# The Gradient and Level Sets

We've defined the directional derivatives of a function, which allow us to determine how a function is changing in various directions.

**Definition 1.** *Consider a function $f : \mathbb{R}^n \to \mathbb{R}$, a point $\vec{a} \in \mathbb{R}^n$, and a direction given by a unit vector $\vec{v} \in \mathbb{R}^n$. Then we define the* directional derivative *of $f$ at $\vec{a}$ in the direction of $\vec{v}$ to be*

$$D_{\vec{v}} f(\vec{a}) = \lim_{h \to 0} \frac{f(\vec{a} + h\vec{v}) - f(\vec{a})}{h},$$

*provided this limit exists.*

However, we would like an easier way to evaluate directional derivatives, that doesn't require the limit definition.

Such a method will require use of the gradient of the function. Recall our definition of the gradient of a function $f : \mathbb{R}^n \to \mathbb{R}$.

**Definition 2.** *Consider a function $f : \mathbb{R}^n \to \mathbb{R}$. The* gradient *of $f$ is the function $\nabla f : \mathbb{R}^n \to \mathbb{R}^n$, defined by*

$$\nabla f = \left( \frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, ..., \frac{\partial f}{\partial x_n} \right).$$

We'll see that this vector turns out to be closely related to directional derivatives.

## The gradient and directional derivatives

Let's suppose we have a differentiable function $f : \mathbb{R}^n \to \mathbb{R}$, and consider our definition of the directional derivative a function $f$ at $\vec{a}$ in the direction of a unit vector $\vec{v}$, which was

$$D_{\vec{v}} f(\vec{a}) = \lim_{h \to 0} \frac{f(\vec{a} + h\vec{v}) - f(\vec{a})}{h}.$$

We'll rewrite this definition by considering another function, $F(h) = f(\vec{a} + h\vec{v})$. Notice that $F$ is a single variable function, and when we rewrite the directional

Learning outcomes:
Author(s):

derivative, we have

$$\begin{aligned} D_{\vec{v}} f(\vec{a}) &= \lim_{h \to 0} \frac{f(\vec{a} + h\vec{v}) - f(\vec{a})}{h} \\ &= \lim_{h \to 0} \frac{F(h) - F(0)}{h - 0} \\ &= F'(0). \end{aligned}$$

So, the directional derivative is just the derivative of this single variable function $F(h)$ evaluated at 0. Revisiting our definition of $F(h)$, we can use the chain rule to find the derivative of $F$.

$$\begin{aligned} \frac{d}{dh} F(h) &= \nabla f(\vec{a} + h\vec{v}) \cdot \frac{d}{dh}(\vec{a} + h\vec{v}) \\ &= \nabla f(\vec{a} + h\vec{v}) \cdot \vec{v} \end{aligned}$$

Evaluating at $h = 0$, we have

$$\begin{aligned} D_{\vec{v}} f(\vec{a}) &= F'(0) \\ &= \nabla f(\vec{a}) \cdot \vec{v}. \end{aligned}$$

Thus, we have arrived at the following result.

**Theorem 1.** *Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable at $\vec{a} \in \mathbb{R}^n$. Then $D_{\vec{v}} f(\vec{a})$ exists for all unit vectors $\vec{v} \in \mathbb{R}^n$, and*

$$D_{\vec{v}} f(\vec{a}) = \nabla f(\vec{a}) \cdot \vec{v}.$$

Let's use this result to compute some directional derivatives.

**Example 1.** *We'll compute the directional derivative of $f(x, y) = x^2 y + y^2$ at $\vec{a} = (2, 0)$, in the direction of $(3, 4)$. (We previously computed this directional derivative using the limit definition.)*

*Since $(3, 4)$ isn't a unit vector, we need to normalize it. Since $\|(3, 4)\| = \sqrt{3^2 + 4^2} = 5$, we'll use the vector $\vec{v} = \left( \frac{3}{5}, \frac{4}{5} \right)$ to compute our desired directional derivative.*

*Next, we'll need the gradient of $f$.*

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

*Since the partial derivatives of $f$ are polynomials, they are continuous, so $f$ is differentiable. Thus, we can use the above theorem to compute the directional derivative.*

*Then, we can compute the directional derivative as*

$$\begin{aligned} D_{\vec{v}} f(\vec{a}) &= \nabla f(\vec{a}) \cdot \vec{v} \\ &= \nabla f(2,0) \cdot \left(\frac{3}{5}, \frac{4}{5}\right) \\ &= (0,4) \cdot \left(\frac{3}{5}, \frac{4}{5}\right) \\ &= \frac{16}{5}. \end{aligned}$$

*This matches what we had previously computed using the definition of directional derivatives.*

**Problem 1** *Compute the directional derivative of $f(x,y,z) = 3xy + xz^2$ at $\vec{a} = (2,0,1)$, in the direction of $(2,2,1)$.*

$$D_{\vec{v}} f(\vec{a}) = \boxed{6}$$

*Compute the directional derivative of $f(x,y,z) = 3xy + xz^2$ at $\vec{a} = (2,0,1)$, in the direction of $(-2,1,-1)$.*

$$D_{\vec{v}} f(\vec{a}) = \boxed{0}$$

# The gradient and level sets

We've shown that for a differentiable function $f : \mathbb{R}^n \to \mathbb{R}$, we can compute directional derivatives as

$$D_{\vec{v}} f(\vec{a}) = \nabla f(\vec{a}) \cdot \vec{v}.$$

What does this mean for the possible values for a directional derivative? Recall that the dot product can be computed as

$$\nabla f(\vec{a}) \cdot \vec{v} = \|\nabla f(\vec{a})\| \, \|\vec{v}\| \cos\theta,$$

where $\theta$ is the angle between the two vectors. Since $\vec{v}$ is a unit vector, we have

$$\|\nabla f(\vec{a})\| \, \|\vec{v}\| \cos\theta = \|\nabla f(\vec{a})\| \cos\theta.$$

Since $-1 \le \cos\theta \le 1$, we have that

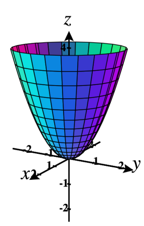
$$-\|\nabla f(\vec{a})\| \le D_{\vec{v}} f(\vec{a}) \le \|\nabla f(\vec{a})\|.$$

In particular, the largest that $D_{\vec{v}}f(\vec{a})$ can be is $\|\nabla f(\vec{a})\|$, and this occurs when $\vec{v}$ points in the same direction as $\nabla f(\vec{a})$, so that $\theta = 0$. Thus, the gradient points in the direction of greatest increase.

On the other hand, the minimum value that $D_{\vec{v}}f(\vec{a})$ can have is $-\|\nabla f(\vec{a})\|$, and this occurs when $\vec{v}$ points in the opposite direction from $\nabla f(\vec{a})$, in the direction of $-\nabla f(\vec{a})$. Thus, $-\nabla f(\vec{a})$ points in the direction of greatest decrease.

Additionally, from $D_{\vec{v}}f(\vec{a}) = \nabla f(\vec{a}) \cdot \vec{v}$, we can see that $\vec{v}$ is perpendicular to $\nabla f(\vec{a})$ if and only if $D_{\vec{v}}f(\vec{a}) = 0$. But what does it mean for $D_{\vec{v}}f(\vec{a})$? This means that there is no instantaneous change in $f$ in the direction of $\vec{v}$, which means that $\vec{v}$ will be a tangent vector to a level curve.

**Example 2.** *Consider the graph of the function $f(x, y) = x^2 + y^2$, which is a paraboloid.*



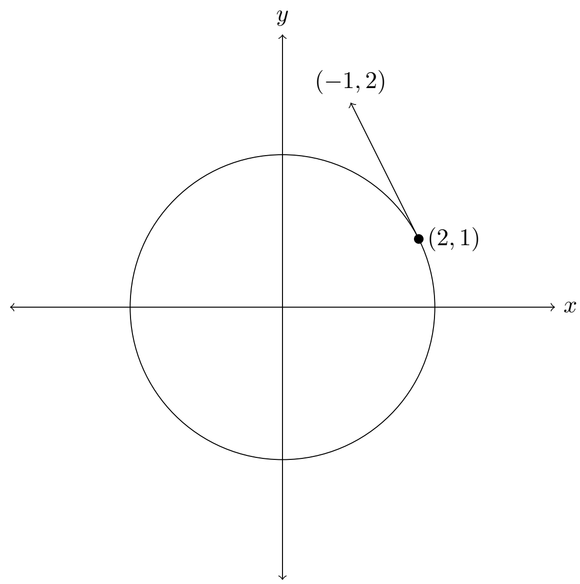

*Let's consider how this function changes near the point $(2, 1)$.*

*The gradient of $f$ at $(2, 1)$ is*

$$\nabla f(2, 1) = \boxed{(4, 2)}.$$

*Consulting the graph of $f$ near the point $(2, 1)$, we can confirm that it increases most rapidly when we move in the direction of $\nabla f(2, 1)$. We can also confirm that it decreases most rapidly when we move in the direction of $-\nabla f(2, 1)$.*

*The point $(2, 1)$ lies on the level curve $x^2 + y^2 = 5$. A tangent vector to this level curve at the point $(2, 1)$ is given by $(-1, 2)$.*




*This vector, $(-1, 2)$, is perpendicular to our gradient $\nabla f(2, 1)$.*

We'll state this observation more formally, and prove that the gradient is perpendicular to the level curves.

**Theorem 2.** *Consider a function $f : \mathbb{R}^n \to \mathbb{R}$, and suppose $f$ is of class $\mathcal{C}^1$. For some constant $c$, consider the level set*

$$S = \{\vec{x} \in \mathbb{R}^n \ : \ f(\vec{x}) = c\}.$$

*Then, for any point $\vec{x}_0$ in $S$, the gradient $\nabla f(\vec{x}_0)$ is perpendicular to $S$.*

***Proof*** We need to show that for any vector $\vec{a}$ which is tangent to $S$ at $\vec{x}_0$, we have that $\vec{a}$ is perpendicular to $\nabla f(\vec{x}_0)$.

If $\vec{a}$ is tangent to $S$, we can find a parametrized curve $\vec{x}(t)$ lying in $S$ such that $\vec{x}_0 = \vec{x}(t_0)$ and $\vec{x}'(t_0) = \vec{a}$. We will show that $\nabla f(\vec{x}_0)$ is perpendicular to $\vec{a} = \vec{x}'(t_0)$.

By the definition of $S$, and since $\vec{x}(t)$ lies in $S$,

$$f(\vec{x}(t)) = c$$

for all $t$. Differentiating both sides of this identity, and using the chain rule on the left side, we obtain

$$\nabla f(\vec{x}(t)) \cdot \vec{x}'(t) = 0.$$

Plugging in $t = t_0$, this gives us

$$\nabla f(\vec{x}(t_0)) \cdot \vec{x}'(t_0) = 0,$$

which we can rewrite as

$$\nabla f(\vec{x}_0) \cdot \vec{x}'(t_0) = 0.$$

Thus, we have shown that $\nabla f(\vec{x}_0)$ is perpendicular to the level set $S$. ■