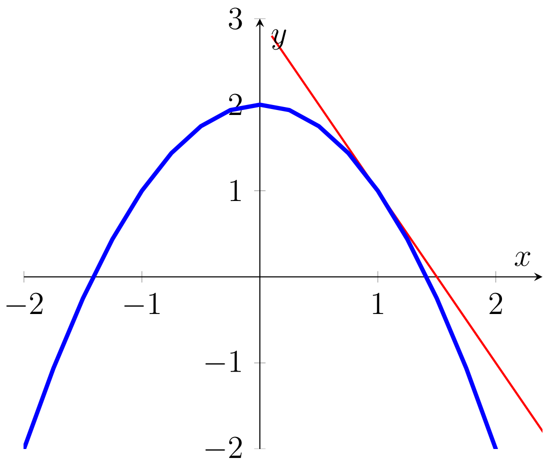
5.4 Directional Derivatives and the Gradient Vector

Objectives

- ☐ I understand the notion of a gradient vector and I know in what direction it points.
- ☐ I know how to calculate a directional derivative in the direction of a given vector $\vec{v}$.
- ☐ I know how to find the unit vector $\vec{u}$ that creates the maximum directional derivative, $D_u f$.

Mathematically speaking, the derivative of a function $f(x)$ at a point x should be thought of as the slope of f at that point x. If we want to be really precise, we should think of it as *the slope of the tangent line of f at x.* This is the best way to conceptually understand a derivative.

For example, if you're given the function $f(x) = -x^2 + 2$, you know the derivative of f at $x = 1$ will be -2. This value corresponds to the slope of the line tangent to f at $x = 1$.



For higher dimensions, we want to find an analogous value. That is, we want to find something that can represent a slope of a tangent. In the last section, we found partial derivatives, but as the word "partial" would suggest, we are not done! These partial derivatives are an intermediate step to the object we wish to find.

Recall that slopes in three dimensions are described with vectors (see section 3.5 Lines and Planes) because *vectors describe movement*. So our true derivative in higher dimensions should be a vector. This vector is called the gradient vector.

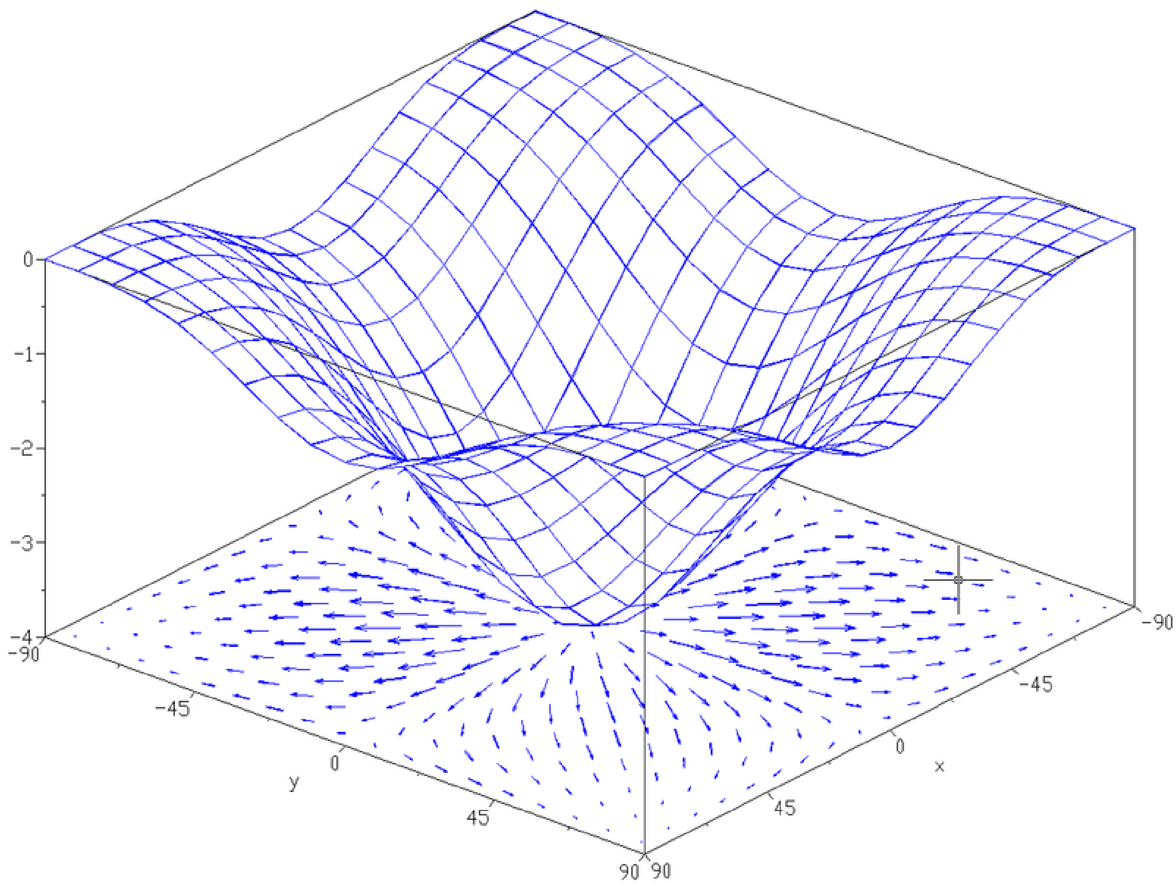
Definition 5.4.1 *The* **gradient vector** *of a function f, denoted ∇f or grad(f), is a vectors whose entries are the partial derivatives of f. That is,*

$$\nabla f(x, y) = \langle f_x(x, y), f_y(x, y) \rangle$$

This is a generalization of a derivative of a function of several variables.

A natural question you may ask is in what direction does the vector point? The answer is that it points in the direction of **steepest ascent**. If you imagine walking on a hilly area, the gradient is a vector that points you toward the steepest climb.

Here is a picture of a three dimensional surface. The arrows at the bottom represent the gradient vectors. They each point in the direction where it is steepest. Longer the vector (the greater its magnitude), the steeper the surface is at that point.



Here is a good thought exercise to test your understanding of gradients. What do you think the gradient vector should be for the function $f(x, y) = x^2 + y^2$ at the point $(0, 0)$?Remember, this function is the paraboloid and $(0, 0)$ is its vertex.

5.4.1 Examples

Example 5.4.1.1 *Find the gradient vector of*

$$f(x,y) = x^2 + y^2$$

What are the gradient vectors at $(1,2),(2,1)$ *and* $(0,0)$*?*

We begin with the formula.

$$\nabla f = \langle f_x, f_y \rangle = \boxed{\langle 2x, 2y \rangle}$$

Now, let us find the gradient at the following points.

- $\nabla f(1,2) = \langle 2, 4 \rangle$
- $\nabla f(2,1) = \langle 4, 2 \rangle$
- $\nabla f(0,0) = \langle 0, 0 \rangle$

Notice that at $(0,0)$ the gradient vector is the zero vector. Since the gradient corresponds to the notion of slope at that point, this is the same as saying the slope is zero.

Example 5.4.1.2 *Find the gradient vector of*

$$f(x,y) = 2xy + x^2 + y$$

What are the gradient vectors at $(1,1),(0,-1)$ *and* $(0,0)$*?*

$$\nabla f = \langle f_x, f_y \rangle = \boxed{\langle 2y + 2x, 2x + 1 \rangle}$$

Now, let us find the gradient at the following points.

- $\nabla f(1,1) = \langle 4, 3 \rangle$
- $\nabla f(0,-1) = \langle -2, 1 \rangle$
- $\nabla f(0,0) = \langle 0, 1 \rangle$

So far, we've learned the definition of the gradient vector and we know that it tells us the direction of steepest ascent. What if, however, we want to know the rate of ascent in another direction? For that, we use something called a directional derivative.

Definition 5.4.2 *The* **directional derivative***, denoted $D_v f(x,y)$, is a derivative of a multivariable function in the direction of a vector $\vec{v}$. It is the scalar projection of the gradient onto $\vec{v}$.*

$$D_v f(x,y) = comp_v \nabla f(x,y) = \frac{\nabla f(x,y) \cdot \vec{v}}{|\vec{v}|}$$

This produces a vector whose magnitude represents the rate a function ascends (how steep it is) at point (x,y) in the direction of $\vec{v}$.

If our function has three inputs, the math works out the same. Suppose $f(x,y,z) = w$. Then,

$$\nabla f(x,y,z) = \langle f_x(x,y,z), f_y(x,y,z), f_z(x,y,z) \rangle$$

and the directional derivative in the direction of $\vec{u}$ is

$$D_v f(x,y,z) = \frac{\nabla f(x,y,z) \cdot \vec{v}}{|\vec{v}|}$$

Let's look at some examples.

5.4.2 Examples

Example 5.4.2.1 *Find the directional derivative of*

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

in the direction of $\vec{v} = \langle 3, 5 \rangle$ at the point $(1,2)$.

First, we find the gradient.

$$f_x(x,y) = \frac{d}{dx}\left(\frac{x}{x^2+y^2}\right)$$

$$= \frac{(x^2+y^2) - x(2x)}{(x^2+y^2)^2}$$

$$= \frac{y^2 - x^2}{(x^2+y^2)^2}$$

$$f_y(x,y) = \frac{d}{dy}\left(\frac{x}{x^2+y^2}\right)$$

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

The gradient is then

$$\nabla f(1,2) = \left\langle \frac{4-1}{(4+1)^2}, \frac{-4}{(4+1)^2} \right\rangle = \left\langle \frac{3}{25}, \frac{-4}{25} \right\rangle = \frac{1}{25}\langle 3, -4\rangle$$

We now find the magnitude of $\vec{v}$. We get

$$|\vec{v}| = \sqrt{9+25} = \sqrt{34}$$

The directional derivative is then

$$D_v f(1,2) = \frac{\nabla f(1,2)\cdot\vec{v}}{|\vec{v}|} = \frac{1}{25\sqrt{34}}\langle 3,-4\rangle\cdot\langle 3,5\rangle = \frac{1}{25\sqrt{34}}(9-20) = \boxed{-\frac{11}{25\sqrt{34}}}$$

Example 5.4.2.2 *Find the directional derivative of*

$$f(x,y,z) = \sqrt{xyz}$$

in the direction of $\vec{v} = \langle -1, -2, 2\rangle$ *at the point* $(3,2,6)$.

First, we find the partial derivatives to define the gradient.

$$f_x(x,y,z) = \frac{yz}{2\sqrt{xyz}}$$
$$f_y(x,y,z) = \frac{xz}{2\sqrt{xyz}}$$
$$f_z(x,y,z) = \frac{xy}{2\sqrt{xyz}}$$

The gradient is

$$\nabla f(3,2,6) = \left\langle \frac{12}{2(6)}, \frac{18}{2(6)}, \frac{6}{2(6)} \right\rangle = \left\langle 1, \frac{3}{2}, \frac{1}{2} \right\rangle = \frac{1}{2}\langle 2,3,1\rangle$$

The magnitude of $\vec{v} = \langle -1, -2, 2 \rangle$ is

$$|\vec{v}| = \sqrt{1+4+4} = 3$$

Therefore, the directional derivative is

$$D_v f(3,2,6) = \frac{\nabla f(3,2,6) \cdot \vec{v}}{|\vec{v}|} = \frac{1}{3(2)} \langle 2,3,1 \rangle \cdot \langle -1,-2,2 \rangle = \frac{1}{6}(-2-6+2) = \boxed{-1}$$

The next natural question is:

In what direction is the derivative maximum?

As we just saw, the directional derivative is calculated by taking the scalar projection of ∇f onto a vector $\vec{v}$. Define θ be the angle between $\vec{v}$ and ∇f. Then,

$$\frac{\nabla f \cdot \vec{v}}{|\vec{v}|} = \frac{|\nabla f||\vec{u}|\cos(\theta)}{|\vec{v}|} = |\nabla f|\cos(\theta)$$

This is maximized if $\theta = 0$. From this, we know the following:

- The maximum rate of change (the largest directional derivative) is $|\nabla f|$.
- This occurs when $\vec{v}$ is *parallel* to ∇f, i.e. when they point in the same direction.

That makes sense since ∇f is the vector pointing toward *steepest ascent*, so it should be the direction with the largest derivative.

You'll typically be asked for the unit vector, $\vec{u}$, that creates the maximum directional derivative. This is because unit vectors are thought of as vectors that just contain information about direction. Based on our discussion above, this will always be

$$\vec{u} = \frac{\nabla f}{|\nabla f|}$$

Let's look at two examples.

5.4.3 Examples

Example 5.4.3.1 1. *Find the maximum rate of change of f at the given point and the direction in which it occurs.*

$$f(s,t) = te^{st}, \quad (0,2)$$

The maximum rate of change is $|\nabla f(0,2)|$. Let's first find the gradient.

$$\nabla f = \left\langle te^{st}, ste^{st} + e^{st} \right\rangle$$

Then

$$|\nabla f(0,2)| = \sqrt{(2)^2 + 1^2} = \boxed{\sqrt{5}}$$

The direction is the unit vector.

$$\frac{\nabla f(0,2)}{|\nabla f(0,2)|} = \boxed{\left\langle \frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}} \right\rangle}$$

Note: for this problem, it may not have been clear which component was the first and which was the second since s and t are atypical variables. For clues about the order, look out how the ordered pairs are defined in the function. It was written as "$f(s,t)$," which tells us our vectors are $\langle s, t \rangle$.

Example 5.4.3.2 2. *Find the maximum rate of change of f at the given point and the direction in which it occurs.*

$$f(x,y,z) = \sqrt{x^2 + y^2 + z^2}, \quad (3, 6, -2)$$

As above, the maximum rate of change is $|\nabla f(3,6,-2)|$.

$$\nabla f(x,y,z) = \left\langle \frac{x}{\sqrt{x^2+y^2+z^2}}, \frac{y}{\sqrt{x^2+y^2+z^2}}, \frac{z}{\sqrt{x^2+y^2+z^2}} \right\rangle$$

Then

$$\nabla f(3,6,-2) = \left\langle \frac{3}{7}, \frac{6}{7}, -\frac{2}{7} \right\rangle$$

The $|\nabla f(3,6,-2)| = 1/7\sqrt{9+36+4} = \boxed{1}$
Since it's already a unit vector, the direction is

$$\boxed{\left\langle \frac{3}{7}, \frac{6}{7}, -\frac{2}{7} \right\rangle}$$

Summary of Ideas: Directional Derivatives and the Gradient Vector

- The **gradient vector** of a function f, denoted ∇f or $grad(f)$, is a vectors whose entries are the partial derivatives of f.

$$\nabla f(x,y) = \langle f_x(x,y), f_y(x,y) \rangle$$

  It is the generalization of a derivative in higher dimensions.

- The gradient points in the direction of **steepest ascent**.

- The **directional derivative**, denoted $D_v f(x,y)$, is a derivative of a $f(x,y)$ in the direction of a vector $\vec{v}$. It is the scalar projection of the gradient onto $\vec{v}$.

$$D_v f(x,y) = \text{comp}_v \nabla f(x,y) = \frac{\nabla f(x,y) \cdot \vec{v}}{|\vec{v}|}$$

  This produces a vector whose magnitude represents the rate a function ascends (how steep it is) at point (x,y) in the direction of $\vec{v}$.

- Both the gradient and the directional derivative work the same in higher variables.

- The **maximum directional derivative** is always $|\nabla f|$.

- This happens in the direction of the unit vector

$$\vec{u} = \frac{\nabla f}{|\nabla f|}$$

  Remember, we use the unit vector as a convention. Any vector parallel to ∇f will work.