8. Gradient and Directional Derivative

In this lecture, we will discuss

- A Review of Gradient of a Function of Two and Three Variables
- Directional Derivative
    - Coordinate Description of the Directional Derivative
    - Maximum Rate of Change of a Function

We will mainly focus on the gradient of a function $f : \mathbb{R}^m \to \mathbb{R}$, where $m = 2$ or 3. Recall the definition:

Definition Gradient of a Function of Two and Three Variables
Let $f : U \subseteq \mathbb{R}^2 \to \mathbb{R}$ be a differentiable function. The gradient of f is the vector field ∇f whose value at a point (x, y) in U is given by

$$\nabla f(x,y) = \left(\frac{\partial f}{\partial x}(x,y), \frac{\partial f}{\partial y}(x,y)\right).$$

The gradient of a differentiable function $f : U \subseteq \mathbb{R}^3 \to \mathbb{R}$ is the vector field

$$\nabla f(x,y,z) = \left(\frac{\partial f}{\partial x}(x,y,z), \frac{\partial f}{\partial y}(x,y,z), \frac{\partial f}{\partial z}(x,y,z)\right).$$

defined on the subset U of $\mathbb{R}^3$.

Example 0

$f: \mathbb{R}^3 \to \mathbb{R}$

Find the gradient of the funtion $f(x,y,z) = xe^{2y}\sin(6z)$ at the point $(1, 1, \pi)$.

Answer:

$$\frac{\partial f}{\partial x} = e^{2y} \sin 6z$$

$$\frac{\partial f}{\partial y} = \underset{\uparrow (2y)'}{2}\, x e^{2y} \sin 6z$$

$$\frac{\partial f}{\partial z} = \underset{\uparrow (6z)'}{6}\, x e^{2y} \cos 6z$$

$$\nabla f(1,1,\pi) = \left(\frac{\partial f}{\partial x}(1,1,\pi), \frac{\partial f}{\partial y}(1,1,\pi), \frac{\partial f}{\partial z}(1,1,\pi)\right)$$

$$= (0, 0, 6e^2) = 0\cdot\vec{i} + 0\cdot\vec{j} + 6e^2 k$$

Directional Derivative

Let $f : \mathbb{R}^2 \to \mathbb{R}$.

Recall the partial derivative of f at (a, b) with respect to x, denoted by $\frac{\partial f}{\partial x}(a, b)$, is defined as (similar defintion for $\frac{\partial f}{\partial y}(a, b)$)

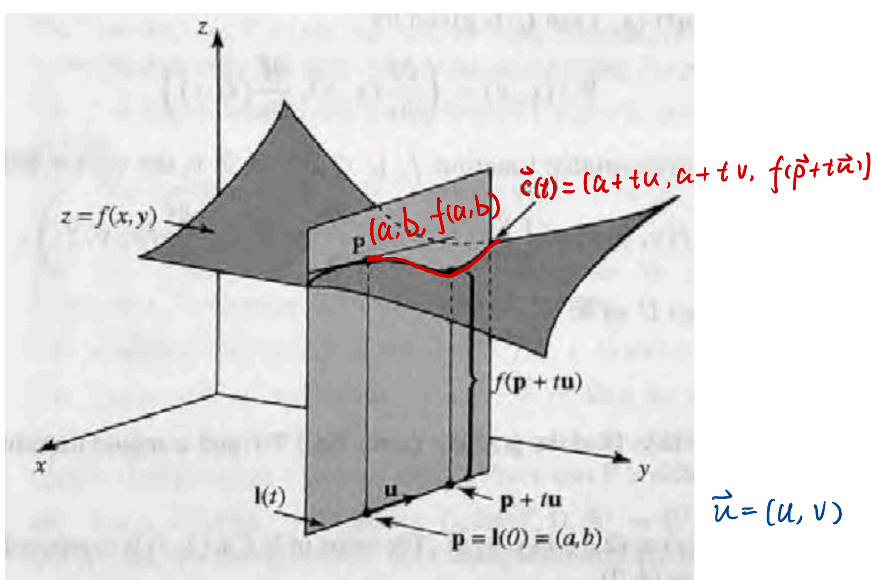
$$\frac{\partial f}{\partial x}(a, b) = \lim_{h \to 0} \frac{f(a+h, b) - f(a, b)}{h}$$

- From the definition, we see that the partial derivative of a function $f(x, y)$ with respect to x or y is the rate of change of $f(x, y)$ in the (positive) x or y direction, respectively.
- 🤔 What about other directions?

  Answer. It turns out that we can find the rate of change in any direction using a more general type of derivative called a *directional derivative.*

Here is a general discussion about how to obtain the definition of the directional derivative:

- Assume that $f(x, y)$ is a function differentiable on an open set $U \subseteq \mathbb{R}^2$, and let $\mathbf{p} = (a, b) \in U$.
- We are going to compute the rate of change of $f(x, y)$ at $\mathbf{p}$ in the direction of a unit vector $\mathbf{u} = (u, v)$.
- Recall $\mathbf{l}(t) = \mathbf{p} + t\mathbf{u} = (a + tu, b + tv)$ represents the line in $\mathbb{R}^2$ that goes through the point $\mathbf{p} = \mathbf{l}(0)$ and whose direction is given by the direction of $\mathbf{u} = (u, v)$.
- Compute the value $f(\mathbf{p} + t\mathbf{u})$ of the function f for each point $\mathbf{p} + t\mathbf{u}$ on $\mathbf{l}(t)$ that is in U.
- The collection of all points $\mathbf{c}(t) = (a + tu, b + tv, f(\mathbf{p} + t\mathbf{u}))$ forms a curve on the surface that is the graph of f.
- Notice that $\mathbf{c}(t)$ belongs to the plane perpendicular to the xy-plane that crosses it along the line $\mathbf{l}(t)$
- Denote the point $\mathbf{c}(0) = (a, b, f(a, b))$ by P, which is a point on the graph of $f(x, y)$
- We can define the directional derivative $D_{\mathbf{u}} f(a, b)$ as the slope of the tangent to $\mathbf{c}(t)$ at P.
- Therefore, $D_{\mathbf{u}} f(a, b)$ describes how f changes in the direction specified by the unit vector $\mathbf{u}$.



Definition Directional Derivative

Let $f : U \subseteq \mathbb{R}^2 \to \mathbb{R}$ be a real-valued differentiable function. The directional derivative of f at the point $\mathbf{p} = (a, b)$ in the direction of the *unit* vector $\mathbf{u} = (u, v)$ is given by

$$D_{\mathbf{u}} f(a,b) = \frac{d}{dt} f(\mathbf{p} + t\mathbf{u})\Big|_{t=0}.$$

Remark. From the limit definition of the derivative, we have

$$D_{\mathbf{u}} f(a,b) = \lim_{t\to 0} \frac{f(\mathbf{p} + t\mathbf{u}) - f(\mathbf{p})}{t} = \lim_{t\to 0} \frac{f((a,b) + t(u,v)) - f(a,b)}{t}.$$

In practice, we often use the following theorem for the computation of directional derivatives.

Theorem 1. Coordinate Description of the Directional Derivative

Let $f : U \subseteq \mathbb{R}^2 \to \mathbb{R}$ be a differentiable function, and $\mathbf{p} = (a, b) \in U$. Then

$$D_{\mathbf{u}} f(a,b) = \nabla f(a,b) \cdot \mathbf{u},$$

where $\mathbf{u} = (u, v)$ is a unit vector in $\mathbb{R}^2$.

Proof. Let $\mathbf{l}(t) = \mathbf{p} + t\mathbf{u} = (a + tu, b + tv)$; then $\mathbf{l}(0) = (a, b) = \mathbf{p}$ and $\mathbf{l}'(0) = \mathbf{u} = (u, v)$. Consider the composition $f(\mathbf{l}(t)) = f(\mathbf{p} + t\mathbf{u})$ of $\mathbf{l}$ and f. By the chain rule,

$$\frac{d}{dt} f(\mathbf{p} + t\mathbf{u}) = \frac{d}{dt} f(\mathbf{l}(t)) = \nabla f(\mathbf{l}(t)) \cdot \mathbf{l}'(t).$$

Consequently,

$$D_{\mathbf{u}} f(a,b) = \frac{d}{dt} f(\mathbf{p} + t\mathbf{u})\Big|_{t=0} = \nabla f(\mathbf{l}(0)) \cdot \mathbf{l}'(0) = \nabla f(a,b) \cdot \mathbf{u}.$$

If f is a function of three variables, then

$$D_{\mathbf{u}} f(a,b,c) = \nabla f(a,b,c) \cdot \mathbf{u},$$

where $\mathbf{u}$ is a unit vector in $\mathbb{R}^3$.

Example 1 Suppose $f(x,y)=\sqrt{\tan(x)+y}$ and $\boldsymbol{u}$ is the unit vector in the direction of $(3,-3)$.

Find the following values:

- $\nabla f(x,y)$
- $\nabla f(0.7,6)$
- $f_u(0.7,6)=D_u f(0.7,6)$

ANS: (1) $\frac{\partial f}{\partial x}=\frac{\partial(\tan x+y)^{\frac{1}{2}}}{\partial x}$ recall $(\tan x)'=\sec^2 x$

$=\frac{1}{2}(\tan x+y)^{-\frac{1}{2}}\cdot\frac{\partial(\tan x+y)}{\partial x}$

$=\frac{\sec^2 x}{2\sqrt{\tan x+y}}$

$\frac{\partial f}{\partial y}=\frac{1}{2\sqrt{\tan x+y}}$

Then $\nabla f(x,y)=\left(\frac{\partial f}{\partial x},\frac{\partial f}{\partial y}\right)=\left(\frac{\sec^2 x}{2\sqrt{\tan x+y}},\ \frac{1}{2\sqrt{\tan x+y}}\right)$

(2) $\nabla f(0.7,6)\approx(0.326758,\ 0.191148)$

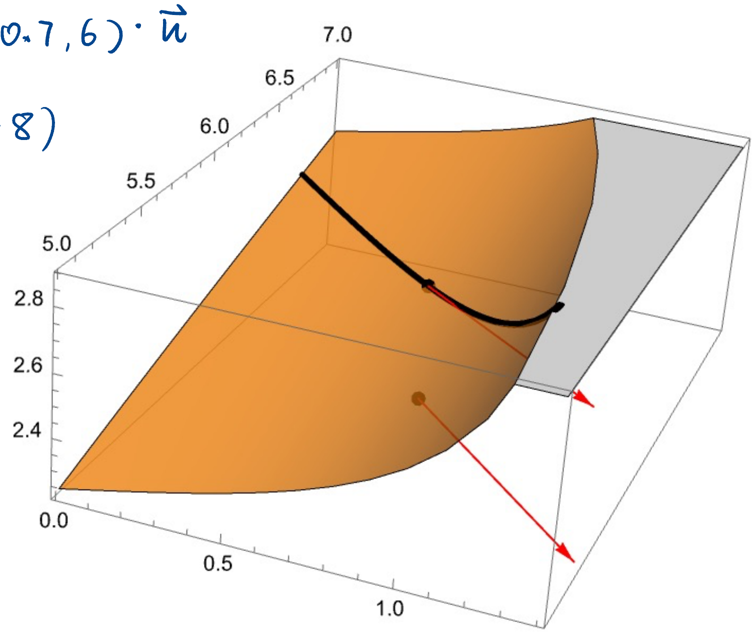
(3) By Thm 1, $D_{\vec{u}}f(a,b)=\nabla f(a,b)\cdot\vec{u}$, where $\vec{u}$ is a unit vector

We first normalize the vector $\vec{v}=(3,-3)$: $\vec{u}=\frac{\vec{v}}{\|\vec{v}\|}=\frac{(3,-3)}{\sqrt{3^2+3^2}}=\left(\frac{1}{\sqrt{2}},-\frac{1}{\sqrt{2}}\right)$

$Df_{\vec{u}}(0.7,6)=\nabla f(0.7,6)\cdot\vec{u}$

$=(0.326758,\ 0.191148)\cdot\left(\frac{1}{\sqrt{2}},-\frac{1}{\sqrt{2}}\right)$

≈ 0.0958906



Example 2

Find the directional derivative of $f(x,y,z) = z^3 - x^2y$ at the point $(-5,2,-3)$ in the direction of the vector $\mathbf{v} = \langle -1,5,3 \rangle$.

ANS:

Normalize $\vec{v} = \langle -1,5,3 \rangle$, we have $\vec{u} = \dfrac{\langle -1,5,3 \rangle}{\sqrt{(-1)^2+5^2+3^2}} = \langle -\frac{1}{\sqrt{35}}, \frac{5}{\sqrt{35}}, \frac{3}{\sqrt{35}} \rangle$

$$\nabla f(x,y,z) = \langle \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \rangle = \langle -2xy, -x^2, 3z^2 \rangle$$

$$\nabla f(-5,2,-3) = \langle 20, -25, 27 \rangle$$

We have

$$\begin{aligned} D_{\vec{u}} f(-5,2,-3) &= \nabla f(-5,2,-3) \cdot \vec{u} \\ &= \langle 20,-25,27 \rangle \cdot \langle -\frac{1}{\sqrt{35}}, \frac{5}{\sqrt{35}}, \frac{3}{\sqrt{35}} \rangle \\ &\approx -10.818 \end{aligned}$$

Example 3

Find the directional derivative of $f(x, y, z) = zy + x^2$ at the point $(2, 3, 1)$ in the direction of a vector making an angle of $\frac{\pi}{4}$ with $\nabla f(2, 3, 1)$.

ANS: $\nabla f(x, y, z) = 2x\vec{i} + z\vec{j} + y\vec{k}$

So $\nabla f(2,3,1) = 4\vec{i} + \vec{j} + 3\vec{k}$

Let $\vec{u}$ be a unit vector making an angle of $\frac{\pi}{4}$ with $\nabla f(2,3,1)$. Then by Thm 1, its directional directive is

$$\nabla f(2,3,1) \cdot \vec{u} = \|\nabla f(2,3,1)\| \cancel{\times} \|\vec{u}\| \cdot \cos\frac{\pi}{4}$$

$$= \sqrt{26} \cdot \frac{\sqrt{2}}{2}$$

Recall

$\vec{u} \cdot \vec{v} = \|\vec{u}\| \cdot \|\vec{v}\| \cdot \cos\theta$

where θ is the angle between $\vec{u}$ and $\vec{v}$.

Theorem 2. Maximum Rate of Change of a Function

Let f be a differentiable function on $U \subseteq \mathbb{R}^2$ (or $\mathbb{R}^3$) and assume that $\nabla f(\mathbf{x}) \neq \mathbf{0}$ for $\mathbf{x} \in U$. The direction of the largest rate of increase in f at $\mathbf{x}$ is given by the vector ∇f.

Proof. Let $\mathbf{u}$ be a unit vector. By Theorem 1,

$$D_{\mathbf{u}} f(\mathbf{x}) = \nabla f(\mathbf{x}) \cdot \mathbf{u} = \|\nabla f(\mathbf{x})\| \|\mathbf{u}\| \cos\theta = \|\nabla f(\mathbf{x})\| \cos\theta$$

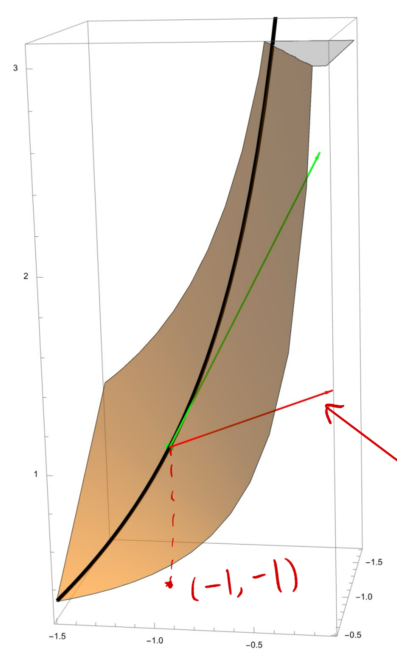
(since $\|\mathbf{u}\| = 1$), where θ denotes the angle between $\nabla f(\mathbf{x})$ and $\mathbf{u}$. Since $-1 \leq \cos\theta \leq 1$, $D_{\mathrm{u}} f(\mathbf{x})$ attains its largest value when $\cos\theta = 1$; that is, when $\theta = 0$. Consequently, maximum directional derivative $D_{\mathbf{u}} f(\mathbf{x})$ at $\mathbf{x}$ occurs in the direction parallel to the vector $\nabla f(\mathbf{x})$.

Example 4

Let $f(x, y) = \frac{x}{y}$, $P = (-1, -1)$ and $\mathbf{v} = \mathbf{i} + 2\mathbf{j}$.

Compute the following

1. The gradient of f.
2. The gradient of f at the point P.
3. The directional derivative of f at P in the direction of $\mathbf{v}$.
4. The maximum rate of change of f at P.
5. The unit direction vector $\mathbf{w}$ in which the maximum rate of change occurs at P.



ANS: 1 The gradient of f

$$\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right) = \left(\frac{1}{y}, -\frac{x}{y^2}\right)$$

2. $\nabla f(-1,-1) = (-1, 1)$

3. We normalize $\vec{v}$, $\vec{u} = \frac{\vec{v}}{\|\vec{v}\|} = \left(\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}\right)$

and $\nabla f(-1,-1) \cdot \vec{u}$

$$= (-1, 1) \cdot \left(\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}\right)$$

$$= \frac{1}{\sqrt{5}} \approx 0.447214.$$

4. By Thm 2, we know the direction of the largest rate of increase in f at P is given by the vector $\vec{\alpha} = \nabla f$ (at P) $= (-1, 1)$.

Normalize it, we have $\vec{w} = \frac{\vec{a}}{\|\vec{a}\|} = \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$

And the maximum rate of change is

$$\nabla f(-1,-1) \cdot \vec{u}_2 = (-1, 1) \cdot \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$$

$$= \sqrt{2} \approx 1.414$$

which is indicated in the green arrow.

5. By the previous disscuss, we know

$$\vec{w} = \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$$

which is the unit vector in the direction of ∇f at point P.

Exercise 5

Let $f(x,y) = x^2 y$ and $\mathbf{u}(t) = \left(t^2, 2t^3\right)$.

1. What is $\nabla f \cdot \mathbf{u}'(t)$?
2. Use the Chain Rule for Paths to compute $\frac{d}{dt} f(\mathbf{u}(t))$ at $t = -1$.

Answer:

1. We compute the partial derivatives of $f(x,y) = x^2 y$

$$\frac{\partial f}{\partial x} = 2xy, \qquad \frac{\partial f}{\partial y} = x^2$$

   The gradient vector is thus

$$\nabla f = \left(2xy, x^2\right)$$

   Also,

$$\mathbf{u}'(t) = \left(2t, 6t^2\right)$$
$$\nabla f \cdot \mathbf{u}'(t) = \left(2xy, x^2\right) \cdot \left(2t, 6t^2\right) = 4xyt + 6x^2t^2$$

2. Using the Chain Rule and substituting $x = t^2, y = 2t^3$ gives

$$\frac{d}{dt} f(\mathbf{u}(t)) = \frac{d}{dt}\left(2t^4t^3\right) = \frac{d}{dt}\left(2t^7\right) = 14t^6$$

   At the point $t = -1$, we obtain

$$\frac{d}{dt} f(\mathbf{u}(t))\Big|_{t=-1} = 14t^6\Big|_{t=-1} = 14$$