

Mathematical Concepts 1

**Solution 1:**

Gradient

- (a) The gradient  $\nabla f(\mathbf{x}) = (2x_1 + x_2, x_2 + x_1)^\top$  is continuous  $\Rightarrow f \in \mathcal{C}^1$ .
- (b) The direction of greatest increase is given by the gradient, i.e.,  $\nabla f(1, 1) = (3, 2)^\top$ .
- (c) Let  $\mathbf{v} \in \mathbb{R}^2$  be a direction with fixed length  $\|\mathbf{v}\|_2 = r > 0$ .  
 The directional derivative  $D_{\mathbf{v}}f(\mathbf{x}) = \nabla f(\mathbf{x})^\top \mathbf{v} = \|\nabla f(\mathbf{x})\|_2 \|\mathbf{v}\|_2 \cos(\theta) = \|\nabla f(\mathbf{x})\|_2 r \cos(\theta)$ . This becomes minimal if  $\theta = \pi$ , i.e., if  $\mathbf{v}$  points in the opposite direction of  $\nabla f \Rightarrow \mathbf{v} = -\nabla f(\mathbf{x})$  if  $r = \|\nabla f(\mathbf{x})\|_2$ . Here, the direction of greatest decrease is given by  $-\nabla f(1, 1) = (-3, -2)^\top$ .
- (d)  $D_{\mathbf{v}}f(\mathbf{x}) = \nabla f(1, 1)^\top \mathbf{v} \stackrel{!}{=} 0 \Rightarrow (3, 2) \cdot \mathbf{v} = 0 \iff \mathbf{v} = \alpha \cdot (-2, 3)^\top$  with  $\alpha \in \mathbb{R}$  and  $\alpha \neq 0$ .
- (e) When we differentiate both sides of the equation  $f(\tilde{\mathbf{x}}(t)) = f(1, 1)$  w.r.t.  $t$  we arrive at  $\frac{\partial f(\tilde{\mathbf{x}}(t))}{\partial t} = 0$ . Via the chain rule it follows that  $\underbrace{\frac{\partial f}{\partial \tilde{\mathbf{x}}}}_{=\nabla f(\tilde{\mathbf{x}})^\top} \frac{\partial \tilde{\mathbf{x}}}{\partial t} = 0$ .
- (f) The gradient is orthogonal to the tangent line of the level curves.

**Solution 2:**

Convexity

- (a) Let  $x, y \in \mathbb{R}$  and  $t \in [0, 1]$  then it holds that

$$\begin{aligned} (f + g)(x + t(y - x)) &= f(x + t(y - x)) + g(x + t(y - x)) \\ &\leq f(x) + t(f(y) - f(x)) + g(x) + t(g(y) - g(x)) && (f, g \text{ are convex}) \\ &= f(x) + g(x) + t(f(y) + g(y) - (f(x) + g(x))) \\ &= (f + g)(x) + t((f + g)(y) - (f + g)(x)). \end{aligned}$$

- (b) Let  $x, y \in \mathbb{R}$  and  $t \in [0, 1]$  then it holds that

$$\begin{aligned} (g \circ f)(x + t(y - x)) &= g(f(x + t(y - x))) \\ &\leq g(f(x) + t(f(y) - f(x))) && (g \text{ is non-decreasing, } f \text{ is convex}) \\ &\leq g(f(x)) + t(g(f(y)) - g(f(x))) && (g \text{ is convex}) \\ &= (g \circ f)(x) + t((g \circ f)(y) - (g \circ f)(x)). \end{aligned}$$

**Solution 3:**

Convexity

Consider the bivariate function  $f : \mathbb{R}^2 \rightarrow \mathbb{R}, (x_1, x_2) \mapsto \exp(\pi \cdot x_1) - \sin(\pi \cdot x_2) + \pi \cdot x_1 \cdot x_2$

- (a)  $\nabla f(\mathbf{x}) = \pi \cdot (\exp(\pi x_1) + x_2, -\cos(\pi x_2) + x_1)^\top$
- (b)  $\nabla^2 f(\mathbf{x}) = \pi \cdot \begin{pmatrix} \pi \exp(\pi x_1) & 1 \\ 1 & \pi \sin(\pi x_2) \end{pmatrix}$
- (c)  $T_{1, \mathbf{a}}(\mathbf{x}) = f(\mathbf{a}) + \nabla f(\mathbf{a})^\top (\mathbf{x} - \mathbf{a}) = 1 + \pi \cdot (2, 1) \cdot (x_1, x_2 - 1)^\top = 1 - \pi + 2\pi x_1 + \pi x_2$

(d)

$$\begin{aligned} T_{2,\mathbf{a}}(\mathbf{x}) &= T_{1,\mathbf{a}}(\mathbf{x}) + \frac{1}{2}(\mathbf{x} - \mathbf{a})^\top \nabla^2 f(\mathbf{a})(\mathbf{x} - \mathbf{a}) \\ &= T_{1,\mathbf{a}}(\mathbf{x}) + \frac{1}{2}\mathbf{x}^\top \nabla^2 f(\mathbf{a})\mathbf{x} - \mathbf{x}^\top \nabla^2 f(\mathbf{a})\mathbf{a} + \frac{1}{2}\mathbf{a}^\top \nabla^2 f(\mathbf{a})\mathbf{a} \end{aligned}$$

With  $\nabla^2 f(\mathbf{a}) = \begin{pmatrix} \pi^2 & \pi \\ \pi & 0 \end{pmatrix}$  we get that

$$\begin{aligned} T_{2,\mathbf{a}}(\mathbf{x}) &= T_{1,\mathbf{a}}(\mathbf{x}) + 0.5\pi^2 x_1^2 \\ &\quad + \pi x_1 x_2 - \pi x_1 \\ &\quad + 0. \end{aligned}$$

.

(e)  $T_{2,\mathbf{a}}(\mathbf{x})$  is multivariate polynomial of degree 2 which means its Hessian is constant and we can directly see that  $\mathbf{H} := \nabla^2 T_{2,\mathbf{a}}(\mathbf{x}) = \nabla^2 f(\mathbf{a})$ . For the eigenvalues of the Hessian it must hold that

$$\begin{aligned} \det(\mathbf{H} - \lambda \mathbf{I}) &= 0 \\ \iff \det \begin{pmatrix} \pi^2 - \lambda & \pi \\ \pi & -\lambda \end{pmatrix} &= 0 \\ \iff (\pi^2 - \lambda) \cdot (-\lambda) - \pi^2 &= 0 \\ \iff \lambda^2 - \pi^2 \lambda - \pi^2 &= 0. \end{aligned}$$

From which it follows that  $\lambda_{1,2} = \frac{\pi^2 \pm \sqrt{\pi^4 + 4\pi^2}}{2} \Rightarrow \lambda_1 \approx 10.785, \lambda_2 \approx -0.915$ . Since  $\lambda_2 < 0$   $T_{2,\mathbf{a}}$  is not convex.