Exercise Sheet – Advanced Calculus III

Taiyang Xu*

18/09/2025, Week 2

Exercise 1 (Warm up: Directional derivatives and gradient)**.** Complete the following exercises.

(1) Let $f(x, y, z) = x^2y + yz^3$. Compute the gradient ∇f at the point $(1, 2, 1)$.

(2) For $f(x, y, z)$ as above, compute the directional derivative of f at $(1, 2, 1)$ in the direction of the vector $\vec{v} = (2, -1, 2)$.

(3) Let $f(x, y) = x^3 - 3xy^2$. Find all points (x, y) where the gradient ∇f is parallel to the vector $(1, 1)$.

Application in ML:

(4) (Linear regression) A common loss function is Mean Squared Error (MSE). For a single data point (x, y), the loss is defined as $L(m, b) = (y - (mx + b))^2$, where m is the slope and b is the y-intercept of the regression line. Compute the gradient $\nabla L(m, b)$ and interpret its components in terms of how they influence the loss.

(5) (Gradient descent step) Suppose you are using gradient descent to minimize the function $J(\theta_0, \theta_1) = \theta_0^2 + 2\theta_1^2$. Calculate the gradient $\nabla J(\theta_0, \theta_1)$ firstly, and write down the update rule for θ_0 and θ_1 using a learning rate α.

Solution 1. (1) The gradient is

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

Compute:

$$\frac{\partial f}{\partial x} = 2xy, \quad \frac{\partial f}{\partial y} = x^2 + z^3, \quad \frac{\partial f}{\partial z} = 3yz^2.$$

At $(1, 2, 1)$:

$$\frac{\partial f}{\partial x} = 2 \cdot 1 \cdot 2 = 4, \quad \frac{\partial f}{\partial y} = 1^2 + 1^3 = 2, \quad \frac{\partial f}{\partial z} = 3 \cdot 2 \cdot 1^2 = 6.$$

So,

$$\nabla f(1, 2, 1) = (4,\ 2,\ 6).$$

(2) The directional derivative in direction $\vec{v} = (2, -1, 2)$ is

$$D_{\vec{v}} f = \nabla f \cdot \frac{\vec{v}}{|\vec{v}|}.$$

Compute $|\vec{v}| = \sqrt{2^2 + (-1)^2 + 2^2} = \sqrt{9} = 3$. So, unit vector $\vec{u} = \left(\frac{2}{3}, -\frac{1}{3}, \frac{2}{3}\right)$. Then,

$$D_{\vec{v}} f = 4 \cdot \frac{2}{3} + 2 \cdot \left(-\frac{1}{3}\right) + 6 \cdot \frac{2}{3} = \frac{8}{3} - \frac{2}{3} + \frac{12}{3} = \frac{18}{3} = 6.$$

*School of Mathematical Sciences, Fudan University, Shanghai 200433, China. Email: tyxu19@fudan.edu.cn

(3) Compute $\nabla f(x,y) = (3x^2 - 3y^2,\ -6xy)$. We want $\nabla f(x,y)$ parallel to $(1,1)$, i.e.,

$$(3x^2 - 3y^2,\ -6xy) = \lambda(1,1).$$

So,

$$3x^2 - 3y^2 = \lambda, \quad -6xy = \lambda.$$

Equate:

$$x^2 - y^2 + 2xy = 0 \implies (x+y)^2 - 2y^2 = 0.$$

So,

$$x + y = \pm\sqrt{2}y.$$

Thus, all points (x,y) with $x = (\sqrt{2}-1)y$ or $x = (-\sqrt{2}-1)y$.

For $\nabla f(x,y)$ to be parallel to $(1,1)$, there must exist $\lambda \in \mathbb{R}$ such that

$$(3x^2 - 3y^2,\ -6xy) = \lambda(1,1).$$

So,

$$3x^2 - 3y^2 = \lambda, \quad -6xy = \lambda.$$

Equate the two expressions for λ:

$$3x^2 - 3y^2 = -6xy \implies x^2 + 2xy - y^2 = 0.$$

This factors as

$$(x+y)^2 - 2y^2 = 0 \implies (x+y)^2 = 2y^2 \implies x + y = \pm\sqrt{2}\, y.$$

Thus,

$$x = (\sqrt{2}-1)y \quad \text{or} \quad x = (-\sqrt{2}-1)y.$$

So all points (x,y) with $x = (\sqrt{2}-1)y$ or $x = (-\sqrt{2}-1)y$ have ∇f parallel to $(1,1)$.

(4)

$$\frac{\partial L}{\partial m} = -2x(y - (mx+b))$$

$$\frac{\partial L}{\partial b} = -2(y - (mx+b))$$

So,

$$\nabla L(m,b) = \big(-2x(y-(mx+b)),\ -2(y-(mx+b))\big)$$

Interpretation: the gradient points in the direction of steepest increase of the loss. The m-component shows how changing the slope affects the loss (scaled by x), and the b-component shows how changing the intercept affects the loss. Both are proportional to the residual $y - (mx+b)$.

(5) The gradient is

$$\nabla J(\theta_0, \theta_1) = \left(\frac{\partial J}{\partial \theta_0}, \frac{\partial J}{\partial \theta_1}\right) = (2\theta_0,\ 4\theta_1).$$

The gradient descent update rule is:

$$\theta_0^{\text{new}} = \theta_0 - \alpha \cdot 2\theta_0 = \theta_0(1 - 2\alpha)$$

$$\theta_1^{\text{new}} = \theta_1 - \alpha \cdot 4\theta_1 = \theta_1(1 - 4\alpha)$$

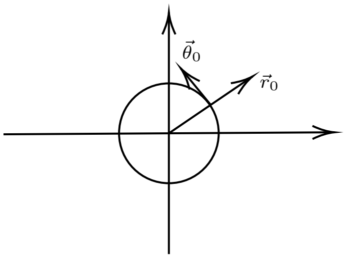
where α is the learning rate.

Exercise 2. Complete the following exercises

(1) Let $u = f(x, y)$, $x = r\cos\theta$, $y = r\sin\theta$. Prove that

$$\nabla u = \frac{\partial f}{\partial r}\vec{r}_0 + \frac{1}{r}\frac{\partial f}{\partial \theta}\vec{\theta}_0,$$

where $\vec{r}_0 = (\cos\theta, \sin\theta)$ and $\vec{\theta}_0 = (\cos(\theta + \frac{\pi}{2}), \cos\theta)$.



(2) We have the following equation:

$$\frac{x^2}{a^2+u} + \frac{y^2}{b^2+u} + \frac{z^2}{c^2+u} = 1.$$

Prove that

$$(\nabla u)^2 = 2 < \vec{A}, \nabla u >,$$

where $\vec{A} := (x, y, z)$.
(Hint: see $u = u(x, y, z)$)

Solution 2. (1) Recall that $\nabla u = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right)$ in Cartesian coordinates.

By the chain rule:

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial r}\frac{\partial r}{\partial x} + \frac{\partial f}{\partial \theta}\frac{\partial \theta}{\partial x},$$
$$\frac{\partial f}{\partial y} = \frac{\partial f}{\partial r}\frac{\partial r}{\partial y} + \frac{\partial f}{\partial \theta}\frac{\partial \theta}{\partial y}.$$

Compute the partial derivatives:

$$r = \sqrt{x^2+y^2} \implies \frac{\partial r}{\partial x} = \frac{x}{r} = \cos\theta, \quad \frac{\partial r}{\partial y} = \frac{y}{r} = \sin\theta,$$

$$\theta = \arctan\left(\frac{y}{x}\right) \implies \frac{\partial \theta}{\partial x} = -\frac{y}{r^2} = -\frac{\sin\theta}{r}, \quad \frac{\partial \theta}{\partial y} = \frac{x}{r^2} = \frac{\cos\theta}{r}.$$

Therefore,

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial r}\cos\theta - \frac{\partial f}{\partial \theta}\frac{\sin\theta}{r},$$
$$\frac{\partial f}{\partial y} = \frac{\partial f}{\partial r}\sin\theta + \frac{\partial f}{\partial \theta}\frac{\cos\theta}{r}.$$

So

$$\nabla u = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right) = \frac{\partial f}{\partial r}(\cos\theta, \sin\theta) + \frac{1}{r}\frac{\partial f}{\partial \theta}(-\sin\theta, \cos\theta).$$

Notice that $(\cos\theta, \sin\theta) = \vec{r}_0$ and $(-\sin\theta, \cos\theta) = (\cos(\theta + \frac{\pi}{2}), \sin(\theta + \frac{\pi}{2})) = \vec{\theta}_0$. Thus,

$$\nabla u = \frac{\partial f}{\partial r}\vec{r}_0 + \frac{1}{r}\frac{\partial f}{\partial \theta}\vec{\theta}_0.$$

(Vector decomposition)

$$\nabla u = < \nabla u, \vec{r}_0 > \vec{r}_0 + < \nabla u, \vec{\theta}_0 > \vec{\theta}_0.$$

Then you can directly calculate $< \nabla u, \vec{r}_0 >$ and $< \nabla u, \vec{\theta}_0 >$.

Exercise 3. Suppose the function $f(x, y)$ has a nonzero directional derivative at some point (x_0, y_0), and the directional derivatives along three different (non-colinear) directions at (x_0, y_0) are equal. Prove that $f(x, y)$ is not differentiable at (x_0, y_0).