

Introduction to Machine Learning
Fall 2025
University of Science and Technology of China

Lecturer: Jie Wang
Posted: Nov. 7, 2025

Homework 4
Due: Nov. 13, 2025

Notice, to get the full credits, please present your solutions step by step.

Exercise 1: Convex Functions

1. Please show that the following functions are convex.

(a) $f(\mathbf{x}) = \sum_{i=1}^k x_{[i]}$ on $\mathbf{dom} f = \mathbb{R}^n$, where $1 \leq k \leq n$ and $x_{[i]}$ denotes the i^{th} largest component of $\mathbf{x}$.

(b) The negative entropy, i.e.,

$$f(\mathbf{p}) = \sum_{i=1}^n p_i \log p_i$$

on $\mathbf{dom} f = \{\mathbf{p} \in \mathbb{R}^n : 0 < p_i \leq 1, \sum_{i=1}^n p_i = 1\}$, where p_i denotes the i^{th} component of $\mathbf{p}$.

(c) The p -norms, i.e., $f(\mathbf{X}) = \|\mathbf{X}\|_p$ on $\mathbf{dom} f = \mathbb{R}^{m \times n}$.

(d) The function $f(\mathbf{X}) = -\log \det \mathbf{X}$ on $\mathbf{dom} f = \mathbf{S}_{++}^n$

2. Please show that a function f is convex if and only if $\mathbf{dom} f$ is convex and its restriction to any line intersecting its domain is convex, i.e., for any $\mathbf{x}_0 \in \mathbf{dom} f$ and $\mathbf{v} \in \mathbb{R}^n$, the function

$$g(t) = f(\mathbf{x}_0 + t\mathbf{v})$$

is convex over its domain $\mathbf{dom} g = \{t \in \mathbb{R} : \mathbf{x}_0 + t\mathbf{v} \in \mathbf{dom} f\}$.

(**Hint:** you may prove the sufficiency by contradiction.)

3. (Optional) Please show that a continuously differentiable function f is strongly convex with parameter $\mu > 0$ if and only if

$$f(\mathbf{y}) \geq f(\mathbf{x}) + \langle \nabla f(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle + \frac{\mu}{2} \|\mathbf{y} - \mathbf{x}\|_2^2, \quad \forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n.$$

4. (Optional) Suppose that f is twice continuously differentiable and strongly convex with parameter $\mu > 0$. Please show that $\mu \leq \lambda_{\min}(\nabla^2 f(\mathbf{x}))$ for any $\mathbf{x} \in \mathbb{R}^n$, where $\lambda_{\min}(\nabla^2 f(\mathbf{x}))$ is the smallest eigenvalue of $\nabla^2 f(\mathbf{x})$.

5. Suppose that $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is twice continuously differentiable, and the gradient of f is Lipschitz continuous, i.e.,

$$\|\nabla f(\mathbf{x}) - \nabla f(\mathbf{y})\|_2 \leq L \|\mathbf{x} - \mathbf{y}\|_2, \quad \forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n,$$

where $L > 0$ is the Lipschitz constant. Please show that $\lambda_{\max}(\nabla^2 f(\mathbf{x})) \leq L$ for any $\mathbf{x} \in \mathbb{R}^n$, where $\lambda_{\max}(\nabla^2 f(\mathbf{x}))$ is the largest eigenvalue of $\nabla^2 f(\mathbf{x})$.

Homework 4

Exercise 2: Operations that Preserve Convexity

1. Let $f : \mathbb{R}^m \rightarrow (-\infty, +\infty]$ be a given convex function, $\mathbf{A} \in \mathbb{R}^{m \times n}$ and $\mathbf{b} \in \mathbb{R}^m$. Please show that

$$F(\mathbf{x}) = f(\mathbf{Ax} + \mathbf{b}), \quad \mathbf{x} \in \mathbb{R}^n.$$

is convex.

2. Let $f_i : \mathbb{R}^n \rightarrow (-\infty, +\infty], i = 1, \dots, m$, be given convex functions. Please show that

$$F(\mathbf{x}) = \sum_{i=1}^m w_i f_i(\mathbf{x})$$

is convex, where $w_i \geq 0, i = 1, \dots, m$.

3. Let $f_i : \mathbb{R}^n \rightarrow (-\infty, +\infty]$ be given convex functions for $i \in I$, where I is an arbitrary index set. Please show that the supremum

$$F(\mathbf{x}) = \sup_{i \in I} f_i(\mathbf{x})$$

is convex.

4. (Optional) Let $\mathbf{A} \in \mathbb{R}^{n \times m}, \mathbf{x}_0 \in \mathbb{R}^n$. The restriction of $f : \mathbb{R}^n \rightarrow \mathbb{R}$ to the affine set $\{\mathbf{Az} + \mathbf{x}_0 | \mathbf{z} \in \mathbb{R}^m\}$ is defined as the function $F : \mathbb{R}^m \rightarrow \mathbb{R}$ with

$$F(\mathbf{z}) = f(\mathbf{Az} + \mathbf{x}_0)$$

on $\text{dom } F = \{\mathbf{z} | \mathbf{Az} + \mathbf{x}_0 \in \text{dom } f\}$. Suppose f is twice differentiable with a convex domain.[1]

- (a) Show that F is convex if and only if for all $\mathbf{z} \in \text{dom } F$, we have

$$\mathbf{A}^\top \nabla^2 f(\mathbf{Az} + \mathbf{x}_0) \mathbf{A} \succeq 0.$$

- (b) Suppose $\mathbf{B} \in \mathbb{R}^{p \times n}$ is a matrix whose nullspace is equal to the range of $\mathbf{A}$, i.e., $\mathbf{AB} = \mathbf{0}$ and $\text{rank}(\mathbf{B}) = n - \text{rank}(\mathbf{A})$. Show that F is convex if for all $\mathbf{z} \in \text{dom } F$, there exists a $\lambda \in \mathbb{R}$ such that

$$\nabla^2 f(\mathbf{Az} + \mathbf{x}_0) + \lambda \mathbf{B}^\top \mathbf{B} \succeq 0.$$

(**Hint:** you can use the result as follows. If $\mathbf{C} \in \mathbb{S}^n$ and $\mathbf{D} \in \mathbb{R}^{p \times n}$, then $\mathbf{x}^\top \mathbf{Cx} \geq 0$ for all $\mathbf{x} \in \mathcal{N}(\mathbf{D})$ if there exists a λ such that $\mathbf{C} + \lambda \mathbf{D}^\top \mathbf{D} \succeq 0$.)

Homework 4

Exercise 3: Subdifferentials

Calculation of subdifferentials (you need to finish at least four of the problems).

1. Let $H \subset \mathbb{R}^n$ be a hyperplane. The extended-value extension of its indicator function I_H is

$$\tilde{I}_H(\mathbf{x}) = \begin{cases} 0, & \mathbf{x} \in H, \\ \infty, & \mathbf{x} \notin H. \end{cases}$$

Find $\partial \tilde{I}_H(\mathbf{x})$.

2. Let $f(\mathbf{x}) = \exp \|\mathbf{x}\|_1$, $\mathbf{x} \in \mathbb{R}^n$. Find $\partial f(\mathbf{x})$.
3. For $\mathbf{x} \in \mathbb{R}^n$, let $x_{[i]}$ be the i^{th} largest component of $\mathbf{x}$. Find the subdifferentials of

$$f(\mathbf{x}) = \sum_{i=1}^k x_{[i]}.$$

4. Let $f(\mathbf{x}) = \|\mathbf{x}\|_\infty$, $\mathbf{x} \in \mathbb{R}^n$. Find $\partial f(\mathbf{x})$.
5. Let $f(X) = \max_{1 \leq i \leq n} |\lambda_i|$, where $X \in \mathbb{S}^n$ and $\lambda_1, \dots, \lambda_n$ are the eigenvalues of X . Find $\partial f(X)$.

(**Hint:** you can refer to Example 7 in Lec06.)

6. (Optional) Let

$$f(\mathbf{x}) = \left(\sum_{i=1}^k x_i^2 \right)^{\frac{1}{2}} + \left(\sum_{i=k+1}^n x_i^2 \right)^{\frac{1}{2}}, \quad \mathbf{x} \in \mathbb{R}^n,$$

where $1 \leq k \leq n-1$. Find $\partial f(\mathbf{x})$.

7. (Optional) Let $f(\mathbf{X}) = \|\mathbf{X}\|_*$ be the trace norm of $\mathbf{X} \in \mathbb{R}^{m \times n}$. Find $\partial f(\mathbf{X})$.

Homework 4

Exercise 4: Convex Functions and Optimization

Consider the problem

$$\min_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x}), \quad (1)$$

where $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is continuously differentiable and convex, and $\mathbf{dom} f$ is closed.

1. Please show that the α -sublevel set of f , i.e., $C_\alpha = \{\mathbf{x} \in \mathbf{dom} f : f(\mathbf{x}) \leq \alpha\}$ is closed.
2. Please give an example to show that Problem (1) may be unsolvable even if f is strictly convex.
3. Suppose that f can attain its minimum. Please show that the optimal set $\mathcal{C} = \{\mathbf{y} : f(\mathbf{y}) = \min_{\mathbf{x}} f(\mathbf{x})\}$ is closed and convex. Does this property still hold if $\mathbf{dom} f$ is not closed?
4. Suppose that f is strongly convex with parameter $\mu > 0$. Please show that Problem (1) admits a unique solution.
5. Consider the problem

$$\min_{\mathbf{w} \in \mathbb{R}^n} f(\mathbf{w}) = \|\mathbf{X}\mathbf{w} - \mathbf{y}\|^2 + \frac{\lambda}{2}\|\mathbf{w}\|^2, \quad (2)$$

where $\mathbf{X} \in \mathbb{R}^{m \times n}$ and $\mathbf{y} \in \mathbb{R}^m$

- (a) Please show that Problem (2) admits a unique solution.
- (b) if $\hat{\mathbf{w}} = \arg \min_{\mathbf{w}} f(\mathbf{w})$, try to show that

$$\langle \nabla f(\hat{\mathbf{w}}), \mathbf{z} - \hat{\mathbf{w}} \rangle = 0, \quad \forall \mathbf{z} \in \mathbb{R}^n,$$

and find the supporting hyperplane of $\mathbf{epi} f$ at the point $(\hat{\mathbf{w}}, f(\hat{\mathbf{w}}))$ using the gradient $\nabla f(\hat{\mathbf{w}})$.

References

- [1] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, Cambridge, UK, 2004.