Introduction to Machine Learning
Fall 2023
University of Science and Technology of China

Lecturer: Jie Wang — Homework 4
Posted: Nov. 1, 2023 — Due: Nov. 8, 2023

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 spectral norm, i.e.,
   $$f(\mathbf{X}) = \|\mathbf{X}\|_2 = \sigma_{\max}(\mathbf{X})$$
   on $\mathbf{dom}\ f = \mathbb{R}^{m\times n}$, where $\sigma_{\max}$ denotes the largest singular value of $\mathbf{X}$.

2. please show the following two equalities:
   $$f(\mathbf{y}) - f(\mathbf{x}) = \int_0^1 \nabla f(\mathbf{x} + t(\mathbf{y} - \mathbf{x}))^\top (\mathbf{y} - \mathbf{x})\mathrm{d}t \tag{1}$$
   $$\nabla f(\mathbf{y}) - \nabla f(\mathbf{x}) = \int_0^1 \nabla^2 f(\mathbf{x} + t(\mathbf{y} - \mathbf{x}))(\mathbf{y} - \mathbf{x})\mathrm{d}t \tag{2}$$
   (**Hint:** you may consider the function $g(t) = f(\mathbf{x}+t(\mathbf{y}-\mathbf{x}))$ and apply the fundamental theorem of calculus.)

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 $f : \mathbb{R}^n \to \mathbb{R}$ 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 $\nabla^2 f(\mathbf{x})$.

6. Consider the problem

$$\min_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x}), \tag{3}$$

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

(a) 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.

(b) Please give an example to show that Problem (3) may be unsolvable even if f is strictly convex.

(c) 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?

(d) Suppose that f is strongly convex with parameter $\mu > 0$. Please show that Problem (3) admits a unique solution.

Solution: ■

Exercise 2: Operations that Preserve Convexity

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

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

   is convex.

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

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

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

3. Let $f_i : \mathbb{R}^n \to (-\infty, +\infty]$ 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.

Solution: ■

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 $f(X) = \max\limits_{1 \le i \le n} |\lambda_i|$, where $X \in \mathbb{S}^n$ and $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of X. Find $\partial f(X)$(**Hint**: you can refer to Example 5 in Lec06).

Solution: ■

References