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:
Matrix Calculus

(a) $\mathbf{Ab} = \begin{pmatrix} 1 & 0 & 3 \\ 2 & 1 & 4 \\ 0 & 3 & 2 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 1\cdot 3 + 0 \cdot 2 + 3 \cdot 1 \\ 2 \cdot 3 + 1 \cdot 2 + 4 \cdot 1 \\ 0 \cdot 3 + 3 \cdot 2 + 2 \cdot 1 \\ 1 \cdot 3 + 1 \cdot 2 + 0 \cdot 1 \end{pmatrix} = \begin{pmatrix} 6 \\ 12 \\ 8 \\ 5 \end{pmatrix}$

(b) $\mathbf{A}$ must be symmetric, such that $\mathbf{A}^T = \mathbf{A}$. As the dimensions of columns and rows of $\mathbf{A}$ do not match, $\mathbf{A}$ cannot be symmtric.

(c) To be invertible, this matrix has to be positive definite and therefore $|\mathbf{A}| > 0$ must hold. The determinant of $\mathbf{A}$ is calculated as $|\mathbf{A}| = 1\cdot 1\cdot 2 + 0 \cdot 4 \cdot 0 + 3 \cdot 2 \cdot 3 - 0 \cdot 1 \cdot 3 - 3 \cdot 4 \cdot 1 - 2 \cdot 2 \cdot 0 = 2 + 0 + 18 - 0 - 12 - 0 = 8$. This is higher than 0 and therefore $\mathbf{A}$ is invertible.

(d) The inverse of a matrix must fulfill $\mathbf{AA}^{-1} = \mathbf{I}$. Hence, we search the entries of $\mathbf{A}^{-1}$ to fulfill that condition:
$\begin{pmatrix} 1 & 0 & 3 \\ 2 & 1 & 4 \\ 0 & 3 & 2 \end{pmatrix} \begin{pmatrix} \bar{a}_{11} & \bar{a}_{12} & \bar{a}_{13} \\ \bar{a}_{21} & \bar{a}_{22} & \bar{a}_{23} \\ \bar{a}_{31} & \bar{a}_{32} & \bar{a}_{33} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$
Gauss-Jordan-algorithm:
$\left(\begin{array}{ccc|ccc} 1 & 0 & 3 & 1 & 0 & 0 \\ 2 & 1 & 4 & 0 & 1 & 0 \\ 0 & 3 & 2 & 0 & 0 & 1 \end{array}\right)$
Multiply row 1 with -2 and add it to row 2:
$\left(\begin{array}{ccc|ccc} 1 & 0 & 3 & 1 & 0 & 0 \\ 0 & 1 & -2 & -2 & 1 & 0 \\ 0 & 3 & 2 & 0 & 0 & 1 \end{array}\right)$
Multiply row 2 with -3 and add it to row 3:
$\left(\begin{array}{ccc|ccc} 1 & 0 & 3 & 1 & 0 & 0 \\ 0 & 1 & -2 & -2 & 1 & 0 \\ 0 & 0 & 8 & 6 & -3 & 1 \end{array}\right)$
Multiply row 3 with 1/8:

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 3 & 1 & 0 & 0 \\ 0 & 1 & -2 & -2 & 1 & 0 \\ 0 & 0 & 1 & 3/4 & -3/8 & 1/8 \end{array}\right)$$
Multiply row 3 with 2 and add it to row 2:
$$\left(\begin{array}{ccc|ccc} 1 & 0 & 3 & 1 & 0 & 0 \\ 0 & 1 & 0 & -1/2 & 1/4 & 1/4 \\ 0 & 0 & 1 & 3/4 & -3/8 & 1/8 \end{array}\right)$$
Multiply row 3 with -3 and add it to row 1:
$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -5/4 & 9/8 & -3/8 \\ 0 & 1 & 0 & -1/2 & 1/4 & 1/4 \\ 0 & 0 & 1 & 3/4 & -3/8 & 1/8 \end{array}\right)$$
Hence, the inverse is
$$\mathbf{A}^{-1} = \begin{pmatrix} -5/4 & 9/8 & -3/8 \\ -1/2 & 1/4 & 1/4 \\ 3/4 & -3/8 & 1/8 \end{pmatrix} = \frac{1}{|\mathbf{A}|} \begin{pmatrix} -10 & 9 & -3 \\ -4 & 2 & 2 \\ 6 & -3 & 1 \end{pmatrix}.$$
Test:
$$\begin{pmatrix} 1 & 0 & 3 \\ 2 & 1 & 4 \\ 0 & 3 & 2 \end{pmatrix} \frac{1}{8} \begin{pmatrix} -10 & 9 & -3 \\ -4 & 2 & 2 \\ 6 & -3 & 1 \end{pmatrix} = \frac{1}{8} \begin{pmatrix} 1 \cdot -10 + 3 \cdot 6 & 1 \cdot 9 - 3 \cdot 3 & 1 \cdot -3 + 3 \cdot 1 \\ 2 \cdot -10 + 1 \cdot -4 + 4 \cdot 6 & 2 \cdot 9 + 1 \cdot 2 + 4 \cdot -3 & 2 \cdot -3 + 1 \cdot 2 + 4 \cdot 1 \\ 3 \cdot -4 + 2 \cdot 6 & 3 \cdot 2 + 2 \cdot -3 & 3 \cdot 2 + 2 \cdot 1 \end{pmatrix} =$$
$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Solution 3:
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 4:
Convexity

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

(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

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

.