

# 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  $\nabla f$  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  $\nabla f$  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  $\theta_0$  and  $\theta_1$  using a learning rate  $\alpha$ .

**Solution 2.** (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  $(1, 2, 1)$ :

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

---

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

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  $\vec{u} = (\frac{2}{3}, -\frac{1}{3}, \frac{2}{3})$ . Then,

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

(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  $\lambda$ :

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

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  $\nabla f$  parallel to  $(1, 1)$ .

(4)

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

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

So,

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

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:

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

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

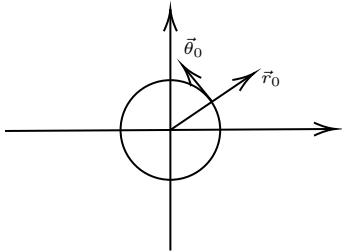
where  $\alpha$  is the learning rate.

**Exercise 3.** 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  $\vec{\theta}_0 = (-\sin \theta, \cos \theta)$ .



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

By the chain rule:

$$\begin{aligned} \frac{\partial f}{\partial x} &= \frac{\partial f}{\partial r} \frac{\partial r}{\partial x} + \frac{\partial f}{\partial \theta} \frac{\partial \theta}{\partial x}, \\ \frac{\partial f}{\partial y} &= \frac{\partial f}{\partial r} \frac{\partial r}{\partial y} + \frac{\partial f}{\partial \theta} \frac{\partial \theta}{\partial y}. \end{aligned}$$

Compute the partial derivatives:

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

Therefore,

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

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

(2) Noticing that

$$(\nabla u)^2 = 2 < \vec{A}, \nabla u > \Leftrightarrow u_x^2 + u_y^2 + u_z^2 = 2(xu_x + yu_y + zu_z).$$

Differentiate

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

with respect to  $x$ :

$$\frac{2x(a^2 + u) - u_x x^2}{(a^2 + u)^2} - \frac{y^2 u_x}{(b^2 + u)^2} - \frac{z^2 u_x}{(c^2 + u)^2} = 0,$$

that is,

$$\frac{2x}{a^2 + u} = \left[ \frac{x^2}{(a^2 + u)^2} + \frac{y^2}{(b^2 + u)^2} + \frac{z^2}{(c^2 + u)^2} \right] u_x. \quad (1)$$

Using similar manners (rotation), we have

$$\frac{2y}{b^2 + u} = \left[ \frac{x^2}{(a^2 + u)^2} + \frac{y^2}{(b^2 + u)^2} + \frac{z^2}{(c^2 + u)^2} \right] u_y, \quad (2)$$

$$\frac{2z}{c^2 + u} = \left[ \frac{x^2}{(a^2 + u)^2} + \frac{y^2}{(b^2 + u)^2} + \frac{z^2}{(c^2 + u)^2} \right] u_z. \quad (3)$$

Multiply  $x$  ( $y$ ,  $z$ ) to the both sides of the above three equations respectively, and add them together (also notice the condition)

$$2 = \left[ \frac{x^2}{(a^2 + u)^2} + \frac{y^2}{(b^2 + u)^2} + \frac{z^2}{(c^2 + u)^2} \right] (zu_x + yu_y + zu_z).$$

Square both sides of the mentioned above three equations (1)–(3), and add them together:

$$4 = \left[ \frac{x^2}{(a^2 + u)^2} + \frac{y^2}{(b^2 + u)^2} + \frac{z^2}{(c^2 + u)^2} \right] (u_x^2 + u_y^2 + u_z^2).$$

Combined the above two equations, we arrive at the desired result.

**Exercise 5.** 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  $(x_0, y_0)$ .

**Solution 6.** Suppose that  $f(x, y)$  is differentiable at  $(x_0, y_0)$ . Therefore there exist partial derivatives of  $f(x, y)$  at  $(x_0, y_0)$ :

$$A := f'_x(x_0, y_0), \quad B := f'_y(x_0, y_0),$$

and

$$df(x_0, y_0) = Adx + Bdy.$$

By the condition, there exist three different direction  $\vec{l}_i$  ( $i = 1, 2, 3$ ) such that

$$\frac{\partial f}{\partial \vec{l}_i}(x_0, y_0) = A \cos \theta_i + B \sin \theta_i = C \neq 0.$$

We have the following linear systems:

$$A \cos \theta_1 + B \sin \theta_1 - C = 0,$$

$$A \cos \theta_2 + B \sin \theta_2 - C = 0,$$

$$A \cos \theta_3 + B \sin \theta_3 - C = 0.$$

We will take advantage of the following truth:

[ The linear equations only has trivial solution ( $A = B = C = 0$ ), if and only if

$$\det \begin{pmatrix} \cos \theta_1 & \sin \theta_1 & 1 \\ \cos \theta_2 & \sin \theta_2 & 1 \\ \cos \theta_3 & \sin \theta_3 & 1 \end{pmatrix} \neq 0.$$

]

Noting that  $\exists \alpha, \beta \neq 0$  such that

$$\vec{l}_1 = \alpha \vec{l}_2 + \beta \vec{l}_3 \implies \cos \theta_1 = \alpha \cos \theta_2 + \beta \cos \theta_3, \quad \sin \theta_1 = \alpha \sin \theta_2 + \beta \sin \theta_3.$$

Square both sides and add them together

$$1 = \alpha^2 + \beta^2 + 2\alpha\beta \cos(\theta_2 - \theta_3) \stackrel{(\theta_2 \neq \theta_3)}{\neq} (\alpha + \beta)^2 \implies \alpha + \beta \neq 1.$$

Then

$$\begin{aligned} \det \begin{pmatrix} \cos \theta_1 & \sin \theta_1 & 1 \\ \cos \theta_2 & \sin \theta_2 & 1 \\ \cos \theta_3 & \sin \theta_3 & 1 \end{pmatrix} &= \det \begin{pmatrix} \alpha \cos \theta_2 + \beta \cos \theta_3 & \alpha \sin \theta_2 + \beta \sin \theta_3 & 1 \\ \cos \theta_2 & \sin \theta_2 & 1 \\ \cos \theta_3 & \sin \theta_3 & 1 \end{pmatrix} \\ &= \det \begin{pmatrix} 0 & 0 & 1 \\ \cos \theta_2 & \sin \theta_2 & 1 - \alpha - \beta \\ \cos \theta_3 & \sin \theta_3 & 1 \end{pmatrix} = (1 - \alpha - \beta) \sin(\theta_3 - \theta_2) \neq 0. \end{aligned}$$

Thus  $(A, B, C) = (0, 0, 0)$ .