Optimization in Machine Learning WS 2024/2025

Exercise sheet 1

Mathematical Concepts 1

Exercise 1: Gradient

Consider the bivariate function $f: \mathbb{R}^2 \to \mathbb{R}, (x_1, x_2) \mapsto x_1^2 + 0.5x_2^2 + x_1x_2$.

- (a) Show that f is smooth (as defined in the lecture).
- (b) Find the direction of greatest increase of f at $\mathbf{x} = (1, 1)$.
- (c) Find the direction of greatest decrease of f at $\mathbf{x} = (1,1)$.
- (d) Find a direction in which f does not instantly change at $\mathbf{x} = (1, 1)$.
- (e) Assume there exists a differentiable parametrization of a curve $\tilde{\mathbf{x}}: \mathbb{R} \to \mathbb{R}^2, t \mapsto \tilde{\mathbf{x}}(t)$ such that $\forall t \in \mathbb{R}: f(\tilde{\mathbf{x}}(t)) = f(1,1)$. Show that at each point of the curve $\tilde{\mathbf{x}}$ the tangent line $\frac{\partial \tilde{\mathbf{x}}}{\partial t}$ is perpendicular to the gradient $\nabla f(\tilde{\mathbf{x}})$.
- (f) Interpret (d), (e) geometrically

Exercise 2: Matrix Calculus

Consider matrix
$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 3 \\ 2 & 1 & 4 \\ 0 & 3 & 2 \\ 1 & 1 & 0 \end{pmatrix}$$
 and vector $\mathbf{b} = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}$

- (a) Calculate the matrix-vector-product of **Ab**
- (b) Which condition must be fulfilled, such that $\mathbf{A}\mathbf{b} = \mathbf{A}^T\mathbf{b}$ holds

Now, matrix **A** is given as
$$\begin{pmatrix} 1 & 0 & 3 \\ 2 & 1 & 4 \\ 0 & 3 & 2 \end{pmatrix}$$

- (c) Show, that this matrix is invertible
- (d) Calculate the inverse of A

Exercise 3: Convexity

Consider two convex functions $f, g : \mathbb{R} \to \mathbb{R}$.

- (a) Show that $f + g : \mathbb{R} \to \mathbb{R}, x \mapsto f(x) + g(x)$ is convex.
- (b) Now, assume that g is additionally non-decreasing, i.e., $g(y) \ge g(x) \ \forall x \in \mathbb{R}, \forall y \in \mathbb{R}$ with y > x. Show that $g \circ f$ is convex.

Exercise 4: Taylor polynomials

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) Compute the gradient of f for an arbitrary \mathbf{x} .

- (b) Compute the Hessian of f for an arbitrary \mathbf{x} .
- (c) State the first order taylor polynomial $T_{1,\mathbf{a}}(\mathbf{x})$ expanded around the point $\mathbf{a}=(0,1).$
- (d) State the second order taylor polynomial $T_{2,\mathbf{a}}(\mathbf{x})$ expanded around the point $\mathbf{a}=(0,1).$