative of \( f \) with respect to \( y \). As before, we denote this partial derivative as \( f_y \) and write

\[
f_y(150, 0.6) = \frac{d}{dy} f(150, y) \big|_{y=0.6} = \lim_{h \to 0} \frac{f(150, 0.6 + h) - f(150, 0.6)}{h}.
\]

As with the partial derivative with respect to \( x \), we may express this quantity more generally at an arbitrary point \((a, b)\). To recap, we have now arrived at the formal definition of the first-order partial derivatives of a function of two variables.

**Definition 10.3.** The first-order partial derivatives of \( f \) with respect to \( x \) and \( y \) at a point \((a, b)\) are, respectively,

\[
f_x(a, b) = \lim_{h \to 0} \frac{f(a + h, b) - f(a, b)}{h}, \quad \text{and} \quad f_y(a, b) = \lim_{h \to 0} \frac{f(a, b + h) - f(a, b)}{h},
\]

provided the limits exist.

**Activity 10.3.**

Consider the function \( f \) defined by

\[
f(x, y) = \frac{xy^2}{x + 1}
\]
at the point \((1, 2)\).

(a) Write the trace \(f(x, 2)\) at the fixed value \(y = 2\). On the left side of Figure 10.11, draw the graph of the trace with \(y = 2\) indicating the scale and labels on the axes. Also, sketch the tangent line at the point \(x = 1\).

(b) Find the partial derivative \(f_x(1, 2)\) and relate its value to the sketch you just made.

(c) Write the trace \(f(1, y)\) at the fixed value \(x = 1\). On the right side of Figure 10.11, draw the graph of the trace with \(x = 1\) indicating the scale and labels on the axes. Also, sketch the tangent line at the point \(y = 2\).

(d) Find the partial derivative \(f_y(1, 2)\) and relate its value to the sketch you just made.

As these examples show, each partial derivative at a point arises as the derivative of a one-variable function defined by fixing one of the coordinates. In addition, we may consider each partial derivative as defining a new function of the point \((x, y)\), just as the derivative \(f'(x)\) defines a new function of \(x\) in single-variable calculus. Due to the connection between one-variable derivatives and partial derivatives, we will often use Leibniz-style notation to denote partial derivatives by writing

\[
\frac{\partial f}{\partial x}(a, b) = f_x(a, b), \quad \text{and} \quad \frac{\partial f}{\partial y}(a, b) = f_y(a, b).
\]

To calculate the partial derivative \(f_x\), we hold \(y\) fixed and thus we treat \(y\) as a constant. In Leibniz notation, observe that

\[
\frac{\partial}{\partial x}(x) = 1 \quad \text{and} \quad \frac{\partial}{\partial x}(y) = 0.
\]
To see the contrast between how we calculate single variable derivatives and partial derivatives, and the difference between the notations \( \frac{d}{dx}[\ ] \) and \( \frac{\partial}{\partial x}[\ ] \), observe that

\[
\frac{d}{dx}[3x^2 - 2x + 3] = 3 \frac{d}{dx}[x^2] - 2 \frac{d}{dx}[x] + \frac{d}{dx}[3] = 3 \cdot 2x - 2,
\]

and

\[
\frac{\partial}{\partial x}[x^2y - xy + 2y] = \frac{\partial}{\partial x}[x^2] - y \frac{\partial}{\partial x}[x] + \frac{\partial}{\partial x}[2y] = y \cdot 2x - y.
\]

Thus, computing partial derivatives is straightforward: we use the standard rules of single variable calculus, but do so while holding one (or more) of the variables constant.

**Activity 10.4.**

(a) If we have the function \( f \) of the variables \( x \) and \( y \) and we want to find the partial derivative \( f_x \), which variable do we treat as a constant? When we find the partial derivative \( f_y \), which variable do we treat as a constant?

(b) If \( f(x, y) = 3x^3 - 2x^2y^5 \), find the partial derivatives \( f_x \) and \( f_y \).

(c) If \( f(x, y) = \frac{xy^2}{x + 1} \), find the partial derivatives \( f_x \) and \( f_y \).

(d) If \( g(r, s) = rs \cos(r) \), find the partial derivatives \( g_r \) and \( g_s \).

(e) Assuming \( f(w, x, y) = (6w + 1) \cos(3x^2 + 4xy^3 + y) \), find the partial derivatives \( f_w \), \( f_x \), and \( f_y \).

(f) Find all possible first-order partial derivatives of \( q(x, t, z) = \frac{x^2t^3}{1 + x^2} \).

**Interpretations of First-Order Partial Derivatives**

Recall that the derivative of a single variable function has a geometric interpretation as the slope of the line tangent to the graph at a given point. Similarly, we have seen that the partial derivatives measure the slope of a line tangent to a trace of a function of two variables as shown in Figure 10.12.

Now we consider the first-order partial derivatives in context. Recall that the difference quotient \( \frac{f(a+h)-f(a)}{h} \) for a function \( f \) of a single variable \( x \) at a point where \( x = a \) tells us the average rate of change of \( f \) over the interval \([a, a+h]\), while the derivative \( f'(a) \) tells us the instantaneous rate of change of \( f \) at \( x = a \). We can use these same concepts to explain the meanings of the partial derivatives in context.

**Activity 10.5.**

The speed of sound \( C \) traveling through ocean water is a function of temperature, salinity and depth. It may be modeled by the function

\[
C = 1449.2 + 4.6T - 0.055T^2 + 0.00029T^3 + (1.34 - 0.01T)(S - 35) + 0.016D.
\]
Here $C$ is the speed of sound in meters/second, $T$ is the temperature in degrees Celsius, $S$ is the salinity in grams/liter of water, and $D$ is the depth below the ocean surface in meters.

(a) State the units in which each of the partial derivatives, $C_T$, $C_S$ and $C_D$, are expressed and explain the physical meaning of each.

(b) Find the partial derivatives $C_T$, $C_S$ and $C_D$.

(c) Evaluate each of the three partial derivatives at the point where $T = 10$, $S = 35$ and $D = 100$. What does the sign of each partial derivatives tell us about the behavior of the function $C$ at the point $(10, 35, 100)$?

Using tables and contours to estimate partial derivatives

Remember that functions of two variables are often represented as either a table of data or a contour plot. In single variable calculus, we saw how we can use the difference quotient to approximate derivatives if, instead of an algebraic formula, we only know the value of the function at a few points. The same idea applies to partial derivatives.

Activity 10.6.

The wind chill, as frequently reported, is a measure of how cold it feels outside when the wind is blowing. In Table 10.1, the wind chill $w$, measured in degrees Fahrenheit, is a function of the wind speed $v$, measured in miles per hour, and the ambient air temperature $T$, also measured in degrees Fahrenheit. We thus view $w$ as being of the form $w = w(v, T)$.

(a) Estimate the partial derivative $w_v(20, -10)$. What are the units on this quantity and what does it mean?
Table 10.1: Wind chill as a function of wind speed and temperature.

<table>
<thead>
<tr>
<th>$v \setminus T$</th>
<th>-30</th>
<th>-25</th>
<th>-20</th>
<th>-15</th>
<th>-10</th>
<th>-5</th>
<th>0</th>
<th>5</th>
<th>10</th>
<th>15</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>-46</td>
<td>-40</td>
<td>-34</td>
<td>-28</td>
<td>-22</td>
<td>-16</td>
<td>-11</td>
<td>-5</td>
<td>1</td>
<td>7</td>
<td>13</td>
</tr>
<tr>
<td>10</td>
<td>-53</td>
<td>-47</td>
<td>-41</td>
<td>-35</td>
<td>-28</td>
<td>-22</td>
<td>-16</td>
<td>-10</td>
<td>-4</td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td>25</td>
<td>-64</td>
<td>-58</td>
<td>-51</td>
<td>-44</td>
<td>-37</td>
<td>-31</td>
<td>-24</td>
<td>-17</td>
<td>-11</td>
<td>-4</td>
<td>3</td>
</tr>
<tr>
<td>30</td>
<td>-67</td>
<td>-60</td>
<td>-53</td>
<td>-46</td>
<td>-39</td>
<td>-33</td>
<td>-26</td>
<td>-19</td>
<td>-12</td>
<td>-5</td>
<td>1</td>
</tr>
<tr>
<td>35</td>
<td>-69</td>
<td>-62</td>
<td>-55</td>
<td>-48</td>
<td>-41</td>
<td>-34</td>
<td>-27</td>
<td>-21</td>
<td>-14</td>
<td>-7</td>
<td>0</td>
</tr>
<tr>
<td>40</td>
<td>-71</td>
<td>-64</td>
<td>-57</td>
<td>-50</td>
<td>-43</td>
<td>-36</td>
<td>-29</td>
<td>-22</td>
<td>-15</td>
<td>-8</td>
<td>-1</td>
</tr>
</tbody>
</table>

(b) Estimate the partial derivative $w_T(20, -10)$. What are the units on this quantity and what does it mean?

(c) Use your results to estimate the wind chill $w(18, -10)$.

(d) Use your results to estimate the wind chill $w(20, -12)$.

(e) Use your results to estimate the wind chill $w(18, -12)$.

Activity 10.7.

Shown below in Figure 10.13 is a contour plot of a function $f$. The value of the function along a few of the contours is indicated to the left of the figure.

Figure 10.13: A contour plot of $f$.

(a) Estimate the partial derivative $f_x(-2, -1)$.

(b) Estimate the partial derivative $f_y(-2, -1)$.

(c) Estimate the partial derivatives $f_x(-1, 2)$ and $f_y(-1, 2)$. 
(d) Locate one point \((x, y)\) where the partial derivative \(f_x(x, y) = 0\).

(e) Locate one point \((x, y)\) where \(f_x(x, y) < 0\).

(f) Locate one point \((x, y)\) where \(f_y(x, y) > 0\).

(g) Suppose you have a different function \(g\), and you know that \(g(2, 2) = 4\), \(g_x(2, 2) > 0\), and \(g_y(2, 2) > 0\). Using this information, sketch a possibility for the contour \(g(x, y) = 4\) passing through \((2, 2)\) on the left side of Figure 10.14. Then include possible contours \(g(x, y) = 3\) and \(g(x, y) = 5\).

(h) Suppose you have yet another function \(h\), and you know that \(h(2, 2) = 4\), \(h_x(2, 2) < 0\), and \(h_y(2, 2) > 0\). Using this information, sketch a possible contour \(h(x, y) = 4\) passing through \((2, 2)\) on the right side of Figure 10.14. Then include possible contours \(h(x, y) = 3\) and \(h(x, y) = 5\).

\[\begin{align*}
\frac{\partial f}{\partial x}(x, y) &= f_x(x, y) = \lim_{h \to 0} \frac{f(x + h, y) - f(x, y)}{h}, \\
\frac{\partial f}{\partial y}(x, y) &= f_y(x, y) = \lim_{h \to 0} \frac{f(x, y + h) - f(x, y)}{h},
\end{align*}\]

where each partial derivative exists only at those points \((x, y)\) for which the limit exists.

- The partial derivative \(f_x(a, b)\) tells us the instantaneous rate of change of \(f\) with respect to \(x\) at \((x, y) = (a, b)\) when \(y\) is fixed at \(b\). Geometrically, the partial derivative \(f_x(a, b)\) tells us the slope of the line tangent to the \(y = b\) trace of the function \(f\) at the point \((a, b, f(a, b))\).
The partial derivative \( f_y(a, b) \) tells us the instantaneous rate of change of \( f \) with respect to \( y \) at \((x, y) = (a, b)\) when \( x \) is fixed at \( a \). Geometrically, the partial derivative \( f_y(a, b) \) tells us the slope of the line tangent to the \( x = a \) trace of the function \( f \) at the point \((a, b, f(a, b))\).

### Exercises

1. The Heat Index, \( I \), (measured in apparent degrees F) is a function of the actual temperature \( T \) outside (in degrees F) and the relative humidity \( H \) (measured as a percentage). A portion of the table which gives values for this function, \( I = I(T, H) \), is reproduced below:

<table>
<thead>
<tr>
<th>( T ) ↓</th>
<th>( H ) →</th>
<th>70</th>
<th>75</th>
<th>80</th>
<th>85</th>
</tr>
</thead>
<tbody>
<tr>
<td>90</td>
<td></td>
<td>106</td>
<td>109</td>
<td>112</td>
<td>115</td>
</tr>
<tr>
<td>92</td>
<td></td>
<td>112</td>
<td>115</td>
<td>119</td>
<td>123</td>
</tr>
<tr>
<td>94</td>
<td></td>
<td>118</td>
<td>122</td>
<td>127</td>
<td>132</td>
</tr>
<tr>
<td>96</td>
<td></td>
<td>125</td>
<td>130</td>
<td>135</td>
<td>141</td>
</tr>
</tbody>
</table>

(a) State the limit definition of the value \( I_T(94, 75) \). Then, estimate \( I_T(94, 75) \), and write one complete sentence that carefully explains the meaning of this value, including its units.

(b) State the limit definition of the value \( I_H(94, 75) \). Then, estimate \( I_H(94, 75) \), and write one complete sentence that carefully explains the meaning of this value, including its units.

(c) Suppose you are given that \( I_T(92, 80) = 3.75 \) and \( I_H(92, 80) = 0.8 \). Estimate the values of \( I(91, 80) \) and \( I(92, 78) \). Explain how the partial derivatives are relevant to your thinking.

(d) On a certain day, at 1 p.m. the temperature is 92 degrees and the relative humidity is 85%. At 3 p.m., the temperature is 96 degrees and the relative humidity 75%. What is the average rate of change of the heat index over this time period, and what are the units on your answer? Write a sentence to explain your thinking.

2. Let \( f(x, y) = \frac{1}{2}xy^2 \) represent the kinetic energy in Joules of an object of mass \( x \) in kilograms with velocity \( y \) in meters per second. Let \((a, b)\) be the point \((4, 5)\) in the domain of \( f \).

(a) Calculate \( f_x(a, b) \).

(b) Explain as best you can in the context of kinetic energy what the partial derivative

\[
f_x(a, b) = \lim_{h \to 0} \frac{f(a + h, b) - f(a, b)}{h}
\]

tells us about kinetic energy.

(c) Calculate \( f_y(a, b) \).
(d) Explain as best you can in the context of kinetic energy what the partial derivative

\[ f_y(a, b) = \lim_{h \to 0} \frac{f(a, b + h) - f(a, b)}{h} \]

tells us about kinetic energy.

(e) Often we are given certain graphical information about a function instead of a rule. We can use that information to approximate partial derivatives. For example, suppose that we are given a contour plot of the kinetic energy function (as in Figure 10.15) instead of a formula. Use this contour plot to approximate \( f_x(4, 5) \) and \( f_y(4, 5) \) as best you can. Compare to your calculations from earlier parts of this exercise.

3. The temperature on an unevenly heated metal plate positioned in the first quadrant of the \( x-y \) plane is given by

\[ C(x, y) = \frac{25xy + 25}{(x - 1)^2 + (y - 1)^2 + 1}. \]

Assume that temperature is measured in degrees Celsius and that \( x \) and \( y \) are each measured in inches. (Note: At no point in the following questions should you expand the denominator of \( C(x, y) \).)

(a) Determine \( \frac{\partial C}{\partial x}(x, y) \) and \( \frac{\partial C}{\partial y}(x, y) \).

(b) If an ant is on the metal plate, standing at the point \((2, 3)\), and starts walking in a direction parallel to the \( y \) axis, at what rate will the temperature he is experiencing change? Explain, and include appropriate units.

(c) If an ant is walking along the line \( y = 3 \), at what instantaneous rate will the temperature he is experiencing change when he passes the point \((1, 3)\)?

(d) Now suppose the ant is stationed at the point \((6, 3)\) and walks in a straight line towards the point \((2, 0)\). Determine the average rate of change in temperature (per unit distance
traveled) the ant encounters in moving between these two points. Explain your reasoning carefully. What are the units on your answer?

4. Consider the function \( f \) defined by \( f(x, y) = 8 - x^2 - 3y^2 \).

(a) Determine \( f_x(x, y) \) and \( f_y(x, y) \).

(b) Find parametric equations in \( \mathbb{R}^3 \) for the tangent line to the trace \( f(x, 1) \) at \( x = 2 \).

(c) Find parametric equations in \( \mathbb{R}^3 \) for the tangent line to the trace \( f(2, y) \) at \( y = 1 \).

(d) State respective direction vectors for the two lines determined in (b) and (c).

(e) Determine the equation of the plane that passes through the point \( (2, 1, f(2, 1)) \) whose normal vector is orthogonal to the direction vectors of the two lines found in (b) and (c).

(f) Use a graphing utility to plot both the surface \( z = 8 - x^2 - 3y^2 \) and the plane from (e) near the point \( (2, 1) \). What is the relationship between the surface and the plane?
10.3 Second-Order Partial Derivatives

Motivating Questions

In this section, we strive to understand the ideas generated by the following important questions:

- Given a function \( f \) of two independent variables \( x \) and \( y \), how are the second-order partial derivatives of \( f \) defined?
- What do the second-order partial derivatives \( f_{xx}, f_{yy}, f_{xy}, \) and \( f_{yx} \) of a function \( f \) tell us about the function’s behavior?

Introduction

Recall that for a single-variable function \( f \), the second derivative of \( f \) is defined to be the derivative of the first derivative. That is, \( f''(x) = \frac{d}{dx}[f'(x)] \), which can be stated in terms of the limit definition of the derivative by writing

\[
f''(x) = \lim_{h \to 0} \frac{f'(x + h) - f'(x)}{h}.
\]

In what follows, we begin exploring the four different second-order partial derivatives of a function of two variables and seek to understand what these various derivatives tell us about the function’s behavior.

Preview Activity 10.3. Once again, let’s consider the function \( f \) defined by \( f(x, y) = \frac{x^2 \sin(2y)}{32} \) that measures a projectile’s range as a function of its initial speed \( x \) and launch angle \( y \). The graph of this function, including traces with \( x = 150 \) and \( y = 0.6 \), is shown in Figure 10.16.

![Figure 10.16: The range function with traces \( y = 0.6 \) and \( x = 150 \).](image)

(a) Compute the partial derivative \( f_x \) and notice that \( f_x \) itself is a new function of \( x \) and \( y \).
(b) We may now compute the partial derivatives of \( f_x \). Find the partial derivative \( f_{xx} = (f_x)_x \) and evaluate \( f_{xx}(150, 0.6) \).

(c) Figure 10.17 shows the trace of \( f \) with \( y = 0.6 \) with three tangent lines included. Explain how your result from part (b) of this preview activity is reflected in this figure.

(d) Determine the partial derivative \( f_y \), and then find the partial derivative \( f_{yy} = (f_y)_y \). Evaluate \( f_{yy}(150, 0.6) \).

(e) Figure 10.18 shows the trace \( f(150, y) \) and includes three tangent lines. Explain how the value of \( f_{yy}(150, 0.6) \) is reflected in this figure.

(f) Because \( f_x \) and \( f_y \) are each functions of both \( x \) and \( y \), they each have two partial derivatives. Not only can we compute \( f_{xx} = (f_x)_x \), but also \( f_{xy} = (f_x)_y \); likewise, in addition to \( f_{yy} = (f_y)_y \), also \( f_{yx} = (f_y)_x \). For the range function \( f(x, y) = \frac{x^2 \sin(2y)}{32} \), use your earlier computations of \( f_x \) and \( f_y \) to now determine \( f_{xy} \) and \( f_{yx} \). Write one sentence to explain how you calculated these “mixed” partial derivatives.
Second-order Partial Derivatives

A function $f$ of two independent variables $x$ and $y$ has two first order partial derivatives, $f_x$ and $f_y$. As we saw in Preview Activity 10.3, each of these first-order partial derivatives has two partial derivatives, giving a total of four second-order partial derivatives:

- $f_{xx} = (f_x)_x = \frac{\partial}{\partial x} \left( \frac{\partial f}{\partial x} \right) = \frac{\partial^2 f}{\partial x^2}$,
- $f_{yy} = (f_y)_y = \frac{\partial}{\partial y} \left( \frac{\partial f}{\partial y} \right) = \frac{\partial^2 f}{\partial y^2}$,
- $f_{xy} = (f_x)_y = \frac{\partial}{\partial y} \left( \frac{\partial f}{\partial x} \right) = \frac{\partial^2 f}{\partial y \partial x}$,
- $f_{yx} = (f_y)_x = \frac{\partial}{\partial x} \left( \frac{\partial f}{\partial y} \right) = \frac{\partial^2 f}{\partial x \partial y}$.

The first two are called *unmixed* second-order partial derivatives while the last two are called the *mixed* second-order partial derivatives.

One aspect of this notation can be a little confusing. The notation

$$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y} \left( \frac{\partial f}{\partial x} \right)$$

means that we first differentiate with respect to $x$ and then with respect to $y$; this can be expressed in the alternate notation $f_{xy} = (f_x)_y$. However, to find the second partial derivative

$$f_{yx} = (f_y)_x$$

we first differentiate with respect to $y$ and then $x$. This means that

$$\frac{\partial^2 f}{\partial y \partial x} = f_{xy}, \quad \text{and} \quad \frac{\partial^2 f}{\partial x \partial y} = f_{yx}.$$  

Be sure to note carefully the difference between Leibniz notation and subscript notation and the order in which $x$ and $y$ appear in each. In addition, remember that anytime we compute a partial derivative, we hold constant the variable(s) other than the one we are differentiating with respect to.

**Activity 10.8.**

Find all second order partial derivatives of the following functions. For each partial derivative you calculate, state explicitly which variable is being held constant.

(a) $f(x, y) = x^2 y^3$
(b) $f(x, y) = y \cos(x)$
(c) \( g(s, t) = st^3 + s^4 \)

(d) How many second order partial derivatives does the function \( h \) defined by \( h(x, y, z) = 9x^9z - xyz + 9 \) have? Find \( h_{xz} \) and \( h_{zx} \).

In Preview Activity 10.3 and Activity 10.8, you may have noticed that the mixed second-order partial derivatives are equal. This observation holds generally and is known as Clairaut’s Theorem.

**Clairaut’s Theorem.** Let \( f \) be a function of two variables for which the partial derivatives \( f_{xy} \) and \( f_{yx} \) are continuous near the point \((a, b)\). Then

\[
f_{xy}(a, b) = f_{yx}(a, b).
\]

**Interpreting the second-order Partial Derivatives**

Recall from single variable calculus that the second derivative measures the instantaneous rate of change of the derivative. This observation is the key to understanding the meaning of the second-order partial derivatives.

Furthermore, we remember that the second derivative of a function at a point provides us with information about the concavity of the function at that point. Since the unmixed second-order partial derivative \( f_{xx} \) requires us to hold \( y \) constant and differentiate twice with respect to \( x \), we may simply view \( f_{xx} \) as the second derivative of a trace of \( f \) where \( y \) is fixed. As such, \( f_{xx} \) will measure the concavity of this trace.

Consider, for example, \( f(x, y) = \sin(x)e^{-y} \). Figure 10.19 shows the graph of this function along with the trace given by \( y = -1.5 \). Also shown are three tangent lines to this trace, with increasing...
Second-Order Partial Derivatives

$x$-values from left to right among the three plots in Figure 10.19.

That the slope of the tangent line is decreasing as $x$ increases is reflected, as it is in one-variable calculus, in the fact that the trace is concave down. Indeed, we see that $f_x(x,y) = \cos(x)e^{-y}$ and so $f_{xx}(x,y) = -\sin(x)e^{-y} < 0$, since $e^{-y} > 0$ for all values of $y$, including $y = -1.5$.

In the following activity, we further explore what second-order partial derivatives tell us about the geometric behavior of a surface.

**Activity 10.9.**

We continue to consider the function $f$ defined by $f(x,y) = \sin(x)e^{-y}$.

(a) In Figure 10.20, we see the trace of $f(x,y) = \sin(x)e^{-y}$ that has $x$ held constant with $x = 1.75$. Write a couple of sentences that describe whether the slope of the tangent lines to this curve increase or decrease as $y$ increases, and, after computing $f_{yy}(x,y)$, explain how this observation is related to the value of $f_{yy}(1.75, y)$. Be sure to address the notion of concavity in your response.

![Figure 10.20: The tangent lines to a trace with increasing $y$.](image)

(b) In Figure 10.21, we start to think about the mixed partial derivative, $f_{xy}$. Here, we first hold $y$ constant to generate the first-order partial derivative $f_x$, and then we hold $x$ constant to compute $f_{xy}$. This leads to first thinking about a trace with $x$ being constant, followed by slopes of tangent lines in the $y$-direction that slide along the original trace. You might think of sliding your pencil down the trace with $x$ constant in a way that its slope indicates $(f_x)_y$ in order to further animate the three snapshots shown in the figure. Based on Figure 10.21, is $f_{xy}(1.75, -1.5)$ positive or negative? Why?

(c) Determine the formula for $f_{xy}(x,y)$, and hence evaluate $f_{xy}(1.75, -1.5)$. How does this value compare with your observations in (b)?

(d) We know that $f_{xx}(1.75, -1.5)$ measures the concavity of the $y = -1.5$ trace, and that $f_{yy}(1.75, -1.5)$ measures the concavity of the $x = 1.75$ trace. What do you think the quantity $f_{xy}(1.75, -1.5)$ measures?
10.3. SECOND-ORDER PARTIAL DERIVATIVES

(e) On Figure 10.21, sketch the trace with \( y = -1.5 \), and sketch three tangent lines whose slopes correspond to the value of \( f_{yx}(x, -1.5) \) for three different values of \( x \), the middle of which is \( x = -1.5 \). Is \( f_{yx}(1.75, -1.5) \) positive or negative? Why? What does \( f_{yx}(1.75, -1.5) \) measure?

Just as with the first-order partial derivatives, we can approximate second-order partial derivatives in the situation where we have only partial information about the function.

Activity 10.10.

As we saw in Activity 10.6, the wind chill \( w(v, T) \), in degrees Fahrenheit, is a function of the wind speed, in miles per hour, and the air temperature, in degrees Fahrenheit. Some values of the wind chill are recorded in Table 10.2.

<table>
<thead>
<tr>
<th>( v )</th>
<th>-30</th>
<th>-25</th>
<th>-20</th>
<th>-15</th>
<th>-10</th>
<th>-5</th>
<th>0</th>
<th>5</th>
<th>10</th>
<th>15</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>-46</td>
<td>-40</td>
<td>-34</td>
<td>-28</td>
<td>-22</td>
<td>-16</td>
<td>-11</td>
<td>-5</td>
<td>1</td>
<td>7</td>
<td>13</td>
</tr>
<tr>
<td>10</td>
<td>-53</td>
<td>-47</td>
<td>-41</td>
<td>-35</td>
<td>-28</td>
<td>-22</td>
<td>-16</td>
<td>-10</td>
<td>-4</td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td>25</td>
<td>-64</td>
<td>-58</td>
<td>-51</td>
<td>-44</td>
<td>-37</td>
<td>-31</td>
<td>-24</td>
<td>-17</td>
<td>-11</td>
<td>-4</td>
<td>3</td>
</tr>
<tr>
<td>30</td>
<td>-67</td>
<td>-60</td>
<td>-53</td>
<td>-46</td>
<td>-39</td>
<td>-33</td>
<td>-26</td>
<td>-19</td>
<td>-12</td>
<td>-5</td>
<td>1</td>
</tr>
<tr>
<td>35</td>
<td>-69</td>
<td>-62</td>
<td>-55</td>
<td>-48</td>
<td>-41</td>
<td>-34</td>
<td>-27</td>
<td>-21</td>
<td>-14</td>
<td>-7</td>
<td>0</td>
</tr>
<tr>
<td>40</td>
<td>-71</td>
<td>-64</td>
<td>-57</td>
<td>-50</td>
<td>-43</td>
<td>-36</td>
<td>-29</td>
<td>-22</td>
<td>-15</td>
<td>-8</td>
<td>-1</td>
</tr>
</tbody>
</table>

Table 10.2: Wind chill as a function of wind speed and temperature.

(a) Estimate the partial derivatives \( w_T(20, -15) \), \( w_T(20, -10) \), and \( w_T(20, -5) \). Use these results to estimate the second-order partial \( w_{TT}(20, -10) \).
(b) In a similar way, estimate the second-order partial \( w_{vv}(20, -10) \).

(c) Estimate the partial derivatives \( w_T(20, -10) \), \( w_T(25, -10) \), and \( w_T(15, -10) \), and use your results to estimate the partial \( w_{Tv}(20, -10) \).

(d) In a similar way, estimate the partial derivative \( w_{vT}(20, -10) \).

(e) Write several sentences that explain what the values \( w_{TT}(20, -10) \), \( w_{vv}(20, -10) \), and \( w_{Tv}(20, -10) \) indicate regarding the behavior of \( w(v, T) \).

As we have found in Activities 10.9 and 10.10, we may think of \( f_{xy} \) as measuring the “twist” of the graph as we increase \( y \) along a particular trace where \( x \) is held constant. In the same way, \( f_{yx} \) measures how the graph twists as we increase \( x \). If we remember that Clairaut’s theorem tells us that \( f_{xy} = f_{yx} \), we see that the amount of twisting is the same in both directions. This twisting is perhaps more easily seen in Figure 10.22, which shows the graph of \( f(x, y) = -xy \), for which \( f_{xy} = -1 \).

Figure 10.22: The graph of \( f(x, y) = -xy \).

Summary

- There are four second-order partial derivatives of a function \( f \) of two independent variables \( x \) and \( y \):
  
  \[
  f_{xx} = (f_x)_x, \quad f_{xy} = (f_x)_y, \quad f_{yx} = (f_y)_x, \quad \text{and} \quad f_{yy} = (f_y)_y.
  \]

- The unmixed second-order partial derivatives, \( f_{xx} \) and \( f_{yy} \), tell us about the concavity of the traces. The mixed second-order partial derivatives, \( f_{xy} \) and \( f_{yx} \), tell us how the graph of \( f \) twists.

Exercises

1. Shown in Figure 10.23 is a contour plot of a function \( f \) with the values of \( f \) labeled on the contours. The point \((2, 1)\) is highlighted in red.
(a) Estimate the partial derivatives $f_x(2,1)$ and $f_y(2,1)$.

(b) Determine whether the second-order partial derivative $f_{xx}(2,1)$ is positive or negative, and explain your thinking.

(c) Determine whether the second-order partial derivative $f_{yy}(2,1)$ is positive or negative, and explain your thinking.

(d) Determine whether the second-order partial derivative $f_{xy}(2,1)$ is positive or negative, and explain your thinking.

(e) Determine whether the second-order partial derivative $f_{yx}(2,1)$ is positive or negative, and explain your thinking.

(f) Consider a function $g$ of the variables $x$ and $y$ for which $g_x(2,2) > 0$ and $g_{xx}(2,2) < 0$. Sketch possible behavior of some contours around $(2,2)$ on the left axes in Figure 10.24.

(g) Consider a function $h$ of the variables $x$ and $y$ for which $h_x(2,2) > 0$ and $h_{xy}(2,2) < 0$. 

Figure 10.23: A contour plot of $f(x,y)$.

Figure 10.24: Plots for contours of $g$ and $h$. 

Sketch possible behavior of some contour lines around \((2, 2)\) on the right axes in Figure 10.24.

2. The Heat Index, \(I\), (measured in apparent degrees F) is a function of the actual temperature \(T\) outside (in degrees F) and the relative humidity \(H\) (measured as a percentage). A portion of the table which gives values for this function, \(I(T, H)\), is reproduced below:

<table>
<thead>
<tr>
<th>(T) ↓</th>
<th>(H \rightarrow)</th>
<th>70</th>
<th>75</th>
<th>80</th>
<th>85</th>
</tr>
</thead>
<tbody>
<tr>
<td>90</td>
<td>106</td>
<td>109</td>
<td>112</td>
<td>115</td>
<td></td>
</tr>
<tr>
<td>92</td>
<td>112</td>
<td>115</td>
<td>119</td>
<td>123</td>
<td></td>
</tr>
<tr>
<td>94</td>
<td>118</td>
<td>122</td>
<td>127</td>
<td>132</td>
<td></td>
</tr>
<tr>
<td>96</td>
<td>125</td>
<td>130</td>
<td>135</td>
<td>141</td>
<td></td>
</tr>
</tbody>
</table>

(a) State the limit definition of the value \(I_{TT}(94, 75)\). Then, estimate \(I_{TT}(94, 75)\), and write one complete sentence that carefully explains the meaning of this value, including units.

(b) State the limit definition of the value \(I_{HH}(94, 75)\). Then, estimate \(I_{HH}(94, 75)\), and write one complete sentence that carefully explains the meaning of this value, including units.

(c) Finally, do likewise to estimate \(I_{HT}(94, 75)\), and write a sentence to explain the meaning of the value you found.

3. The temperature on a heated metal plate positioned in the first quadrant of the \(x\)-\(y\) plane is given by

\[
C(x, y) = 25e^{-(x-1)^2-(y-1)^3}.
\]

Assume that temperature is measured in degrees Celsius and that \(x\) and \(y\) are each measured in inches.

(a) Determine \(C_{xx}(x, y)\) and \(C_{yy}(x, y)\). Do not do any additional work to algebraically simplify your results.

(b) Calculate \(C_{xx}(1.1, 1.2)\). Suppose that an ant is walking past the point \((1.1, 1.2)\) along the line \(y = 1.2\). Write a sentence to explain the meaning of the value of \(C_{xx}(1.1, 1.2)\), including units.

(c) Calculate \(C_{yy}(1.1, 1.2)\). Suppose instead that an ant is walking past the point \((1.1, 1.2)\) along the line \(x = 1.1\). Write a sentence to explain the meaning of the value of \(C_{yy}(1.1, 1.2)\), including units.

(d) Determine \(C_{xy}(x, y)\) and hence compute \(C_{xy}(1.1, 1.2)\). What is the meaning of this value? Explain, in terms of an ant walking on the heated metal plate.

4. Let \(f(x, y) = 8 - x^2 - y^2\) and \(g(x, y) = 8 - x^2 + 4xy - y^2\).

(a) Determine \(f_x, f_y, f_{xx}, f_{yy}, f_{xy}, f_{yx}\), and \(f_{yx}\).

(b) Evaluate each of the partial derivatives in (a) at the point \((0, 0)\).
(c) What do the values in (b) suggest about the behavior of $f$ near $(0,0)$? Plot a graph of $f$ and compare what you see visually to what the values suggest.

(d) Determine $g_x, g_y, g_{xx}, g_{yy}, g_{xy},$ and $g_{yx}$.

(e) Evaluate each of the partial derivatives in (d) at the point $(0,0)$.

(f) What do the values in (e) suggest about the behavior of $g$ near $(0,0)$? Plot a graph of $g$ and compare what you see visually to what the values suggest.

(g) What do the functions $f$ and $g$ have in common at $(0,0)$? What is different? What do your observations tell you regarding the importance of a certain second-order partial derivative?

5. Let $f(x, y) = \frac{1}{2}xy^2$ represent the kinetic energy in Joules of an object of mass $x$ in kilograms with velocity $y$ in meters per second. Let $(a,b)$ be the point $(4,5)$ in the domain of $f$.

(a) Calculate $\frac{\partial^2 f}{\partial x^2}$ at the point $(a,b)$. Then explain as best you can what this second order partial derivative tells us about kinetic energy.

(b) Calculate $\frac{\partial^2 f}{\partial y^2}$ at the point $(a,b)$. Then explain as best you can what this second order partial derivative tells us about kinetic energy.

(c) Calculate $\frac{\partial^2 f}{\partial y \partial x}$ at the point $(a,b)$. Then explain as best you can what this second order partial derivative tells us about kinetic energy.

(d) Calculate $\frac{\partial^2 f}{\partial x \partial y}$ at the point $(a,b)$. Then explain as best you can what this second order partial derivative tells us about kinetic energy.
10.4 Linearization: Tangent Planes and Differentials

Motivating Questions

In this section, we strive to understand the ideas generated by the following important questions:

- What does it mean for a function of two variables to be locally linear at a point?
- How do we find the equation of the plane tangent to a locally linear function at a point?
- What does it mean to say that a multivariable function is differentiable?
- What is the differential of a multivariable function of two variables and what are its uses?

Introduction

One of the central concepts in single variable calculus is that the graph of a differentiable function, when viewed on a very small scale, looks like a line. We call this line the tangent line and measure its slope with the derivative. In this section, we will extend this concept to functions of several variables.

Let’s see what happens when we look at the graph of a two-variable function on a small scale. To begin, let’s consider the function $f$ defined by

$$f(x, y) = 6 - \frac{x^2}{2} - y^2,$$

whose graph is shown in Figure 10.25.

![Figure 10.25: The graph of $f(x, y) = 6 - \frac{x^2}{2} - y^2$.](image)

We choose to study the behavior of this function near the point $(x_0, y_0) = (1, 1)$. In particular, we wish to view the graph on an increasingly small scale around this point, as shown in the two plots in Figure 10.26.
Just as the graph of a differentiable single-variable function looks like a line when viewed on a small scale, we see that the graph of this particular two-variable function looks like a plane, as seen in Figure 10.27. In the following preview activity, we explore how to find the equation of this plane.

In what follows, we will also use the important fact\(^1\) that the plane passing through \((x_0, y_0, z_0)\) may be expressed in the form \(z = z_0 + a(x - x_0) + b(y - y_0)\), where \(a\) and \(b\) are constants.

\(^1\)As we saw in Section 9.5, the equation of a plane passing through the point \((x_0, y_0, z_0)\) may be written in the form \(Ax + By + Cz + D = 0\). If the plane is not vertical, then \(C \neq 0\), and we can rearrange this and hence write \(C(z - z_0) = -A(x - x_0) - B(y - y_0)\) and thus

\[
  z = z_0 - \frac{A}{C}(x - x_0) - \frac{B}{C}(y - y_0)
  = z_0 + a(x - x_0) + b(y - y_0)
\]

where \(a = -A/C\) and \(b = -B/C\), respectively.
Preview Activity 10.4. Let \( f(x, y) = 6 - \frac{x^2}{2} - y^2 \), and let \((x_0, y_0) = (1, 1)\).

(a) Evaluate \( f(x, y) = 6 - \frac{x^2}{2} - y^2 \) and its partial derivatives at \((x_0, y_0)\); that is, find \( f(1, 1) \), \( f_x(1, 1) \), and \( f_y(1, 1) \).

(b) We know one point on the tangent plane; namely, the \( z \)-value of the tangent plane agrees with the \( z \)-value on the graph of \( f(x, y) = 6 - \frac{x^2}{2} - y^2 \) at the point \((x_0, y_0)\). In other words, both the tangent plane and the graph of the function \( f \) contain the point \((x_0, y_0, z_0)\). Use this observation to determine \( z_0 \) in the expression \( z = z_0 + a(x - x_0) + b(y - y_0) \).

(c) Sketch the traces of \( f(x, y) = 6 - \frac{x^2}{2} - y^2 \) for \( y = y_0 = 1 \) and \( x = x_0 = 1 \) below in Figure 10.28.

(d) Determine the equation of the tangent line of the trace that you sketched in the previous part with \( y = 1 \) (in the \( x \) direction) at the point \( x_0 = 1 \).

Figure 10.28: The traces of \( f(x, y) \) with \( y = y_0 = 1 \) and \( x = x_0 = 1 \).

Figure 10.29: The traces of \( f(x, y) \) and the tangent plane.
(e) Figure 10.29 shows the traces of the function and the traces of the tangent plane. Explain how the tangent line of the trace of \( f \), whose equation you found in the last part of this activity, is related to the tangent plane. How does this observation help you determine the constant \( a \) in the equation for the tangent plane \( z = z_0 + a(x - x_0) + b(y - y_0) \)? (Hint: How do you think \( f_x(x_0, y_0) \) should be related to \( z_x(x_0, y_0) \)?)

(f) In a similar way to what you did in (d), determine the equation of the tangent line of the trace with \( x = 1 \) at the point \( y_0 = 1 \). Explain how this tangent line is related to the tangent plane, and use this observation to determine the constant \( b \) in the equation for the tangent plane \( z = z_0 + a(x - x_0) + b(y - y_0) \). (Hint: How do you think \( f_y(x_0, y_0) \) should be related to \( z_y(x_0, y_0) \)?)

(g) Finally, write the equation \( z = z_0 + a(x - x_0) + b(y - y_0) \) of the tangent plane to the graph of \( f(x, y) = 6 - x^2/2 - y^2 \) at the point \( (x_0, y_0) = (1, 1) \).

The tangent plane

Before stating the formula for the equation of the tangent plane at a point for a general function \( f = f(x, y) \), we need to discuss a mild technical condition. As we have noted, when we look at the graph of a single-variable function on a small scale near a point \( x_0 \), we expect to see a line; in this case, we say that \( f \) is locally linear near \( x_0 \) since the graph looks like a linear function locally around \( x_0 \). Of course, there are functions, such as the absolute value function given by \( f(x) = |x| \), that are not locally linear at every point. In single-variable calculus, we learn that if the derivative of a function exists at a point, then the function is guaranteed to be locally linear there.

In a similar way, we say that a two-variable function \( f \) is locally linear near \( (x_0, y_0) \) provided that the graph of \( f \) looks like a plane when viewed on a small scale near \( (x_0, y_0) \). There are, of course, functions that are not locally linear at some points \( (x_0, y_0) \). However, it turns out that if the first-order partial derivatives, \( f_x \) and \( f_y \), are continuous near \( (x_0, y_0) \), then \( f \) is locally linear at \( (x_0, y_0) \) and the graph looks like a plane, which we call the tangent plane, when viewed on a small scale. Moreover, when a function is locally linear at a point, we will also say it is differentiable at that point.

If \( f \) is a function of the independent variables \( x \) and \( y \) and both \( f_x \) and \( f_y \) exist and are continuous in an open disk containing the point \( (x_0, y_0) \), then \( f \) is differentiable at \( (x_0, y_0) \).

So, whenever a function \( z = f(x, y) \) is differentiable at a point \( (x_0, y_0) \), it follows that the function has a tangent plane at \( (x_0, y_0) \). Viewed up close, the tangent plane and the function are then virtually indistinguishable. In addition, as in Preview Activity 10.4, we find the following
If $f(x, y)$ has continuous first-order partial derivatives, then the equation of the plane tangent to the graph of $f$ at the point $(x_0, y_0, f(x_0, y_0))$ is

$$z = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0).$$  \hspace{1cm} (10.1)

Finally, one important note about the form of the equation for the tangent plane, $z = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0)$. Say, for example, that we have the particular tangent plane $z = 7 - 2(x - 3) + 4(y + 1)$. Observe that we can immediately read from this form that $f_x(3, -1) = -2$ and $f_y(3, -1) = 4$; furthermore, $f_x(3, -1) = -2$ is the slope of the trace to both $f$ and the tangent plane in the $x$-direction at $(-3, 1)$. In the same way, $f_y(3, -1) = 4$ is the slope of the trace of both $f$ and the tangent plane in the $y$-direction at $(3, -1)$.

Activity 10.11.

Find the equation of the tangent plane to $f(x, y) = x^2y$ at the point $(1, 2)$.

\hspace{1cm} ◀

Linearization

In single variable calculus, an important use of the tangent line is to approximate the value of a differentiable function. Near the point $x_0$, the tangent line to the graph of $f$ at $x_0$ is close to the graph of $f$ near $x_0$, as shown in Figure 10.30.

![Figure 10.30: The linearization of the single-variable function $f(x)$.](image)

In this single-variable setting, we let $L$ denote the function whose graph is the tangent line, and thus

$$L(x) = f(x_0) + f'(x_0)(x - x_0)$$
Furthermore, observe that \( f(x) \approx L(x) \) near \( x_0 \). We call \( L \) the linearization of \( f \).

In the same way, the tangent plane to the graph of a differentiable function \( z = f(x,y) \) at a point \((x_0, y_0)\) provides a good approximation of \( f(x,y) \) near \((x_0, y_0)\). Here, we define the linearization, \( L \), to be the two-variable function whose graph is the tangent plane, and thus

\[
L(x, y) = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0).
\]

Finally, note that \( f(x, y) \approx L(x, y) \) for points near \((x_0, y_0)\). This is illustrated in Figure 10.31.

![Figure 10.31: The linearization of \( f(x, y) \).](image)

**Activity 10.12.**

In what follows, we find the linearization of several different functions that are given in algebraic, tabular, or graphical form.

(a) Find the linearization \( L(x, y) \) for the function \( g \) defined by

\[
g(x, y) = \frac{x}{x^2 + y^2}
\]

at the point \((1, 2)\). Then use the linearization to estimate the value of \( g(0.8, 2.3) \).

(b) Table 10.3 provides a collection of values of the wind chill \( w(v, T) \), in degrees Fahrenheit, as a function of wind speed, in miles per hour, and temperature, also in degrees Fahrenheit.

Use the data to first estimate the appropriate partial derivatives, and then find the linearization \( L(v, T) \) at the point \((25, -10)\). Finally, use the linearization to estimate \( w(25, -12) \), \( w(23, -10) \), and \( w(23, -12) \).

(c) Figure 10.32 gives a contour plot of a differentiable function \( f \).

After estimating appropriate partial derivatives, determine the linearization \( L(x, y) \) at the point \((2, 1)\), and use it to estimate \( f(2.2, 1) \), \( f(2, 0.8) \), and \( f(2.2, 0.8) \).
### Differentials

As we have seen, the linearization $L(x, y)$ enables us to estimate the value of $f(x, y)$ for points $(x, y)$ near the base point $(x_0, y_0)$. Sometimes, however, we are more interested in the change in $f$ as we move from the base point $(x_0, y_0)$ to another point $(x, y)$.

Figure 10.33 illustrates this situation. Suppose we are at the point $(x_0, y_0)$, and we know the value $f(x_0, y_0)$ of $f$ at $(x_0, y_0)$. If we consider the displacement $(\Delta x, \Delta y)$ to a new point $(x, y) = (x_0 + \Delta x, y_0 + \Delta y)$, we would like to know how much the function has changed. We denote this change by $\Delta f$, where

$$\Delta f = f(x, y) - f(x_0, y_0).$$

A simple way to estimate the change $\Delta f$ is to approximate it by $df$, which represents the change

<table>
<thead>
<tr>
<th>$v \backslash T$</th>
<th>-30</th>
<th>-25</th>
<th>-20</th>
<th>-15</th>
<th>-10</th>
<th>-5</th>
<th>0</th>
<th>5</th>
<th>10</th>
<th>15</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>-46</td>
<td>-40</td>
<td>-34</td>
<td>-28</td>
<td>-22</td>
<td>-16</td>
<td>-11</td>
<td>-5</td>
<td>1</td>
<td>7</td>
<td>13</td>
</tr>
<tr>
<td>10</td>
<td>-53</td>
<td>-47</td>
<td>-41</td>
<td>-35</td>
<td>-28</td>
<td>-22</td>
<td>-16</td>
<td>-10</td>
<td>4</td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td>25</td>
<td>-64</td>
<td>-58</td>
<td>-51</td>
<td>-44</td>
<td>-37</td>
<td>-31</td>
<td>-24</td>
<td>-17</td>
<td>11</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>30</td>
<td>-67</td>
<td>-60</td>
<td>-53</td>
<td>-46</td>
<td>-39</td>
<td>-33</td>
<td>-26</td>
<td>-19</td>
<td>-12</td>
<td>-5</td>
<td>1</td>
</tr>
<tr>
<td>35</td>
<td>-69</td>
<td>-62</td>
<td>-55</td>
<td>-48</td>
<td>-41</td>
<td>-34</td>
<td>-27</td>
<td>-21</td>
<td>-14</td>
<td>-7</td>
<td>0</td>
</tr>
<tr>
<td>40</td>
<td>-71</td>
<td>-64</td>
<td>-57</td>
<td>-50</td>
<td>-43</td>
<td>-36</td>
<td>-29</td>
<td>-22</td>
<td>-15</td>
<td>-8</td>
<td>-1</td>
</tr>
</tbody>
</table>

Table 10.3: Wind chill as a function of wind speed and temperature.
in the linearization $L(x, y)$ as we move from $(x_0, y_0)$ to $(x, y)$. This gives

$$\Delta f \approx df = L(x, y) - f(x_0, y_0)$$
$$= [f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0)] - f(x_0, y_0)$$
$$= f_x(x_0, y_0)\Delta x + f_y(x_0, y_0)\Delta y.$$  

For consistency, we will denote the change in the independent variables as $dx = \Delta x$ and $dy = \Delta y$.

$$\Delta f \approx df = f_x(x_0, y_0)\ dx + f_y(x_0, y_0)\ dy. \quad (10.2)$$

Expressed equivalently in Leibniz notation, we have

$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy. \quad (10.3)$$

We call the quantities $dx$, $dy$, and $df$ differentials, and we think of them as measuring small changes in the quantities $x$, $y$, and $f$. Equations (10.2) and (10.3) express the relationship between these changes. Equation (10.3) resembles an important idea from single-variable calculus: when $y$ depends on $x$, it follows in the notation of differentials that

$$dy = y'\ dx = \frac{dy}{dx}\ dx.$$  

We will illustrate the use of differentials with an example. Suppose we have a machine that manufactures rectangles of width $x = 20$ cm and height $y = 10$ cm. However, the machine isn’t perfect, and therefore the width could be off by $dx = \Delta x = 0.2$ cm and the height could be off by $dy = \Delta y = 0.4$ cm.

The area of the rectangle is

$$A(x, y) = xy,$$

so that the area of a perfectly manufactured rectangle is $A(20, 10) = 200$ square centimeters. Since the machine isn’t perfect, we would like to know how much the area of a given manufactured
rectangle could differ from the perfect rectangle. We will estimate the uncertainty in the area using (10.2), and find that
\[ \Delta A \approx dA = A_x(20,10) \, dx + A_y(20,10) \, dy. \]
Since \( A_x = y \) and \( A_y = x \), we have
\[ \Delta A \approx dA = 10 \, dx + 20 \, dy = 10 \cdot 0.2 + 20 \cdot 0.4 = 10. \]
That is, we estimate that the area in our rectangles could be off by as much as 10 square centimeters.

**Activity 10.13.**

The questions in this activity explore the differential in several different contexts.

(a) Suppose that the elevation of a landscape is given by the function \( h \), where we additionally know that \( h(3,1) = 4.35 \), \( h_x(3,1) = 0.27 \), and \( h_y(3,1) = -0.19 \). Assume that \( x \) and \( y \) are measured in miles in the easterly and northerly directions, respectively, from some base point \((0,0)\).
Your GPS device says that you are currently at the point \((3,1)\). However, you know that the coordinates are only accurate to within 0.2 units; that is, \( dx = \Delta x = 0.2 \) and \( dy = \Delta y = 0.2 \). Estimate the uncertainty in your elevation using differentials.

(b) The pressure, volume, and temperature of an ideal gas are related by the equation
\[ P = P(T,V) = 8.31 T/V, \]
where \( P \) is measured in kilopascals, \( V \) in liters, and \( T \) in kelvin. Find the pressure when the volume is 12 liters and the temperature is 310 K. Use differentials to estimate the change in the pressure when the volume increases to 12.3 liters and the temperature decreases to 305 K.

(c) Refer to Table 10.3, the table of values of the wind chill \( w(v,T) \), in degrees Fahrenheit, as a function of temperature, also in degrees Fahrenheit, and wind speed, in miles per hour.
Suppose your anemometer says the wind is blowing at 25 miles per hour and your thermometer shows a reading of \(-15^\circ \) degrees. However, you know your thermometer is only accurate to within \( 2^\circ \) degrees and your anemometer is only accurate to within 3 miles per hour. What is the wind chill based on your measurements? Estimate the uncertainty in your measurement of the wind chill.

**Summary**

- A function \( f \) of two independent variables is locally linear at a point \((x_0,y_0)\) if the graph of \( f \) looks like a plane as we zoom in on the graph around the point \((x_0,y_0)\). In this case, the equation of the tangent plane is given by
\[ z = f(x_0,y_0) + f_x(x_0,y_0)(x-x_0) + f_y(x_0,y_0)(y-y_0). \]
The tangent plane \( L(x, y) = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0), \) when considered as a function, is called the linearization of a differentiable function \( f \) at \((x_0, y_0)\) and may be used to estimate values of \( f(x, y) \); that is, \( f(x, y) \approx L(x, y) \) for points \((x, y)\) near \((x_0, y_0)\).

A function \( f \) of two independent variables is differentiable at \((x_0, y_0)\) provided that both \( f_x \) and \( f_y \) exist and are continuous in an open disk containing the point \((x_0, y_0)\).

The differential \( df \) of a function \( f = f(x, y) \) is related to the differentials \( dx \) and \( dy \) by
\[
df = f_x(x_0, y_0)\,dx + f_y(x_0, y_0)\,dy.
\]
We can use this relationship to approximate small changes in \( f \) that result from small changes in \( x \) and \( y \).

**Exercises**

1. Let \( f \) be the function defined by \( f(x, y) = x^{1/3}y^{1/3} \), whose graph is shown in Figure 10.34.

   ![Figure 10.34: The surface for \( f(x, y) = x^{1/3}y^{1/3} \).](image)

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

   What does this limit tell us about \( f_x(0, 0) \)?

   (b) Note that \( f(x, y) = f(y, x) \), and this symmetry implies that \( f_x(0, 0) = f_y(0, 0) \). So both partial derivatives of \( f \) exist at \((0, 0)\). A picture of the surface defined by \( f \) near \((0, 0)\) is shown in Figure 10.34. Based on this picture, do you think \( f \) is locally linear at \((0, 0)\)? Why?

   (c) Show that the curve where \( x = y \) on the surface defined by \( f \) is not differentiable at 0. What does this tell us about the local linearity of \( f \) at \((0, 0)\)?
(d) Is the function \( f \) defined by \( f(x, y) = \frac{x^2}{y^2 + 1} \) locally linear at \((0, 0)\)? Why or why not?

2. Let \( g \) be a function that is differentiable at \((-2, 5)\) and suppose that its tangent plane at this point is given by \( z = -7 + 4(x + 2) - 3(y - 5) \).

(a) Determine the values of \( g(-2, 5) \), \( g_x(-2, 5) \), and \( g_y(-2, 5) \). Write one sentence to explain your thinking.

(b) Estimate the value of \( g(-1.8, 4.7) \). Clearly show your work and thinking.

(c) Given changes of \( dx = -0.34 \) and \( dy = 0.21 \), estimate the corresponding change in \( g \) that is given by its differential, \( dg \).

(d) Suppose that another function \( h \) is also differentiable at \((-2, 5)\), but that its tangent plane at \((-2, 5)\) is given by \( 3x + 2y - 4z = 9 \). Determine the values of \( h(-2, 5) \), \( h_x(-2, 5) \), and \( h_y(-2, 5) \), and then estimate the value of \( h(-1.8, 4.7) \). Clearly show your work and thinking.

3. In the following questions, we determine and apply the linearization for several different functions.

(a) Find the linearization \( L(x, y) \) for the function \( f \) defined by \( f(x, y) = \cos(x)(2e^{2y} + e^{-2y}) \) at the point \((x_0, y_0) = (0, 0)\). Hence use the linearization to estimate the value of \( f(0.1, 0.2) \). Compare your estimate to the actual value of \( f(0.1, 0.2) \).

(b) The Heat Index, \( I \), (measured in apparent degrees F) is a function of the actual temperature \( T \) outside (in degrees F) and the relative humidity \( H \) (measured as a percentage). A portion of the table which gives values for this function, \( I = I(T, H) \), is provided below:

<table>
<thead>
<tr>
<th>( T ) ( \downarrow ) ( H ) ( \rightarrow )</th>
<th>70</th>
<th>75</th>
<th>80</th>
<th>85</th>
</tr>
</thead>
<tbody>
<tr>
<td>90</td>
<td>106</td>
<td>109</td>
<td>112</td>
<td>115</td>
</tr>
<tr>
<td>92</td>
<td>112</td>
<td>115</td>
<td>119</td>
<td>123</td>
</tr>
<tr>
<td>94</td>
<td>118</td>
<td>122</td>
<td>127</td>
<td>132</td>
</tr>
<tr>
<td>96</td>
<td>125</td>
<td>130</td>
<td>135</td>
<td>141</td>
</tr>
</tbody>
</table>

Suppose you are given that \( I_T(94, 75) = 3.75 \) and \( I_H(94, 75) = 0.9 \). Use this given information and one other value from the table to estimate the value of \( I(93.1, 77) \) using the linearization at \((94, 75)\). Using proper terminology and notation, explain your work and thinking.

(c) Just as we can find a local linearization for a differentiable function of two variables, we can do so for functions of three or more variables. By extending the concept of the local linearization from two to three variables, find the linearization of the function \( h(x, y, z) = e^{2x}(y + z^2) \) at the point \((x_0, y_0, z_0) = (0, 1, -2)\). Then, use the linearization to estimate the value of \( h(-0.1, 0.9, -1.8) \).

4. In the following questions, we investigate two different applied settings using the differential.
(a) Let \( f \) represent the vertical displacement in centimeters from the rest position of a string (like a guitar string) as a function of the distance \( x \) in centimeters from the fixed left end of the string and \( y \) the time in seconds after the string has been plucked.\(^2\) A simple model for \( f \) could be

\[
    f(x, y) = \cos(x) \sin(2y).
\]

Use the differential to approximate how much more this vibrating string is vertically displaced from its position at \((a, b) = \left(\frac{\pi}{4}, \frac{\pi}{3}\right)\) if we decrease \( a \) by 0.01 cm and increase the time by 0.1 seconds. Compare to the value of \( f \) at the point \( \left(\frac{\pi}{4} - 0.01, \frac{\pi}{3} + 0.1\right) \).

(b) Resistors used in electrical circuits have colored bands painted on them to indicate the amount of resistance and the possible error in the resistance. When three resistors, whose resistances are \( R_1, R_2, \) and \( R_3 \), are connected in parallel, the total resistance \( R \) is given by

\[
    \frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3}.
\]

Suppose that the resistances are \( R_1 = 25\Omega \), \( R_2 = 40\Omega \), and \( R_3 = 50\Omega \). Find the total resistance \( R \).

If you know each of \( R_1, R_2, \) and \( R_3 \) with a possible error of 0.5%, estimate the maximum error in your calculation of \( R \).

\(^2\)An interesting video of this can be seen at https://www.youtube.com/watch?v=TKF6nFzpHBUA.
10.5 The Chain Rule

Motivating Questions

In this section, we strive to understand the ideas generated by the following important questions:

- What is the Chain Rule and how do we use it to find a derivative?
- How can we use a tree diagram to guide us in applying the Chain Rule?

Introduction

In single-variable calculus, we encountered situations in which some quantity \( z \) depends on \( y \) and, in turn, \( y \) depends on \( x \). A change in \( x \) produces a change in \( y \), which consequently produces a change in \( z \). Using the language of differentials that we saw in the previous section, these changes are naturally related by

\[
\frac{dz}{dy} \ dy \quad \text{and} \quad \frac{dy}{dx} \ dx.
\]

In terms of instantaneous rates of change, we then have

\[
\frac{dz}{dx} = \frac{dz}{dy} \frac{dy}{dx} \ dx = \frac{dz}{dx} \ dx
\]

and thus

\[
\frac{dz}{dx} = \frac{dz}{dy} \frac{dy}{dx}.
\]

This most recent equation we call the Chain Rule.

In the case of a function \( f \) of two variables where \( z = f(x, y) \), it might be that both \( x \) and \( y \) depend on another variable \( t \). A change in \( t \) then produces changes in both \( x \) and \( y \), which then cause \( z \) to change. In this section we will see how to find the change in \( z \) that is caused by a change in \( t \), leading us to multivariable versions of the Chain Rule involving both regular and partial derivatives.

Preview Activity 10.5. Suppose you are driving around in the \( x\)-\( y \) plane in such a way that your position \( r(t) \) at time \( t \) is given by function

\[
r(t) = (x(t), y(t)) = (2 - t^2, t^3 + 1).
\]

The path taken is shown on the left of Figure 10.35.

Suppose, furthermore, that the temperature at a point in the plane is given by

\[
T(x, y) = 10 - \frac{1}{2} x^2 - \frac{1}{5} y^2,
\]

and note that the surface generated by \( T \) is shown on the right of Figure 10.35. Therefore, as time passes, your position \( (x(t), y(t)) \) changes, and, as your position changes, the temperature \( T(x, y) \) also changes.
(a) The position function $r$ provides a parameterization $x = x(t)$ and $y = y(t)$ of the position at time $t$. By substituting $x(t)$ for $x$ and $y(t)$ for $y$ in the formula for $T$, we can write $T = T(x(t), y(t))$ as a function of $t$. Make these substitutions to write $T$ as a function of $t$ and then use the Chain Rule from single variable calculus to find $\frac{dT}{dt}$. (Do not do any algebra to simplify the derivative, either before taking the derivative, nor after.)

(b) Now we want to understand how the result from part (a) can be obtained from $T$ as a multivariable function. Recall from the previous section that small changes in $x$ and $y$ produce a change in $T$ that is approximated by

$$\Delta T \approx T_x \Delta x + T_y \Delta y.$$ 

The Chain Rule tells us about the instantaneous rate of change of $T$, and this can be found as

$$\lim_{\Delta t \to 0} \frac{\Delta T}{\Delta t} = \lim_{\Delta t \to 0} \frac{T_x \Delta x + T_y \Delta y}{\Delta t}. \quad (10.4)$$

Use equation (10.4) to explain why the instantaneous rate of change of $T$ that results from a change in $t$ is

$$\frac{dT}{dt} = \frac{\partial T}{\partial x} \frac{dx}{dt} + \frac{\partial T}{\partial y} \frac{dy}{dt}. \quad (10.5)$$

(c) Using the original formulas for $T$, $x$, and $y$ in the problem statement, calculate all of the derivatives in Equation (10.5) (with $T_x$ and $T_y$ in terms of $x$ and $y$, and $x'$ and $y'$ in terms of $t$), and hence write the right-hand side of Equation (10.5) in terms of $x$, $y$, and $t$.

(d) Compare the results of parts (a) and (c). Write a couple of sentences that identify specifically how each term in (c) relates to a corresponding terms in (a). This connection between parts (a) and (c) provides a multivariable version of the Chain Rule.
The Chain Rule

As Preview Activity 10.3 suggests, the following version of the Chain Rule holds in general.

Let \( z = f(x, y) \), where \( f \) is a differentiable function of the independent variables \( x \) and \( y \), and let \( x \) and \( y \) each be differentiable functions of an independent variable \( t \). Then

\[
\frac{dz}{dt} = \frac{\partial z}{\partial x} \frac{dx}{dt} + \frac{\partial z}{\partial y} \frac{dy}{dt}.
\]

(10.6)

It is important to note the differences among the derivatives in (10.6). Since \( z \) is a function of the two variables \( x \) and \( y \), the derivatives in the Chain Rule for \( z \) with respect to \( x \) and \( y \) are partial derivatives. However, since \( x = x(t) \) and \( y = y(t) \) are functions of the single variable \( t \), their derivatives are the standard derivatives of functions of one variable. When we compose \( z \) with \( x(t) \) and \( y(t) \), we then have \( z \) as a function of the single variable \( t \), making the derivative of \( z \) with respect to \( t \) a standard derivative from single variable calculus as well.

To understand why this Chain Rule works in general, suppose that some quantity \( z \) depends on \( x \) and \( y \) so that

\[
\frac{dz}{dt} = \frac{\partial z}{\partial x} dx + \frac{\partial z}{\partial y} dy.
\]

(10.7)

Next, suppose that \( x \) and \( y \) each depend on another quantity \( t \), so that

\[
dx = \frac{dx}{dt} dt \quad \text{and} \quad dy = \frac{dy}{dt} dt.
\]

(10.8)

Combining Equations (10.7) and (10.8), we find that

\[
\frac{dz}{dt} = \frac{\partial z}{\partial x} \frac{dx}{dt} dt + \frac{\partial z}{\partial y} \frac{dy}{dt} dt = \frac{dz}{dt} dt,
\]

which is the Chain Rule in this particular context, as expressed in Equation 10.6.


In the following questions, we apply the recently-developed Chain Rule in several different contexts.

(a) Suppose that we have a function \( z \) defined by \( z(x, y) = x^2 + xy^3 \). In addition, suppose that \( x \) and \( y \) are restricted to points that move around the plane by following a circle of radius 2 centered at the origin that is parameterized by

\[
x(t) = 2 \cos(t), \quad \text{and} \quad y(t) = 2 \sin(t).
\]

Use the Chain Rule to find the resulting instantaneous rate of change \( \frac{dz}{dt} \).
(b) Suppose that the temperature on a metal plate is given by the function $T$ with

$$T(x, y) = 100 - (x^2 + 4y^2),$$

where the temperature is measured in degrees Fahrenheit and $x$ and $y$ are each measured in feet.

i. Find $T_x$ and $T_y$. What are the units on these partial derivatives?

ii. Suppose an ant is walking along the $x$-axis at the rate of 2 feet per minute toward the origin. When the ant is at the point $(2, 0)$, what is the instantaneous rate of change in the temperature $dT/dt$ that the ant experiences. Include units on your response.

iii. Suppose instead that the ant walks along an ellipse with $x = 6 \cos(t)$ and $y = 3 \sin(t)$, where $t$ is measured in minutes. Find $dT/dt$ at $t = \pi/6$, $t = \pi/4$, and $t = \pi/3$. What does this seem to tell you about the path along which the ant is walking?

(c) Suppose that you are walking along a surface whose elevation is given by a function $f$. Furthermore, suppose that if you consider how your location corresponds to points in the $x$-$y$ plane, you know that when you pass the point $(2, 1)$, your velocity vector is $v = \langle -1, 2 \rangle$. If some contours of $f$ are as shown in Figure 10.36, estimate the rate of change $df/dt$ when you pass through $(2, 1)$.

![Figure 10.36: Some contours of $f$.](image)

Tree Diagrams

Up to this point, we have applied the Chain Rule to situations where we have a function $z$ of variables $x$ and $y$, with both $x$ and $y$ depending on another single quantity $t$. We may apply the Chain Rule, however, when $x$ and $y$ each depend on more than one quantity, or when $z$ is a function of more than two variables. It can be challenging to keep track of all the dependencies.
among the variables, and thus a tree diagram can be a useful tool to organize our work. For example, suppose that \( z \) depends on \( x \) and \( y \), and \( x \) and \( y \) both depend on \( t \). We may represent these relationships using the tree diagram shown at left Figure 10.37. We place the dependent variable at the top of the tree and connect it to the variables on which it depends one level below. We then connect each of those variables to the variable on which each depends.

![Figure 10.37: A tree diagram illustrating dependencies.](image)

To represent the Chain Rule, we label every edge of the diagram with the appropriate derivative or partial derivative, as seen at right in Figure 10.37. To calculate an overall derivative according to the Chain Rule, we construct the product of the derivatives along all paths connecting the variables and then add all of these products. For example, the diagram at right in 10.37 illustrates the Chain Rule

\[
\frac{dz}{dt} = \frac{\partial z}{\partial x} \frac{dx}{dt} + \frac{\partial z}{\partial y} \frac{dy}{dt}.
\]

Activity 10.15.

(a) Figure 10.38 shows the tree diagram we construct when (a) \( z \) depends on \( w, x, \) and \( y \), (b) \( w, x, \) and \( y \) each depend on \( u \) and \( v \), and (c) \( u \) and \( v \) depend on \( t \).

i. Label the edges with the appropriate derivatives.

ii. Use the Chain Rule to write \( \frac{dz}{dt} \).

(b) Suppose that \( z = x^2 - 2xy^2 \) and that

\[
x = r \cos(\theta) \\
y = r \sin(\theta).
\]

i. Construct a tree diagram representing the dependencies of \( z \) on \( x \) and \( y \) and \( x \) and \( y \) on \( r \) and \( \theta \).

ii. Use the tree diagram to find \( \frac{\partial z}{\partial r} \).

iii. Now suppose that \( r = 3 \) and \( \theta = \pi/6 \). Find the values of \( x \) and \( y \) that correspond to these given values of \( r \) and \( \theta \), and then use the Chain Rule to find the value of the partial derivative \( \frac{\partial z}{\partial \theta} \) at \((3, \pi/6)\).
Summary

- The Chain Rule is a tool for differentiating a composite for functions. In its simplest form, it says that if \( f(x, y) \) is a function of two variables and \( x(t) \) and \( y(t) \) depend on \( t \), then

\[
\frac{df}{dt} = \frac{\partial f}{\partial x} \frac{dx}{dt} + \frac{\partial f}{\partial y} \frac{dy}{dt}.
\]

- A tree diagram can be used to represent the dependence of variables on other variables. By following the links in the tree diagram, we can form chains of partial derivatives or derivatives that can be combined to give a desired partial derivative.

Exercises

1. Find the indicated derivative. In each case, state the version of the Chain Rule that you are using.

   (a) \( \frac{df}{dt} \), if \( f(x, y) = 2x^2y, x = \cos(t), \) and \( y = \ln(t) \).

   (b) \( \frac{\partial f}{\partial w} \), if \( f(x, y) = 2x^2y, x = w + z^2, \) and \( y = \frac{2z+1}{w} \).

   (c) \( \frac{\partial f}{\partial v} \), if \( f(x, y, z) = 2x^2y + z^3, x = u - v + 2w, y = w2^v - u^3, \) and \( z = u^2 - v \).

2. Let \( z = u^2 - v^2 \) and suppose that

\[
\begin{align*}
  u &= e^x \cos(y) \\
  v &= e^x \sin(y)
\end{align*}
\]
(a) Find the values of \( u \) and \( v \) that correspond to \( x = 0 \) and \( y = 2\pi/3 \).

(b) Use the Chain Rule to find the general partial derivatives

\[
\frac{\partial z}{\partial x} \quad \text{and} \quad \frac{\partial z}{\partial y}
\]

and then determine both \( \frac{\partial z}{\partial x} \bigg|_{(0, \frac{2\pi}{3})} \) and \( \frac{\partial z}{\partial y} \bigg|_{(0, \frac{2\pi}{3})} \).

3. Suppose that \( T = x^2 + y^2 - 2z \) where

\[
x = \rho \sin(\phi) \cos(\theta) \\
y = \rho \sin(\phi) \sin(\theta) \\
z = \rho \cos(\phi)
\]

(a) Construct a tree diagram representing the dependencies among the variables.

(b) Apply the chain rule to find the partial derivatives

\[
\frac{\partial T}{\partial \rho}, \frac{\partial T}{\partial \phi}, \text{ and } \frac{\partial T}{\partial \theta}.
\]

4. Suppose that the temperature on a metal plate is given by the function \( T \) with

\[
T(x, y) = 100 - (x^2 + 4y^2),
\]

where the temperature is measured in degrees Fahrenheit and \( x \) and \( y \) are each measured in feet. Now suppose that an ant is walking on the metal plate in such a way that it walks in a straight line from the point \((1, 4)\) to the point \((5, 6)\).

(a) Find parametric equations \((x(t), y(t))\) for the ant’s coordinates as it walks the line from \((1, 4)\) to \((5, 6)\).

(b) What can you say about \( \frac{dx}{dt} \) and \( \frac{dy}{dt} \) for every value of \( t \)?

(c) Determine the instantaneous rate of change in temperature with respect to \( t \) that the ant is experiencing at the moment it is halfway from \((1, 4)\) to \((5, 6)\), using your parametric equations for \( x \) and \( y \). Include units on your answer.

5. There are several proposed formulas to approximate the surface area of the human body. One model\(^3\) uses the formula

\[
A(h, w) = 0.0072h^{0.725}w^{0.425},
\]

where \( A \) is the surface area in square meters, \( h \) is the height in centimeters, and \( w \) is the weight in kilograms.

Since a person’s height $h$ and weight $w$ change over time, $h$ and $w$ are functions of time $t$. Let us think about what is happening to a child whose height is 60 centimeters and weight is 9 kilograms. Suppose, furthermore, that $h$ is increasing at an instantaneous rate of 20 centimeters per year and $w$ is increasing at an instantaneous rate of 5 kg per year.

Determine the instantaneous rate at which the child’s surface area is changing at this point in time.

6. Let $z = f(x, y) = 50 - (x + 1)^2 - (y + 3)^2$ and $z = h(x, y) = 24 - 2x - 6y$.

Suppose a person is walking on the surface $z = f(x, y)$ in such a way that she walks the curve which is the intersection of $f$ and $h$.

(a) Show that $x(t) = 4 \cos(t)$ and $y(t) = 4 \sin(t)$ is a parameterization of the “shadow” in the $x$-$y$ plane of the curve that is the intersection of the graphs of $f$ and $h$.

(b) Use the parameterization from part (a) to find the instantaneous rate at which her height is changing with respect to time at the instant $t = 2\pi/3$.

7. The voltage $V$ (in volts) across a circuit is given by Ohm’s Law: $V = IR$, where $I$ is the current (in amps) in the circuit and $R$ is the resistance (in ohms). Suppose we connect two resistors with resistances $R_1$ and $R_2$ in parallel as shown in Figure 10.39. The total resistance $R$ in the circuit is then given by

$$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2}.$$ 

(a) Assume that the current, $I$, and the resistances, $R_1$ and $R_2$, are changing over time, $t$. Use the Chain Rule to write a formula for $\frac{dV}{dt}$.

(b) Suppose that, at some particular point in time, we measure the current to be 3 amps and that the current is increasing at $\frac{1}{10}$ amps per second, while resistance $R_1$ is 2 ohms and decreasing at the rate of 0.2 ohms per second and $R_2$ is 1 ohm and increasing at the rate of 0.5 ohms per second. At what rate is the voltage changing at this point in time?
10.6 Directional Derivatives and the Gradient

Motivating Questions

In this section, we strive to understand the ideas generated by the following important questions:

- The partial derivatives of a function $f$ tell us the rate of change of $f$ in the direction of the coordinate axes. How can we measure the rate of change of $f$ in other directions?
- What is the gradient of a function and what does it tell us?

Introduction

The partial derivatives of a function tell us the instantaneous rate at which the function changes as we hold all but one independent variable constant and allow the remaining independent variable to change. It is natural to wonder how we can measure the rate at which a function changes in directions other than parallel to a coordinate axes. In what follows, we investigate this question, and see how the rate of change in any given direction is connected to the rates of change given by the standard partial derivatives.

Preview Activity 10.6. Let’s consider the function $f$ defined by

$$f(x, y) = 30 - x^2 - \frac{1}{2}y^2,$$

and suppose that $f$ measures the temperature, in degrees Celsius, at a given point in the plane, where $x$ and $y$ are measured in feet. Assume that the positive $x$-axis points due east, while the positive $y$-axis points due north. A contour plot of $f$ is shown in Figure 10.40

![Contour Plot](image)

Figure 10.40: A contour plot of $f(x, y) = 30 - x^2 - \frac{1}{2}y^2$. 
(a) Suppose that a person is walking due east, and thus parallel to the \(x\)-axis. At what instantaneous rate is the temperature changing at the moment she passes the point \((2, 1)\)? What are the units on this rate of change?

(b) Next, determine the instantaneous rate of change of temperature at the point \((2, 1)\) if the person is instead walking due north. Again, include units on your result.

(c) Now, rather than walking due east or due north, let’s suppose that the person is walking with velocity given by the vector \(\mathbf{v} = \langle 3, 4 \rangle\), where time is measured in seconds. Note that the person’s speed is thus \(|\mathbf{v}| = 5\) feet per second.

Find parametric equations for the person’s path; that is, parameterize the line through \((2, 1)\) using the direction vector \(\mathbf{v} = \langle 3, 4 \rangle\). Let \(x(t)\) denote the \(x\)-coordinate of the line, and \(y(t)\) its \(y\)-coordinate.

(d) With the parameterization in (c), we can now view the temperature \(f\) as not only a function of \(x\) and \(y\), but also of time, \(t\). Hence, use the chain rule to determine the value of \(\frac{df}{dt} \bigg|_{t=0}\). What are the units on your answer? What is the practical meaning of this result?

**Directional Derivatives**

Given a function \(z = f(x, y)\), the partial derivative \(f_x(x_0, y_0)\) measures the instantaneous rate of change of \(f\) as only the \(x\) variable changes; likewise, \(f_y(x_0, y_0)\) measures the rate of change of \(f\) at \((x_0, y_0)\) as only \(y\) changes. Note particularly that \(f_x(x_0, y_0)\) is measured in “units of \(f\) per unit of change in \(x\),” and that the units on \(f_y(x_0, y_0)\) are similar.

In Preview Activity 10.6, we saw how we could measure the rate of change of \(f\) in a situation where both \(x\) and \(y\) were changing; in that activity, however, we found that this rate of change was measured in “units of \(f\) per unit of time.” In a given unit of time, we may move more than one unit of distance. In fact, in Preview Activity 10.6, in each unit increase in time we move a distance of \(|\mathbf{v}| = 5\) feet. To generalize the notion of partial derivatives to any direction of our choice, we instead want to have a rate of change whose units are “units of \(f\) per unit of distance in the given direction.”

In this light, in order to formally define the derivative in a particular direction of motion, we want to represent the change in \(f\) for a given unit change in the direction of motion. We can represent this unit change in direction with a unit vector, say \(\mathbf{u} = \langle u_1, u_2 \rangle\). If we move a distance \(h\) in the direction of \(\mathbf{u}\) from a fixed point \((x_0, y_0)\), we then arrive at the new point \((x_0 + u_1h, y_0 + u_2h)\).

It now follows that the slope of the secant line to the curve on the surface through \((x_0, y_0)\) in the direction of \(\mathbf{u}\) through the points \((x_0, y_0)\) and \((x_0 + u_1h, y_0 + u_2h)\) is

\[
m_{\text{sec}} = \frac{f(x_0 + u_1h, y_0 + u_2h) - f(x_0, y_0)}{h}.
\]

(10.9)
10.6. DIRECTIONAL DERIVATIVES AND THE GRADIENT

To get the instantaneous rate of change of \( f \) in the direction \( \mathbf{u} = \langle u_1, u_2 \rangle \), we must take the limit of the quantity in Equation (10.9) as \( h \to 0 \). Doing so results in the formal definition of the directional derivative.

**Definition 10.4.** Let \( f = f(x, y) \) be given. The derivative of \( f \) at the point \((x, y)\) in the direction of the unit vector \( \mathbf{u} = \langle u_1, u_2 \rangle \) is denoted \( D_{\mathbf{u}} f(x, y) \) and is given by

\[
D_{\mathbf{u}} f(x, y) = \lim_{h \to 0} \frac{f(x + u_1 h, y + u_2 h) - f(x, y)}{h}
\]

for those values of \( x \) and \( y \) for which the limit exists.

The quantity \( D_{\mathbf{u}} f(x, y) \) is called a directional derivative. When we evaluate the directional derivative \( D_{\mathbf{u}} f(x, y) \) at a point \((x_0, y_0)\), the result \( D_{\mathbf{u}} f(x_0, y_0) \) tells us the instantaneous rate at which \( f \) changes at \((x_0, y_0)\) per unit increase in the direction of the vector \( \mathbf{u} \). In addition, the quantity \( D_{\mathbf{u}} f(x_0, y_0) \) tells us the slope of the line tangent to the surface in the direction of \( \mathbf{u} \) at the point \((x_0, y_0, f(x_0, y_0))\).

**Computing the Directional Derivative**

In a similar way to how we developed shortcut rules for standard derivatives in single variable calculus, and for partial derivatives in multivariable calculus, we can also find a way to evaluate directional derivatives without resorting to the limit definition found in Equation (10.10). We do so using a very similar approach to our work in Preview Activity 10.6.

Suppose we consider the situation where we are interested in the instantaneous rate of change of \( f \) at a point \((x_0, y_0)\) in the direction \( \mathbf{u} = \langle u_1, u_2 \rangle \), where \( \mathbf{u} \) is a unit vector. The variables \( x \) and \( y \) are therefore changing according to the parameterization

\[
x = x_0 + u_1 t \quad \text{and} \quad y = y_0 + u_2 t.
\]

Observe that \( \frac{dx}{dt} = u_1 \) and \( \frac{dy}{dt} = u_2 \) for all values of \( t \). Since \( \mathbf{u} \) is a unit vector, it follows that a point moving along this line moves one unit of distance per one unit of time; that is, each single unit of time corresponds to movement of a single unit of distance in that direction. This observation allows us to use the Chain Rule to calculate the directional derivative, which measures the instantaneous rate of change of \( f \) with respect to change in the direction \( \mathbf{u} \).

In particular, by the Chain Rule, it follows that

\[
D_{\mathbf{u}} f(x_0, y_0) = f_x(x_0, y_0) \frac{dx}{dt} \bigg|_{(x_0, y_0)} + f_y(x_0, y_0) \frac{dy}{dt} \bigg|_{(x_0, y_0)} = f_x(x_0, y_0) u_1 + f_y(x_0, y_0) u_2.
\]

This now allows us to compute the directional derivative at an arbitrary point according to the
Given a differentiable function \( f = f(x, y) \) and a unit vector \( \mathbf{u} = \langle u_1, u_2 \rangle \), we may compute \( D_{\mathbf{u}}f(x, y) \) by
\[
D_{\mathbf{u}}f(x, y) = f_x(x, y)u_1 + f_y(x, y)u_2. \tag{10.11}
\]

**Note well:** To use Equation (10.11), we must have a unit vector \( \mathbf{u} = \langle u_1, u_2 \rangle \) in the direction of motion. In the event that we have a direction prescribed by a non-unit vector, we must first scale the vector to have length 1.

**Activity 10.16.**
Let \( f(x, y) = 3xy - x^2y^3 \).

(a) Determine \( f_x(x, y) \) and \( f_y(x, y) \).

(b) Use Equation (10.11) to determine \( D_1f(x, y) \) and \( D_2f(x, y) \). What familiar function is \( D_1f \)? What familiar function is \( D_2f \)?

(c) Use Equation (10.11) to find the derivative of \( f \) in the direction of the vector \( \mathbf{v} = \langle 2, 3 \rangle \) at the point \( (1, -1) \). Remember that a unit direction vector is needed.

The Gradient

Via the Chain Rule, we have seen that for a given function \( f = f(x, y) \), its instantaneous rate of change in the direction of a unit vector \( \mathbf{u} = \langle u_1, u_2 \rangle \) is given by
\[
D_{\mathbf{u}}f(x_0, y_0) = f_x(x_0, y_0)u_1 + f_y(x_0, y_0)u_2. \tag{10.12}
\]

Recalling that the dot product of two vectors \( \mathbf{v} = \langle v_1, v_2 \rangle \) and \( \mathbf{u} = \langle u_1, u_2 \rangle \) is computed by
\[
\mathbf{v} \cdot \mathbf{u} = v_1u_1 + v_2u_2,
\]
we see that we may recast Equation (10.12) in a way that has geometric meaning. In particular, we see that \( D_{\mathbf{u}}f(x_0, y_0) \) is the dot product of the vector \( \langle f_x(x_0, y_0), f_y(x_0, y_0) \rangle \) and the vector \( \mathbf{u} \).

We call this vector formed by the partial derivatives of \( f \) the **gradient** of \( f \) and denote it
\[
\nabla f(x_0, y_0) = \langle f_x(x_0, y_0), f_y(x_0, y_0) \rangle.
\]

We read \( \nabla f \) as “the gradient of \( f \),” “grad \( f \)” or “del \( f \).” Notice that \( \nabla f \) varies from point to point. In the following activity, we investigate some of what the gradient tells us about the behavior of a function \( f \).

---

The symbol \( \nabla \) is called **nabla**, which comes from a Greek word for a certain type of harp that has a similar shape.
Activity 10.17.

Let’s consider the function $f$ defined by $f(x, y) = x^2 - y^2$. Some contours for this function are shown in Figure 10.41.

![Figure 10.41: Contours of $f(x, y) = x^2 - y^2$.](image)

(a) Find the gradient $\nabla f(x, y)$.

(b) For each of the following points $(x_0, y_0)$, evaluate the gradient $\nabla f(x_0, y_0)$ and sketch the gradient vector with its tail at $(x_0, y_0)$. Some of the vectors are too long to fit onto the plot, but we’d like to draw them to scale; to do so, scale each vector by a factor of 1/4.

- $(x_0, y_0) = (2, 0)$
- $(x_0, y_0) = (0, 2)$
- $(x_0, y_0) = (2, 2)$
- $(x_0, y_0) = (2, 1)$
- $(x_0, y_0) = (-3, 2)$
- $(x_0, y_0) = (-2, -4)$
- $(x_0, y_0) = (0, 0)$

(c) What do you notice about the relationship between the gradient at $(x_0, y_0)$ and the contour line passing through that point?

(d) Does $f$ increase or decrease in the direction of $\nabla f(x_0, y_0)$? Provide a justification for your response.

As a vector, $\nabla f(x_0, y_0)$ defines a direction and a length. As we will soon see, both of these convey important information about the behavior of $f$ near $(x_0, y_0)$. 

\[ \nabla f(x, y) = \left( \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right) \]
The Direction of the Gradient

Remember that the dot product also conveys information about the angle between the two vectors. If \( \theta \) is the angle between \( \nabla f(x_0, y_0) \) and \( u \) (where \( u \) is a unit vector), then we also have that

\[
D_u f(x_0, y_0) = \nabla f(x_0, y_0) \cdot u = |\nabla f(x_0, y_0)||u| \cos(\theta).
\]

In particular, when \( \theta \) is a right angle, as shown on the left of Figure 10.42, then \( D_u f(x_0, y_0) = 0 \), because \( \cos(\theta) = 0 \). Since the value of the directional derivative is 0, this means that \( f \) is unchanging in this direction, and hence \( u \) must be tangent to the contour of \( f \) that passes through \( (x_0, y_0) \). In other words, \( \nabla f(x_0, y_0) \) is orthogonal to the contour through \( (x_0, y_0) \). This shows that the gradient vector at a given point is always perpendicular to the contour passing through the point, confirming that what we saw in part (c) of Activity 10.17 holds in general.

Moreover, when \( \theta \) is an acute angle, it follows that \( \cos(\theta) > 0 \) so since

\[
D_u f(x_0, y_0) = |\nabla f(x_0, y_0)||u| \cos(\theta),
\]

and therefore \( D_u f(x_0, y_0) > 0 \), as shown in the middle image in Figure 10.42. This means that \( f \) is increasing in any direction where \( \theta \) is acute. In a similar way, when \( \theta \) is an obtuse angle, then \( \cos(\theta) < 0 \) so \( D_u f(x_0, y_0) < 0 \), as seen on the right in Figure 10.42. This means that \( f \) is decreasing in any direction for which \( \theta \) is obtuse.

Finally, as we can see in the following activity, we may also use the gradient to determine the directions in which the function is increasing and decreasing most rapidly.

**Activity 10.18.**

In this activity we investigate how the gradient is related to the directions of greatest increase and decrease of a function. Let \( f \) be a differentiable function and \( u \) a unit vector.

(a) Let \( \theta \) be the angle between \( \nabla f(x_0, y_0) \) and \( u \). Explain why

\[
D_u f(x_0, y_0) = |\langle f_x(x_0, y_0), f_y(x_0, y_0)\rangle| \cos(\theta).
\]
(b) At the point \((x_0, y_0)\), the only quantity in Equation (10.13) that can change is \(\theta\) (which determines the direction \(u\) of travel). Explain why \(\theta = 0\) makes the quantity
\[
|\langle f_x(x_0, y_0), f_y(x_0, y_0) \rangle| \cos(\theta)
\]
as large as possible.

(c) When \(\theta = 0\), in what direction does the unit vector \(u\) point relative to \(\nabla f(x_0, y_0)\)? Why? What does this tell us about the direction of greatest increase of \(f\) at the point \((x_0, y_0)\)?

(d) In what direction, relative to \(\nabla f(x_0, y_0)\), does \(f\) decrease most rapidly at the point \((x_0, y_0)\)?

(e) State the unit vectors \(u\) and \(v\) (in terms of \(\nabla f(x_0, y_0)\)) that provide the directions of greatest increase and decrease for the function \(f\) at the point \((x_0, y_0)\). What important assumption must we make regarding \(\nabla f(x_0, y_0)\) in order for these vectors to exist?

\[
\text{The Length of the Gradient}
\]

Having established in Activity 10.18 that the direction in which a function increases most rapidly at a point \((x_0, y_0)\) is the unit vector \(u\) in the direction of the gradient, (that is, \(u = \frac{1}{|\nabla f(x_0, y_0)|} \nabla f(x_0, y_0)\), provided that \(\nabla f(x_0, y_0) \neq 0\), it is also natural to ask, “in the direction of greatest increase for \(f\) at \((x_0, y_0)\), what is the value of the rate of increase?” In this situation, we are asking for the value of \(D_u f(x_0, y_0)\) where \(u = \frac{1}{|\nabla f(x_0, y_0)|} \nabla f(x_0, y_0)\).

Using the now familiar way to compute the directional derivative, we see that
\[
D_u f(x_0, y_0) = \nabla f(x_0, y_0) \cdot \left( \frac{1}{|\nabla f(x_0, y_0)|} \nabla f(x_0, y_0) \right).
\]

Next, we recall two important facts about the dot product: (i) \(w \cdot (cv) = c(w \cdot v)\) for any scalar \(c\), and (ii) \(w \cdot w = |w|^2\). Applying these properties to the most recent equation involving the directional derivative, we find that
\[
D_u f(x_0, y_0) = \frac{1}{|\nabla f(x_0, y_0)|} (\nabla f(x_0, y_0) \cdot \nabla f(x_0, y_0)) = \frac{1}{|\nabla f(x_0, y_0)|} |\nabla f(x_0, y_0)|^2.
\]

Finally, since \(\nabla f(x_0, y_0)\) is a nonzero vector, its length \(|\nabla f(x_0, y_0)|\) is a nonzero scalar, and thus we can simplify the preceding equation to establish that
\[
D_u f(x_0, y_0) = |\nabla f(x_0, y_0)|.
\]

We summarize our most recent work by stating two important facts about the gradient.

Let \(f\) be a differentiable function and \((x_0, y_0)\) a point for which \(\nabla f(x_0, y_0) \neq 0\). Then \(\nabla f(x_0, y_0)\) points in the direction of greatest increase of \(f\) at \((x_0, y_0)\), and the instantaneous rate of change of \(f\) in that direction is the length of the gradient vector. That is, if \(u = \frac{1}{|\nabla f(x_0, y_0)|} \nabla f(x_0, y_0)\), then \(u\) is a unit vector in the direction of greatest increase of \(f\) at \((x_0, y_0)\), and \(D_u f(x_0, y_0) = |\nabla f(x_0, y_0)|\).
Activity 10.19. 

Consider the function \( f \) defined by \( f(x, y) = 2x^2 - xy + 2y. \)

(a) Find the gradient \( \nabla f(1, 2) \) and sketch it on Figure 10.43.

![Figure 10.43: A plot for the gradient \( \nabla f(1, 2) \).](image)

(b) Sketch the unit vector \( \mathbf{z} = \left\langle -\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}} \right\rangle \) on Figure 10.43 with its tail at \((1, 2)\). Now find the directional derivative \( D_{\mathbf{z}}f(1, 2) \).

(c) What is the slope of the graph of \( f \) in the direction \( \mathbf{z} \)? What does the sign of the directional derivative tell you?

(d) Consider the vector \( \mathbf{v} = \langle 2, -1 \rangle \) and sketch \( \mathbf{v} \) on Figure 10.43 with its tail at \((1, 2)\). Find a unit vector \( \mathbf{w} \) pointing in the same direction of \( \mathbf{v} \). Without computing \( D_{\mathbf{w}}f(1, 2) \), what do you know about the sign of this directional derivative? Now verify your observation by computing \( D_{\mathbf{w}}f(1, 2) \).

(e) In which direction (that is, for what unit vector \( \mathbf{u} \)) is \( D_{\mathbf{u}}f(1, 2) \) the greatest? What is the slope of the graph in this direction?

(f) Corresponding, in which direction is \( D_{\mathbf{u}}f(1, 2) \) least? What is the slope of the graph in this direction?

(g) Sketch two unit vectors \( \mathbf{u} \) for which \( D_{\mathbf{u}}f(1, 2) = 0 \) and then find component representations of these vectors.

(h) Suppose you are standing at the point \((3, 3)\). In which direction should you move to cause \( f \) to increase as rapidly as possible? At what rate does \( f \) increase in this direction?
Applications

The gradient finds many natural applications. For example, situations often arise – for instance, constructing a road through the mountains or planning the flow of water across a landscape – where we are interested in knowing the direction in which a function is increasing or decreasing most rapidly.

For example, consider a two-dimensional version of how a heat-seeking missile might work.\(^5\) Suppose that the temperature surrounding a fighter jet can be modeled by the function \(T\) defined by

\[
T(x, y) = \frac{100}{1 + (x - 5)^2 + 4(y - 2.5)^2},
\]

where \((x, y)\) is a point in the plane of the fighter jet and \(T(x, y)\) is measured in degrees Celsius. Some contours and gradients \(\nabla T\) are shown on the left in Figure 10.44.

![Figure 10.44: Contours and gradient for \(T(x, y)\) and the missile’s path.](image)

A heat-seeking missile will always travel in the direction in which the temperature increases most rapidly; that is, it will always travel in the direction of the gradient \(\nabla T\). If a missile is fired from the point \((2, 4)\), then its path will be that shown on the right in Figure 10.44.

In the final activity of this section, we consider several questions related to this context of a heat-seeking missile, and foreshadow some upcoming work in Section 10.7.

**Activity 10.20.**

(a) The temperature \(T(x, y)\) has its maximum value at the fighter jet’s location. State the fighter jet’s location and explain how Figure 10.44 tells you this.

(b) Determine \(\nabla T\) at the fighter jet’s location and give a justification for your response.

(c) Suppose that a different function \(f\) has a local maximum value at \((x_0, y_0)\). Sketch the behavior of some possible contours near this point. What is \(\nabla f(x_0, y_0)\)?

(d) Suppose that a function \(g\) has a local minimum value at \((x_0, y_0)\). Sketch the behavior of some possible contours near this point. What is \(\nabla g(x_0, y_0)\)?

---

\(^5\)This application is borrowed from United States Air Force Academy Department of Mathematical Sciences at [http://www.nku.edu/~longa/classes/mat320/mathematica/multcalc.htm](http://www.nku.edu/~longa/classes/mat320/mathematica/multcalc.htm).
(e) If a function $g$ has a local minimum at $(x_0, y_0)$, what is the direction of greatest increase of $g$ at $(x_0, y_0)$?

Summary

- The directional derivative of $f$ at the point $(x, y)$ in the direction of the unit vector $u = \langle u_1, u_2 \rangle$ is
  \[ D_u f(x, y) = \lim_{h \to 0} \frac{f(x + u_1 h, y + u_2 h) - f(x, y)}{h} \]
  for those values of $x$ and $y$ for which the limit exists. In addition, $D_u f(x, y)$ measures the slope of the graph of $f$ when we move in the direction $u$. Alternatively, $D_u f(x_0, y_0)$ measures the instantaneous rate of change of $f$ in the direction $u$ at $(x_0, y_0)$.

- The gradient of a function $f = f(x, y)$ at a point $(x_0, y_0)$ is the vector
  \[ \nabla f(x_0, y_0) = \langle f_x(x_0, y_0), f_y(x_0, y_0) \rangle. \]

- The directional derivative in the direction $u$ may be computed by
  \[ D_u f(x_0, y_0) = \nabla f(x_0, y_0) \cdot u. \]

- At any point where the gradient is nonzero, gradient is orthogonal to the contour through that point and points in the direction in which $f$ increases most rapidly; moreover, the slope of $f$ in this direction equals the length of the gradient $|\nabla f(x_0, y_0)|$. Similarly, the opposite of the gradient points in the direction of greatest decrease, and that rate of decrease is the opposite of the length of the gradient.

Exercises

1. Let $E(x, y) = \frac{100}{1 + (x - 5)^2 + 4(y - 2.5)^2}$ represent the elevation on a land mass at location $(x, y)$. Suppose that $E, x, \text{ and } y$ are all measured in meters.
   (a) Find $E_x(x, y)$ and $E_y(x, y)$.
   (b) Let $u$ be a unit vector in the direction of $(-4, 3)$. Determine $D_u E(3, 4)$. What is the practical meaning of $D_u E(3, 4)$ and what are its units?
   (c) Find the direction of greatest increase in $E$ at the point $(3, 4)$.
   (d) Find the instantaneous rate of change of $E$ in the direction of greatest decrease at the point $(3, 4)$. Include units on your answer.
   (e) At the point $(3, 4)$, find a direction $w$ in which the instantaneous rate of change of $E$ is 0.

2. Let $f(x, y) = x^2 + 3y^2$. 

 experimentation icon
(a) Find \( \nabla f(x, y) \) and \( \nabla f(1, 2) \).

(b) Find the direction of greatest increase in \( f \) at the point \((1, 2)\). Explain. A graph of the surface defined by \( f \) is shown in Figure 10.45. Illustrate this direction on the surface.

(c) A contour diagram of \( f \) is shown in Figure 10.46. Illustrate your calculation from (b) on this contour diagram.

(d) Find a direction \( \mathbf{w} \) for which the slope of the tangent line to the surface generated by \( f \) at the point \((1, 2)\) is zero in the direction \( \mathbf{w} \).

3. The properties of the gradient that we have observed for functions of two variables also hold for functions of more variables. In this problem, we consider a situation where there are three independent variables. Suppose that the temperature in a region of space is described by

\[
T(x, y, z) = 100e^{-x^2 - y^2 - z^2}
\]

and that you are standing at the point \((1, 2, -1)\).

(a) Find the instantaneous rate of change of the temperature in the direction of \( \mathbf{v} = (0, 1, 2) \) at the point \((1, 2, -1)\). Remember that you should first find a unit vector in the direction of \( \mathbf{v} \).

(b) In what direction from the point \((1, 2, -1)\) would you move to cause the temperature to decrease as quickly as possible?

(c) How fast does the temperature decrease in this direction?

(d) Find a direction in which the temperature does not change at \((1, 2, -1)\).
4. Figure 10.47 shows a plot of the gradient $\nabla f$ at several points for some function $f = f(x, y)$.
   
   (a) Consider each of the three indicated points, and draw, as best as you can, the contour through that point.
   
   (b) Beginning at each point, draw a curve on which $f$ is continually decreasing.
10.7 Optimization

Motivating Questions

In this section, we strive to understand the ideas generated by the following important questions:

• How can we find the points at which $f(x, y)$ has a local maximum or minimum?
• How can we determine whether critical points of $f(x, y)$ are local maxima or minima?
• How can we find the absolute maximum and minimum of $f(x, y)$ on a closed and bounded domain?

Introduction

We learn in single-variable calculus that the derivative is a useful tool for finding the local maxima and minima of functions, and that these ideas may often be employed in applied settings. In particular, if a function $f$, such as the one shown in Figure 10.48 is everywhere differentiable, we know that the tangent line is horizontal at any point where $f$ has a local maximum or minimum. This, of course, means that the derivative $f'$ is zero at any such point. Hence, one way that we seek extreme values of a given function is to first find where the derivative of the function is zero.

In multivariable calculus, we are often similarly interested in finding the greatest and/or least value(s) that a function may achieve. Moreover, there are many applied settings in which a quantity of interest depends on several different variables. In the following preview activity, we begin to see how some key ideas in multivariable calculus can help us answer such questions by thinking about the geometry of the surface generated by a function of two variables.

Preview Activity 10.7. Let $z = f(x, y)$ be a differentiable function, and suppose that at the point $(x_0, y_0)$, $f$ achieves a local maximum. That is, the value of $f(x_0, y_0)$ is greater than the value...
of \( f(x, y) \) for all \((x, y)\) nearby \((x_0, y_0)\). You might find it helpful to sketch a rough picture of a possible function \( f \) that has this property.

(a) If we consider the trace given by holding \( y = y_0 \) constant, then the single-variable function defined by \( f(x, y_0) \) must have a local maximum at \( x_0 \). What does this say about the value of the partial derivative \( f_x(x_0, y_0) \)?

(b) In the same way, the trace given by holding \( x = x_0 \) constant has a local maximum at \( y = y_0 \). What does this say about the value of the partial derivative \( f_y(x_0, y_0) \)?

(c) What may we now conclude about the gradient \( \nabla f(x_0, y_0) \) at the local maximum? How is this consistent with the statement “\( f \) increases most rapidly in the direction \( \nabla f(x_0, y_0) \)”?

(d) How will the tangent plane to the surface \( z = f(x, y) \) appear at the point \((x_0, y_0, f(x_0, y_0))\)?

(e) By first computing the partial derivatives, find any points at which \( f(x, y) = 2x - x^2 - (y + 2)^2 \) may have a local maximum.

Exrema and Critical Points

One of the important applications of single-variable calculus is the use of derivatives to identify local extremes of functions (that is, local maxima and local minima). Using the tools we have developed so far, we can naturally extend the concept of local maxima and minima to several-variable functions.

**Definition 10.5.** Let \( f \) be a function of two variables \( x \) and \( y \).

- The function \( f \) has a **local maximum** at a point \((x_0, y_0)\) provided that \( f(x, y) \leq f(x_0, y_0) \) for all points \((x, y)\) near \((x_0, y_0)\). In this situation we say that \( f(x_0, y_0) \) is a **local maximum value**.

- The function \( f \) has a **local minimum** at a point \((x_0, y_0)\) provided that \( f(x, y) \geq f(x_0, y_0) \) for all points \((x, y)\) near \((x_0, y_0)\). In this situation we say that \( f(x_0, y_0) \) is a **local minimum value**.

- An **absolute maximum point** is a point \((x_0, y_0)\) for which \( f(x, y) \leq f(x_0, y_0) \) for all points \((x, y)\) in the domain of \( f \). The value of \( f \) at an absolute maximum point is the **maximum value** of \( f \).

- An **absolute minimum point** is a point such that \( f(x, y) \geq f(x_0, y_0) \) for all points \((x, y)\) in the domain of \( f \). The value of \( f \) at an absolute minimum point is the **maximum value** of \( f \).
We use the term **extremum point** to refer to any point \((x_0, y_0)\) at which \(f\) has a local maximum or minimum. In addition, the function value \(f(x_0, y_0)\) at an extremum is called an **extremal value**. Figure 10.49 illustrates the graphs of two functions that have an absolute maximum and minimum, respectively, at the origin \((x_0, y_0) = (0, 0)\).

![Figure 10.49: An absolute maximum and an absolute minimum](image)

In single-variable calculus, we saw that the extrema of a continuous function \(f\) always occur at **critical points**, values of \(x\) where \(f\) fails to be differentiable or where \(f'(x) = 0\). Said differently, critical points provide the locations where extrema of a function may appear. Our work in Preview Activity 10.7 suggests that something similar happens with two-variable functions.

Suppose that a continuous function \(f\) has an extremum at \((x_0, y_0)\). In this case, the trace \(f(x, y_0)\) has an extremum at \(x_0\), which means that \(x_0\) is a critical value of \(f(x, y_0)\). Therefore, either \(f_x(x_0, y_0)\) does not exist or \(f_x(x_0, y_0) = 0\). Similarly, either \(f_y(x_0, y_0)\) does not exist or \(f_y(x_0, y_0) = 0\). This implies that the extrema of a two-variable function occur at points that satisfy the following definition.

**Definition 10.6.** A **critical point** \((x_0, y_0)\) of a function \(f = f(x, y)\) is a point in the domain of \(f\) at which \(f_x(x_0, y_0) = 0\) and \(f_y(x_0, y_0) = 0\), or such that one of \(f_x(x_0, y_0)\) or \(f_y(x_0, y_0)\) fails to exist.

We can therefore find critical points of a function \(f\) by computing partial derivatives and identifying any values of \((x, y)\) for which one of the partials doesn’t exist or for which both partial derivatives are simultaneously zero. For the latter, note that we have to solve the system of equations

\[
\begin{align*}
    f_x(x, y) &= 0 \\
    f_y(x, y) &= 0.
\end{align*}
\]
Activity 10.21.

Find the critical points of each of the following functions. Then, using appropriate technology (e.g., Wolfram|Alpha or CalcPlot3D⁶), plot the graphs of the surfaces near each critical value and compare the graph to your work.

(a) \( f(x, y) = 2 + x^2 + y^2 \)
(b) \( f(x, y) = 2 + x^2 - y^2 \)
(c) \( f(x, y) = 2x - x^2 - \frac{1}{3}y^2 \)
(d) \( f(x, y) = |x| + |y| \)
(e) \( f(x, y) = 2xy - 4x + 2y - 3 \).

Classifying Critical Points: The Second Derivative Test

While the extrema of a continuous function \( f \) always occur at critical points, it is important to note that not every critical point leads to an extremum. Recall, for instance, \( f(x) = x^3 \) from single variable calculus. We know that \( x_0 = 0 \) is a critical point since \( f'(x_0) = 0 \), but \( x_0 = 0 \) is neither a local maximum nor a local minimum of \( f \).

A similar situation may arise in a multivariable setting. Consider the function \( f \) defined by \( f(x, y) = x^2 - y^2 \) whose graph and contour plot are shown in Figure 10.50. Because \( \nabla f = (2x, -2y) \), we see that the origin \((x_0, y_0) = (0, 0)\) is a critical point. However, this critical point is neither a local maximum or minimum; the origin is a local minimum on the trace defined by \( y = 0 \), while the origin is a local maximum on the trace defined by \( x = 0 \). We call such a critical point a saddle point due to the shape of the graph near the critical point.

Figure 10.50: A saddle point.

As in single-variable calculus, we would like to have some sort of test to help us identify whether a critical point is a local maximum, local minimum, or neither. Before describing the test that accomplishes this classification in a two-variable setting, we recall the Second Derivative Test from single-variable calculus.

**The Second Derivative Test** for single-variable functions. If \( x_0 \) is a critical point of a function \( f \) so that \( f'(x_0) = 0 \) and if \( f''(x_0) \) exists, then

- if \( f''(x_0) < 0 \), \( x_0 \) is a local maximum,
- if \( f''(x_0) > 0 \), \( x_0 \) is a local minimum, and
- if \( f''(x_0) = 0 \), this test yields no information.

For the analogous test for functions of two variables, we have to consider all of the second-order partial derivatives.

**The Second Derivative Test** for two-variable functions. Suppose \((x_0, y_0)\) is a critical point of the function \( f \) for which \( f_x(x_0, y_0) = 0 \) and \( f_y(x_0, y_0) = 0 \). Let \( D \) be the quantity defined by

\[
D = f_{xx}(x_0, y_0)f_{yy}(x_0, y_0) - f_{xy}(x_0, y_0)^2.
\]

- If \( D > 0 \) and \( f_{xx}(x_0, y_0) < 0 \), then \( f \) has a local maximum at \((x_0, y_0)\).
- If \( D > 0 \) and \( f_{xx}(x_0, y_0) > 0 \), then \( f \) has a local minimum at \((x_0, y_0)\).
- If \( D < 0 \), then \( f \) has a saddle point at \((x_0, y_0)\).
- If \( D = 0 \), then this test yields no information about what happens at \((x_0, y_0)\).

The quantity \( D \) is called the *discriminant* of the function \( f \) at \((x_0, y_0)\).

To properly understand the origin of the Second Derivative Test, we could introduce a “second-order directional derivative.” If this second-order directional derivative were negative in every direction, for instance, we could guarantee that the critical point is a local maximum. A complete justification of the Second Derivative Test requires key ideas from linear algebra that are beyond the scope of this course, so instead of presenting a detailed explanation, we will accept this test as stated and demonstrate its use in three basic examples.
Example 10.1. Let \( f(x, y) = 4 - x^2 - y^2 \) as shown in Figure 10.51. Critical points occur when \( f_x = -2x = 0 \) and \( f_y = -2y = 0 \) so the origin \((x_0, y_0) = (0, 0)\) is the only critical point. We then find that

\[
f_{xx}(0, 0) = -2, \quad f_{yy}(0, 0) = -2, \quad \text{and} \quad f_{xy}(0, 0) = 0,
\]

giving \( D = (-2)(-2) - 0^2 = 4 > 0 \). We then consider \( f_{xx}(0, 0) = -2 < 0 \) and conclude, from the Second Derivative Test, that \( f(0, 0) = 4 \) is a local maximum value.

Example 10.2. Let \( f(x, y) = x^2 + y^2 \) as shown in Figure 10.52. Critical points occur when \( f_x = 2x = 0 \) and \( f_y = 2y = 0 \) so the origin \((x_0, y_0) = (0, 0)\) is the only critical point. We then find that

\[
f_{xx}(0, 0) = 2, \quad f_{yy}(0, 0) = 2, \quad \text{and} \quad f_{xy}(0, 0) = 0,
\]

giving \( D = 2 \cdot 2 - 0^2 = 4 > 0 \). We then consider \( f_{xx}(0, 0) = 2 > 0 \) and conclude, from the Second Derivative Test, that \( f(0, 0) = 0 \) is a local minimum value.

Example 10.3. Let \( f(x, y) = x^2 - y^2 \) as shown in Figure 10.53. Critical points occur when \( f_x = 2x = 0 \) and \( f_y = -2y = 0 \) so the origin \((x_0, y_0) = (0, 0)\) is the only critical point. We then find that

\[
f_{xx}(0, 0) = 2, \quad f_{yy}(0, 0) = -2, \quad \text{and} \quad f_{xy}(0, 0) = 0,
\]

giving \( D = 2 \cdot (-2) - 0^2 = -4 < 0 \). We then conclude, from the Second Derivative Test, that \((0, 0)\) is a saddle point.

Activity 10.22.

Find the critical points of the following functions and use the Second Derivative Test to classify the critical points.

(a) \( f(x, y) = 3x^3 + y^2 - 9x + 4y \)

(b) \( f(x, y) = xy + \frac{2}{x} + \frac{4}{y} \)

(c) \( f(x, y) = x^3 + y^3 - 3xy \).

As we learned in single-variable calculus, finding extremal values of functions can be particularly useful in applied settings. For instance, we can often use calculus to determine the least
expensive way to construct something or to find the most efficient route between two locations. The same possibility holds in settings with two or more variables.

**Activity 10.23.**

While the quantity of a product demanded by consumers is often a function of the price of the product, the demand for a product may also depend on the price of other products. For instance, the demand for blue jeans at Old Navy may be affected not only by the price of the jeans themselves, but also by the price of khakis.

Suppose we have two goods whose respective prices are $p_1$ and $p_2$. The demand for these goods, $q_1$ and $q_2$, depend on the prices as

\[
q_1 = 150 - 2p_1 - p_2 \quad (10.14)
\]

\[
q_2 = 200 - p_1 - 3p_2 \quad (10.15)
\]

The seller would like to set the prices $p_1$ and $p_2$ in order to maximize revenue. We will assume that the seller meets the full demand for each product. Thus, if we let $R$ be the revenue obtained by selling $q_1$ items of the first good at price $p_1$ per item and $q_2$ items of the second good at price $p_2$ per item, we have

\[
R = p_1 q_1 + p_2 q_2.
\]

We can then write the revenue as a function of just the two variables $p_1$ and $p_2$ by using Equations (10.14) and (10.15), giving us

\[
R(p_1, p_2) = p_1 (150 - 2p_1 - p_2) + p_2 (200 - p_1 - 3p_2) = 150p_1 + 200p_2 - 2p_1p_2 - 2p_1^2 - 3p_2^2.
\]

A graph of $R$ as a function of $p_1$ and $p_2$ is shown in Figure 10.54.

![Figure 10.54: A revenue function.](image)

(a) Find all critical points of the revenue function, $R$. 

[Image of a 3D graph showing a revenue function with labeled axes and a defined range for $p_1$ and $p_2$.]
(b) Apply the Second Derivative Test to determine the type of any critical points.
(c) Where should the seller set the prices \( p_1 \) and \( p_2 \) to maximize the revenue?

### Optimization on a Restricted Domain

The Second Derivative Test helps us classify critical points of a function, but it does not tell us if the function actually has an absolute maximum or minimum at each such point. For single-variable functions, the Extreme Value Theorem told us that a continuous function on a closed interval \([a,b]\) always has both an absolute maximum and minimum on that interval, and that these absolute extremes must occur at either an endpoint or at a critical point. Thus, to find the absolute maximum and minimum, we determine the critical points in the interval and then evaluate the function at these critical point(s) and at the endpoints of the interval. A similar approach works for functions of two variables.

For functions of two variables, closed and bounded regions play the role that closed intervals did for functions of a single variable. A closed region is a region that contains its boundary (the unit disk \( x^2 + y^2 \leq 1 \) is closed, while its interior \( x^2 + y^2 < 1 \) is not, for example), while a bounded region is one that does not stretch to infinity in any direction. Just as for functions of a single variable, continuous functions of several variables that are defined on closed, bounded regions must have absolute maxima and minima in those regions.

#### The Extreme Value Theorem

Let \( f = f(x,y) \) be a continuous function on a closed and bounded region \( R \). Then \( f \) has an absolute maximum and an absolute minimum in \( R \).

The absolute extremes must occur at either a critical point in the interior of \( R \) or at a boundary point of \( R \). We therefore must test both possibilities, as we demonstrate in the following example.

**Example 10.4.** Suppose the temperature \( T \) at each point on the circular plate \( x^2 + y^2 \leq 1 \) is given by

\[
T(x, y) = 2x^2 + y^2 - y.
\]

The domain \( R = \{(x, y) : x^2 + y^2 \leq 1\} \) is a closed and bounded region, as shown on the left of Figure 10.55, so the Extreme Value Theorem assures us that \( T \) has an absolute maximum and minimum on the plate. The graph of \( T \) over its domain \( R \) is shown in Figure 10.55. We will find the hottest and coldest points on the plate.

If the absolute maximum or minimum occurs inside the disk, it will be at a critical point so we begin by looking for critical points inside the disk. To do this, notice that critical points are given by the conditions \( T_x = 4x = 0 \) and \( T_y = 2y - 1 = 0 \). This means that there is one critical point of the function at the point \((x_0, y_0) = (0, 1/2)\), which lies inside the disk.

We now find the hottest and coldest points on the boundary of the disk, which is the circle of radius 1. As we have seen, the points on the unit circle can be parametrized as

\[
x(t) = \cos(t), \quad y(t) = \sin(t),
\]
where $0 \leq t \leq 2\pi$. The temperature at a point on the circle is then described by

$$T(x(t), y(t)) = 2\cos^2(t) + \sin^2(t) - \sin(t).$$

To find the hottest and coldest point on the boundary, we look for the critical points of this single-variable function on the interval $0 \leq t \leq 2\pi$. We have

$$\frac{dT}{dt} = -4\cos(t)\sin(t) + 2\cos(t)\sin(t) - \cos(t) = -2\cos(t)\sin(t) - \cos(t) = \cos(t)(-2\sin(t) - 1) = 0.$$ 

This shows that we have critical points when $\cos(t) = 0$ or $\sin(t) = -1/2$. This occurs when $t = \pi/2, 3\pi/2, 7\pi/6$, and $11\pi/6$. Since we have $x(t) = \cos(t)$ and $y(t) = \sin(t)$, the corresponding points are

- $(x, y) = (0, 1)$ when $t = \frac{\pi}{2}$,
- $(x, y) = \left(\frac{-\sqrt{3}}{2}, -\frac{1}{2}\right)$ when $t = \frac{7\pi}{6}$,
- $(x, y) = (0, -1)$ when $t = \frac{3\pi}{2}$,
- $(x, y) = \left(\frac{\sqrt{3}}{2}, -\frac{1}{2}\right)$ when $t = \frac{11\pi}{6}$.

These are the critical points of $T$ on the boundary and so this collection of points includes the hottest and coldest points on the boundary.

We now have a list of candidates for the hottest and coldest points: the critical point in the interior of the disk and the critical points on the boundary. We find the hottest and coldest points by evaluating the temperature at each of these points, and find that

- $T(0, \frac{1}{2}) = -\frac{1}{4}$,
- $T\left(\frac{-\sqrt{3}}{2}, -\frac{1}{2}\right) = \frac{9}{4}$,
- $T(0, 1) = 0$,
- $T\left(\frac{\sqrt{3}}{2}, -\frac{1}{2}\right) = \frac{9}{4}$,
- $T(0, -1) = 2$,
So the maximum value of $T$ on the disk $x^2 + y^2 \leq 1$ is $\frac{9}{4}$, which occurs at the two points $\left(\pm \frac{\sqrt{3}}{2}, -\frac{1}{2}\right)$ on the boundary, and the minimum value of $T$ on the disk is $-\frac{1}{4}$ which occurs at the critical point $(0, \frac{1}{2})$ in the interior of $R$.

From this example, we see that we use the following procedure for determining the absolute maximum and absolute minimum of a function on a closed and bounded domain.

**Step 1:** Find all critical points of the function in the interior of the domain.

**Step 2:** Find all the critical points of the function on the boundary of the domain. Working on the boundary of the domain reduces this part of the problem to one or more single variable optimization problems.

**Step 3:** Evaluate the function at each of the points found in Steps 1 and 2.

**Step 4:** The maximum value of the function is the largest value obtained in Step 3, and the minimum value of the function is the smallest value obtained in Step 3.

**Activity 10.24.**

Let $f(x, y) = x^2 - 3y^2 - 4x + 6y$ with triangular domain $R$ whose vertices are at $(0, 0)$, $(4, 0)$, and $(0, 4)$. The domain $R$ and a graph of $f$ on the domain appear in Figure 10.56.

![Figure 10.56: The domain of $f(x, y) = x^2 - 3y^2 - 4x + 6y$ and its graph.](image)

(a) Find all of the critical points of $f$ in $R$.

(b) Parameterize the horizontal leg of the triangular domain, and find the critical points of $f$ on that leg.

(c) Parameterize the vertical leg of the triangular domain, and find the critical points of $f$ on that leg.

(d) Parameterize the hypotenuse of the triangular domain, and find the critical points of $f$ on the hypotenuse.
(e) Find the absolute maximum and absolute minimum value of $f$ on $R$. 

\[\]$
Summary

- To find the extrema of a function $f = f(x, y)$, we first find the critical points, which are points where one of the partials of $f$ fails to exist, or where $f_x = 0$ and $f_y = 0$.
- The Second Derivative Test helps determine whether a critical point is a local maximum, local minimum, or saddle point.
- If $f$ is defined on a closed and bounded domain, we find the absolute maxima and minima by finding the critical points in the interior of the domain, finding the critical points on the boundary, and testing the value of $f$ at both sets of critical points.

Exercises

1. Respond to each of the following prompts to solve the given optimization problem.
   
   (a) Let $f(x, y) = \sin(x) + \cos(y)$. Determine the absolute maximum and minimum values of $f$. At what points do these extreme values occur?
   
   (b) For a certain differentiable function $F$ of two variables $x$ and $y$, its partial derivatives are
       
       $$F_x(x, y) = x^2 - y - 4 \quad \text{and} \quad F_y(x, y) = -x + y - 2.$$  
       
       Find each of the critical points of $F$, and classify each as a local maximum, local minimum, or a saddle point.
   
   (c) Determine all critical points of $T(x, y) = 48 + 3xy - x^2y - xy^2$ and classify each as a local maximum, local minimum, or saddle point.
   
   (d) Find and classify all critical points of $g(x, y) = \frac{x^2}{2} + 3y^3 + 9y^2 - 3xy + 9y - 9x$
   
   (e) Find and classify all critical points of $z = f(x, y) = ye^{-x^2-2y^2}$.
   
   (f) Determine the absolute maximum and absolute minimum of $f(x, y) = 2 + 2x + 2y - x^2 - y^2$ on the triangular plate in the first quadrant bounded by the lines $x = 0$, $y = 0$, and $y = 9 - x$.
   
   (g) Determine the absolute maximum and absolute minimum of $f(x, y) = 2 + 2x + 2y - x^2 - y^2$ over the closed disk of points $(x, y)$ such that $(x - 1)^2 + (y - 1)^2 \leq 1$.
   
   (h) Find the point on the plane $z = 6 - 3x - 2y$ that lies closest to the origin.

2. If a continuous function $f$ of a single variable has two critical numbers $c_1$ and $c_2$ at which $f$ has relative maximum values, then $f$ must have another critical number $c_3$, because "it is impossible to have two mountains without some sort of valley in between. The other critical point can be a saddle point (a pass between the mountains) or a local minimum (a true valley).")

---

7From Calculus in Vector Spaces by Lawrence J. Corwin and Robert H. Szczarb.
Consider the function \( f \) defined by \( f(x, y) = 4x^2e^y - 2x^4 - e^{4y} \). Show that \( f \) has exactly two critical points of \( f \), and that \( f \) has relative maximum values at each of these critical points. Explain how function illustrates that it really is possible to have two mountains without some sort of valley in between. Use appropriate technology to draw the surface defined by \( f \) to see graphically how this happens.

3. If a continuous function \( f \) of a single variable has exactly one critical number with a relative maximum at that critical point, then the value of \( f \) at that critical point is an absolute maximum. In this exercise we see that the same is not always true for functions of two variables. Let \( f(x, y) = 3xe^y - x^3 - e^{3y} \). Show that \( f \) has exactly one critical point, has a relative maximum value at that critical point, but that \( f \) has no absolute maximum value. Use appropriate technology to draw the surface defined by \( f \) to see graphically how this happens.

4. A manufacturer wants to procure rectangular boxes to ship its product. The boxes must contain 20 cubic feet of space. To be durable enough to ensure the safety of the product, the material for the sides of the boxes will cost $0.10 per square foot, while the material for the top and bottom will cost $0.25 per square foot. In this activity we will help the manufacturer determine the box of minimal cost.

(a) What quantities are constant in this problem? What are the variables in this problem? Provide appropriate variable labels. What, if any, restrictions are there on the variables?

(b) Using your variables from (a), determine a formula for the total cost \( C \) of a box.

(c) Your formula in part (b) might be in terms of three variables. If so, find a relationship between the variables, and then use this relationship to write \( C \) as a function of only two independent variables.

(d) Find the dimensions that minimize the cost of a box. Be sure to verify that you have a minimum cost.

5. A rectangular box with length \( x \), width \( y \), and height \( z \) is being built. The box is positioned so that one corner is stationed at the origin and the box lies in the first octant where \( x \), \( y \), and \( z \) are all positive. There is an added constraint on how the box is constructed: it must fit underneath the plane with equation \( x + 2y + 3z = 6 \). In fact, we will assume that the corner of the box “opposite” the origin must actually lie on this plane. The basic problem is to find the maximum volume of the box.

(a) Sketch the plane \( x + 2y + 3z = 6 \), as well as a picture of a potential box. Label everything appropriately.

(b) Explain how you can use the fact that one corner of the box lies on the plane to write the volume of the box as a function of \( x \) and \( y \) only. Do so, and clearly show the formula you find for \( V(x, y) \).

---

\(^8\)From Ira Rosenholz in the Problems Section of the Mathematics Magazine, Vol. 60 NO. 1, February 1987.

(c) Find all critical points of $V$. (Note that when finding the critical points, it is essential that you factor first to make the algebra easier.)

(d) Without considering the current applied nature of the function $V$, classify each critical point you found above as a local maximum, local minimum, or saddle point of $V$.

(e) Determine the maximum volume of the box, justifying your answer completely with an appropriate discussion of the critical points of the function.

(f) Now suppose that we instead stipulated that, while the vertex of the box opposite the origin still had to lie on the plane, we were only going to permit the sides of the box, $x$ and $y$, to have values in a specified range (given below). That is, we now want to find the maximum value of $V$ on the closed, bounded region

$$\frac{1}{2} \leq x \leq 1, \quad 1 \leq y \leq 2.$$ 

Find the maximum volume of the box under this condition, justifying your answer fully.

6. The airlines place restrictions on luggage that can be carried onto planes.

- A carry-on bag can weigh no more than 40 lbs.
- The length plus width plus height of a bag cannot exceed 45 inches.
- The bag must fit in an overhead bin.

Let $x$, $y$, and $z$ be the length, width, and height (in inches) of a carry on bag. In this problem we find the dimensions of the bag of largest volume, $V = xyz$, that satisfies the second restriction. Assume that we use all 45 inches to get a maximum volume. (Note that this bag of maximum volume might not satisfy the third restriction.)

(a) Write the volume $V = V(x, y)$ as a function of just the two variables $x$ and $y$.

(b) Explain why the domain over which $V$ is defined is the triangular region $R$ with vertices $(0,0)$, $(45,0)$, and $(0,45)$.

(c) Find the critical points, if any, of $V$ in the interior of the region $R$.

(d) Find the maximum value of $V$ on the boundary of the region $R$, and the determine the dimensions of a bag with maximum volume on the entire region $R$. (Note that most carry-on bags sold today measure 22 by 14 by 9 inches with a volume of 2772 cubic inches, so that the bags will fit into the overhead bins.)

7. According to *The Song of Insects* by G.W. Pierce (Harvard College Press, 1948) the sound of striped ground crickets chirping, in number of chirps per second, is related to the temperature. So the number of chirps per second could be a predictor of temperature. The data Pierce collected is shown in the table below, where $x$ is the (average) number of chirps per second and $y$ is the temperature in degrees Fahrenheit.
A scatter plot of the data is given below. The relationship between $x$ and $y$ is not exactly linear, but looks to have a linear pattern. It could be that the relationship is really linear but experimental error causes the data to be slightly inaccurate. Or perhaps the data is not linear, but only approximately linear.

If we want to use the data to make predictions, then we need to fit a curve of some kind to the data. Since the cricket data appears roughly linear, we will fit a linear function $f$ of the form $f(x) = mx + b$ to the data. We will do this in such a way that we minimize the sums of the squares of the distances between the $y$ values of the data and the corresponding $y$ values of the line defined by $f$. This type of fit is called a least squares approximation. If the data is represented by the points $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$, then the square of the distance between $y_i$ and $f(x_i)$ is $(f(x_i) - y_i)^2 = (mx_i + b - y_i)^2$. So our goal is to minimize the sum of these squares, of minimize the function $S$ defined by

$$S(m, b) = \sum_{i=1}^{n} (mx_i + b - y_i)^2.$$ 

(a) Calculate $S_m$ and $S_b$.

(b) Solve the system $S_m(m, b) = 0$ and $S_b(m, b) = 0$ to show that the critical point satisfies

$$m = \frac{n \sum_{i=1}^{n} x_i y_i - (\sum_{i=1}^{n} x_i)(\sum_{i=1}^{n} y_i)}{n \sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2}$$

$$b = \frac{(\sum_{i=1}^{n} y_i)(\sum_{i=1}^{n} x_i^2) - (\sum_{i=1}^{n} x_i)(\sum_{i=1}^{n} x_i y_i)}{n \sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2}.$$
(Hint: Don’t be daunted by these expressions, the system $S_m(m, b) = 0$ and $S_b(m, b) = 0$ is a system of two linear equations in the unknowns $m$ and $b$. It might be easier to let $r = \sum_{i=1}^{n} x_i^2$, $s = \sum_{i=1}^{n} x_i$, $t = \sum_{i=1}^{n} y_i$, and $u = \sum_{i=1}^{n} x_i y_i$ and write your equations using these constants.)

(c) Use the Second Derivative Test to explain why the critical point gives a local minimum. Can you then explain why the critical point gives an absolute minimum?

(d) Use the formula from part (b) to find the values of $m$ and $b$ that give the line of best fit in the least squares sense to the cricket data. Draw your line on the scatter plot to convince yourself that you have a well-fitting line.
Motivating Questions

In this section, we strive to understand the ideas generated by the following important questions:

- What geometric condition enables us to optimize a function \( f = f(x, y) \) subject to a constraint given by \( g(x, y) = k \), where \( k \) is a constant?
- How can we exploit this geometric condition to find the extreme values of a function subject to a constraint?

Introduction

We previously considered how to find the extreme values of functions on both unrestricted domains and on closed, bounded domains. Other types of optimization problems involve maximizing or minimizing a quantity subject to an external constraint. In these cases the extreme values frequently won’t occur at the points where the gradient is zero, but rather at other points that satisfy an important geometric condition. These problems are often called constrained optimization problems and can be solved with the method of Lagrange Multipliers, which we study in this section.

Preview Activity 10.8. According to U.S. postal regulations, the girth plus the length of a parcel sent by mail may not exceed 108 inches, where by “girth” we mean the perimeter of the smallest end. Our goal is to find the largest possible volume of a rectangular parcel with a square end that can be sent by mail.\(^{10}\) If we let \( x \) be the length of the side of one square end of the package and \( y \) the length of the package, then we want to maximize the volume \( f(x, y) = x^2 y \) of the box subject to the constraint that the girth (4\(x\)) plus the length (\(y\)) is as large as possible, or \( 4x + y = 108 \). The equation \( 4x + y = 108 \) is thus an external constraint on the variables.

(a) The constraint equation involves the function \( g \) that is given by

\[
 g(x, y) = 4x + y.
\]

Explain why the constraint is a contour of \( g \), and is therefore a two-dimensional curve.

(b) Figure 10.57 shows the graph of the constraint equation \( g(x, y) = 108 \) along with a few contours of the volume function \( f \). Since our goal is to find the maximum value of \( f \) subject to the constraint \( g(x, y) = 108 \), we want to find the point on our constraint curve that intersects the contours of \( f \) at which \( f \) has its largest value.

\(^{10}\)We solved this applied optimization problem in single variable Active Calculus, so it may look familiar. We take a different approach in this section, and this approach allows us to view most applied optimization problems from single variable calculus as constrained optimization problems, as well as provide us tools to solve a greater variety of optimization problems.
Figure 10.57: Contours of \( f \) and the constraint equation \( g(x, y) = 108 \).

i. Points \( A \) and \( B \) in Figure 10.57 lie on a contour of \( f \) and on the constraint equation \( g(x, y) = 108 \). Explain why neither \( A \) nor \( B \) provides a maximum value of \( f \) that satisfies the constraint.

ii. Points \( C \) and \( D \) in Figure 10.57 lie on a contour of \( f \) and on the constraint equation \( g(x, y) = 108 \). Explain why neither \( C \) nor \( D \) provides a maximum value of \( f \) that satisfies the constraint.

iii. Based on your responses to parts i. and ii., draw the contour of \( f \) on which you believe \( f \) will achieve a maximum value subject to the constraint \( g(x, y) = 108 \). Explain why you drew the contour you did.

(c) Recall that \( g(x, y) = 108 \) is a contour of the function \( g \), and that the gradient of a function is always orthogonal to its contours. With this in mind, how should \( \nabla f \) and \( \nabla g \) be related at the optimal point? Explain.

\[ \nabla \]

Constrained Optimization and Lagrange Multipliers

In Preview Activity 10.8, we considered an optimization problem where there is an external constraint on the variables, namely that the girth plus the length of the package cannot exceed 108 inches. We saw that we can create a function \( g \) from the constraint, specifically \( g(x, y) = 4x + y \). The constraint equation is then just a contour of \( g \), \( g(x, y) = c \), where \( c \) is a constant (in our case 108). Figure 10.58 illustrates that the volume function \( f \) is maximized, subject to the constraint \( g(x, y) = c \), when the graph of \( g(x, y) = c \) is tangent to a contour of \( f \). Moreover, the value of \( f \) on this contour is the sought maximum value. To find this point where the graph of the constraint is tangent to a contour of \( f \), recall that \( \nabla f \) is perpendicular to the contours of \( f \) and \( \nabla g \) is perpendicular to the contour of \( g \). At such a point, the vectors \( \nabla g \) and \( \nabla f \) are parallel, and thus we need
to determine the points where this occurs. Recall that two vectors are parallel if one is a nonzero scalar multiple of the other, so we therefore look for values of a parameter \( \lambda \) that make

\[
\nabla f = \lambda \nabla g.
\] (10.16)

The constant \( \lambda \) is called a Lagrange multiplier.

To find the values of \( \lambda \) that satisfy (10.16) for the volume function in Preview Activity 10.8, we calculate both \( \nabla f \) and \( \nabla g \). Observe that

\[
\nabla f = 2xy\mathbf{i} + x^2\mathbf{j} \quad \text{and} \quad \nabla g = 4\mathbf{i} + \mathbf{j},
\]

and thus we need a value of \( \lambda \) so that

\[
2xy\mathbf{i} + x^2\mathbf{j} = \lambda (4\mathbf{i} + \mathbf{j}).
\]

Equating components in the most recent equation and incorporating the original constraint, we have three equations

\[
\begin{align*}
2xy &= \lambda (4) \quad (10.17) \\
x^2 &= \lambda (1) \quad (10.18) \\
4x + y &= 108 \quad (10.19)
\end{align*}
\]

in the three unknowns \( x, y, \) and \( \lambda \). First, note that if \( \lambda = 0 \), then equation (10.18) shows that \( x = 0 \). From this, Equation (10.19) tells us that \( y = 108 \). So the point \((0, 108)\) is a point we need to consider. Next, provided that \( \lambda \neq 0 \) (from which it follows that \( x \neq 0 \) by Equation (10.18)), we may divide both sides of Equation (10.17) by the corresponding sides of (10.18) to eliminate \( \lambda \), and thus find that

\[
\frac{2y}{x} = 4, \quad \text{so} \quad y = 2x.
\]
Substituting into Equation (10.19) gives us
\[ 4x + 2x = 108 \]
or
\[ x = 18. \]
Thus we have \( y = 2x = 36 \) and \( \lambda = x^2 = 324 \) as another point to consider. So the points at which the gradients of \( f \) and \( g \) are parallel, and thus at which \( f \) may have a maximum or minimum subject to the constraint, are \((0, 108)\) and \((18, 36)\). By evaluating the function \( f \) at these points, we see that we maximize the volume when the length of the square end of the box is 18 inches and the length is 36 inches, for a maximum volume of \( f(18, 36) = 11664 \) cubic inches. Since \( f(0, 108) = 0 \), we obtain a minimum value at this point.

We summarize the process of Lagrange multipliers as follows.

The general technique for optimizing a function \( f = f(x, y) \) subject to a constraint \( g(x, y) = c \) is to solve the system \( \nabla f = \lambda \nabla g \) for \( x, y, \) and \( \lambda \). We then evaluate the function \( f \) at each point \((x, y)\) that results from a solution to the system in order to find the optimum values of \( f \) subject to the constraint.

Activity 10.25.

A cylindrical soda can holds about 355 cc of liquid. In this activity, we want to find the dimensions of such a can that will minimize the surface area.

(a) What are the variables in this problem? What restriction(s), if any, are there on these variables?

(b) What quantity do we want to optimize in this problem? What equation describes the constraint?

(c) Find \( \lambda \) and the values of your variables that satisfy Equation (10.16) in the context of this problem.

(d) Determine the dimensions of the pop can that give the desired solution to this constrained optimization problem.

The method of Lagrange multipliers also works for functions of more than two variables.


Use the method of Lagrange multipliers to find the dimensions of the least expensive packing crate with a volume of 240 cubic feet when the material for the top costs $2 per square foot, the bottom is $3 per square foot and the sides are $1.50 per square foot.
Summary

- The extrema of a function \( f = f(x, y) \) subject to a constraint \( g(x, y) = c \) occur at points for which the contour of \( f \) is tangent to the curve that represents the constraint equation. This occurs when

\[
\nabla f = \lambda \nabla g.
\]

- We use the condition \( \nabla f = \lambda \nabla g \) to generate a system of equations, together with the constraint \( g(x, y) = c \), that may be solved for \( x, y \), and \( \lambda \). Once we have all the solutions, we evaluate \( f \) at each of the \((x, y)\) points to determine the extrema.

Exercises

1. The Cobb-Douglas production function is used in economics to model production levels based on labor and equipment. Suppose we have a specific Cobb-Douglas function of the form

\[
f(x, y) = 50x^{0.4}y^{0.6},
\]

where \( x \) is the dollar amount spent on labor and \( y \) the dollar amount spent on equipment. Use the method of Lagrange multipliers to determine how much should be spent on labor and how much on equipment to maximize productivity if we have a total of \$1.5 million dollars to invest in labor and equipment.

2. Use the method of Lagrange multipliers to find the point on the line \( x - 2y = 5 \) that is closest to the point \((1, 3)\). To do so, respond to the following prompts.

   (a) Write the function \( f = f(x, y) \) that measures the square of the distance from \((x, y)\) to \((1, 3)\). (The extrema of this function are the same as the extrema of the distance function, but \( f(x, y) \) is simpler to work with.)

   (b) What is the constraint \( g(x, y) = c \)?

   (c) Write the equations resulting from \( \nabla f = \lambda \nabla g \) and the constraint. Find all the points \((x, y)\) satisfying these equations.

   (d) Test all the points you found to determine the extrema.

3. Apply the Method of Lagrange Multipliers solve each of the following constrained optimization problems.

   (a) Determine the absolute maximum and absolute minimum values of \( f(x, y) = (x - 1)^2 + (y - 2)^2 \) subject to the constraint that \( x^2 + y^2 = 16 \).

   (b) Determine the points on the sphere \( x^2 + y^2 + z^2 = 4 \) that are closest to and farthest from the point \((3, 1, -1)\). (As in the preceding exercise, you may find it simpler to work with the square of the distance formula, rather than the distance formula itself.)
(c) Find the absolute maximum and minimum of \( f(x, y, z) = x^2 + y^2 + z^2 \) subject to the constraint that \((x - 3)^2 + (y + 2)^2 + (z - 5)^2 \leq 16\). (Hint: here the constraint is a closed, bounded region. Use the boundary of that region for applying Lagrange Multipliers, but don’t forget to also test any critical values of the function that lie in the interior of the region.)

4. There is a useful interpretation of the Lagrange multiplier \( \lambda \). Assume that we want to optimize a function \( f \) with constraint \( g(x, y) = c \). Recall that an optimal solution occurs at a point \((x_0, y_0)\) where \( \nabla f = \lambda \nabla g \). As the constraint changes, so does the point at which the optimal solution occurs. So we can think of the optimal point as a function of the parameter \( c \), that is \( x_0 = x_0(c) \) and \( y_0 = y_0(c) \). The optimal value of \( f \) subject to the constraint can then be considered as a function of \( c \) defined by \( f(x_0(c), y_0(c)) \). The Chain Rule shows that

\[
\frac{df}{dc} = \frac{\partial f}{\partial x_0} \frac{dx_0}{dc} + \frac{\partial f}{\partial y_0} \frac{dy_0}{dc}.
\]

(a) Use the fact that \( \nabla f = \lambda \nabla g \) at \((x_0, y_0)\) to explain why

\[
\frac{df}{dc} = \lambda \frac{dg}{dc}.
\]

(b) Use the fact that \( g(x, y) = c \) to show that

\[
\frac{df}{dc} = \lambda.
\]

Conclude that \( \lambda \) tells us the rate of change of the function \( f \) as the parameter \( c \) increases (or by approximately how much the optimal value of the function \( f \) will change if we increase the value of \( c \) by 1 unit).

(c) Suppose that \( \lambda = 324 \) at the point where the package described in Preview Activity 10.8 has its maximum volume. Explain in context what the value 324 tells us about the package.