The Gradient

We've given a formal definition for differentiability of a function $f : \mathbb{R}^2 \to \mathbb{R}$,

Definition 1. *Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$, and suppose that the partial derivatives f_x and f_y are defined at the point $(x, y) = (a, b)$. Define the linear function*

$$h(x, y) = f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b).$$

We say that f is differentiable *at $(x, y) = (a, b)$ if*

$$\lim_{(x,y)\to(a,b)} \frac{f(x, y) - h(x, y)}{\|(x, y) - (a, b)\|} = 0.$$

If either of the partial derivatives $f_x(a, b)$ and $f_y(a, b)$ do not exist, or the above limit does not exist or is not 0, then f is not differentiable at (a, b).

The idea behind this definition is that $h(x, y)$ will be a "good" linear approximation to $f(x, y)$ near (a, b) if f is differentiable at (a, b).

We would now like to define differentiability for scalar-valued functions of more than two variables, so functions from $\mathbb{R}^n$ to $\mathbb{R}$. This definition will closely resemble our definition above, which handles the case $n = 2$. For example, in the case $n = 3$, we will use the linear function

$$h(x, y, z) = f(a, b, c) + f_x(a, b, c)(x - a) + f_y(a, b, c)(y - b) + f_z(a, b, c)(z - c).$$

For larger n, we'll define a similar function h, but this notation will quickly become unwieldy! In order to simplify notation, we'll now introduce a new object to organize our partial derivatives: the gradient of a scalar-valued function.

The gradient

In order to organize our information about partial derivatives, and streamline our definition of differentiability for functions $\mathbb{R}^n \to \mathbb{R}$, we now define the gradient of a scalar-valued function.

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*

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

Learning outcomes:
Author(s):

The gradient vector will be a useful computation tool in general, not only for defining differentiability.

Example 1. *For $f(x,y,z) = x^2 + ye^z$, we can compute the partial derivatives*

$$\begin{aligned} f_x(x,y,z) &= 2x, \\ f_y(x,y,z) &= e^z, \\ f_z(x,y,z) &= ye^z. \end{aligned}$$

Then the gradient of f is

$$\nabla f = (2x, e^z ye^z).$$

Problem 1 *Find the gradient of each function.*

$f(x,y,z) = \sin(xyz)$

$$\nabla f(x,y,z) = \boxed{(yz\cos(xyz), xz\cos(xyz), xy\cos(xyz))}$$

$g(x,y) = x^2e^y + y$

$$\nabla g(x,y) = \boxed{(2xe^y, x^2e^y + 1)}$$

$h(x_1,x_2,x_3,x_4) = x_1^2x_2 + x_1x_3 + x_2x_4$

$$\nabla h(x_1,x_2,x_3,x_4) = \boxed{(2x_1x_2 + x_3, x_1^2 + x_4, x_1, x_2)}$$

Differentiability

Now that we've defined the gradient, let's revisit our definition of differentiability for a function from $\mathbb{R}^2$ to $\mathbb{R}$. We used the function

$$h(x,y) = f(a,b) + f_x(a,b)(x-a) + f_y(a,b)(y-b).$$

Looking at the terms $f_x(a,b)(x-a) + f_y(a,b)(y-b)$, we can rewrite this as a dot product of two vectors:

$$f_x(a,b)(x-a) + f_y(a,b)(y-b) = (f_x(a,b), f_y(a,b)) \cdot (x-a, y-b).$$

The first vector is the gradient of f evaluated at (a,b), so we can rewrite this as

$$(f_x(a,b), f_y(a,b)) \cdot (x-a, y-b) = \nabla f(a,b) \cdot (x-a, y-b).$$

If we take $\vec{x} = (x,y)$ and $\vec{a} = (a,b)$, we can write this as

$$\nabla f(\vec{a}) \cdot (\vec{x} - \vec{a}).$$

With these notational changes in mind, we now define differentiability for a function $f : \mathbb{R}^n \to \mathbb{R}$.

Definition 3. *Consider a function $f : \mathbb{R}^n \to \mathbb{R}$. For a point $\vec{a} \in \mathbb{R}^n$, define*

$$h(\vec{x}) = f(\vec{a}) + \nabla f(\vec{a}) \cdot (\vec{x} - \vec{a}).$$

We say that f is differentiable *at $\vec{a}$ if*

$$\lim_{\vec{x} \to \vec{a}} \frac{f(\vec{x}) - h(\vec{x})}{\|\vec{x} - \vec{a}\|} = 0.$$

If f is differentiable, we say that $h(\vec{x})$ is the tangent hyperplane *to f at $\vec{a}$.*

If any of the partial derivatives of f do not exist, or the above limit does not exist or is not 0, then f is not differentiable at $\vec{a}$.

Example 2. *We'll use this definition of differentiability to prove that the function $f(x, y, z) = xy + z$ is differentiable at $(1, 1, 1)$.*

First, we find the gradient of f.

$$\nabla f(x, y, z) = \boxed{(y, x, 1)}$$

At the point $(1, 1, 1)$, we have

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

From this, we find the formula for $h(x, y, z)$.

$$\begin{aligned} h(x, y, z) &= f(1, 1, 1) + \nabla f(1, 1, 1) \cdot ((x, y, z) - (1, 1, 1)) \\ &= 2 + (1, 1, 1) \cdot (x - 1, y - 1, z - 1) \\ &= \boxed{x + y + z - 1} \end{aligned}$$

Next, we evaluate the limit

$$\begin{aligned} \lim_{(x,y,z) \to (1,1,1)} \frac{f(x, y, z) - h(x, y, z)}{\|(x, y, z) - (1, 1, 1)\|} &= \lim_{(x,y,z) \to (1,1,1)} \frac{(xy + z) - (x + y + z - 1)}{\sqrt{(x-1)^2 + (y-1)^2 + (z-1)^2}} \\ &= \lim_{(x,y,z) \to (1,1,1)} \frac{xy - x - y + 1}{\sqrt{(x-1)^2 + (y-1)^2 + (z-1)^2}}. \end{aligned}$$

To evaluate this limit, we switch to translated spherical coordinates

$$\begin{aligned} x &= 1 + \rho \cos\theta \sin\phi, \\ y &= 1 + \rho \sin\theta \sin\phi, \\ z &= 1 + \rho \cos\phi. \end{aligned}$$

Making this change, we obtain

$$\begin{aligned} \lim_{(x,y,z) \to (1,1,1)} \frac{xy - x - y + 1}{\sqrt{(x-1)^2 + (y-1)^2 + (z-1)^2}} &= \lim_{\rho \to 0} \frac{(1 + \rho\cos\theta\sin\phi)(1 + \rho\sin\theta\sin\phi) - (1 + \rho\cos\theta\sin\phi) - (1}{|\rho|} \\ &= \lim_{\rho \to 0} \frac{\rho^2 \cos\theta \sin\theta \sin^2\phi}{|\rho|}. \end{aligned}$$

Since $-|\rho| \leq \dfrac{\rho^2 \cos\theta \sin\theta \sin^2\phi}{|\rho|} \leq |\rho|$, *we use the squeeze theorem to obtain*

$$\lim_{\rho \to 0} \frac{\rho^2 \cos\theta \sin\theta \sin^2\phi}{|\rho|} = 0.$$

Thus, we have shown that $f(x, y, z) = xy + z$ *is differentiable at* $(1, 1, 1)$.