Introduction to Machine Learning

Fall 2019

University of Science and Technology of China

Lecturer: Jie Wang | Homework 2
Posted: Mar. 12, 2020 | Due: Mar. 19, 2020
Name: San Zhang | ID: PBXXXXXXXX

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

Exercise 1: Limit and Limit Points 10pts

1. Show that $\{\mathbf{x}_n\}$ in $\mathbb{R}^n$ converges to $\mathbf{x} \in \mathbb{R}^n$ if and only if $\{\mathbf{x}_n\}$ is bounded and has a unique limit point $\mathbf{x}$.

2. (**Limit Points of a Set**). Let C be a subset of $\mathbb{R}^n$. A point $\mathbf{x} \in \mathbb{R}^n$ is called a limit point of C if there is a sequence $\{\mathbf{x}_n\}$ in C such that $\mathbf{x}_n \to \mathbf{x}$ and $\mathbf{x}_n \neq \mathbf{x}$ for all positive integers n. If $\mathbf{x} \in C$ and $\mathbf{x}$ is not a limit point of C, then $\mathbf{x}$ is called an isolated point of C. Let C' be the set of limit points of the set C. Please show the following statements.

   (a) If $C = (0,1) \cup \{2\} \subset \mathbb{R}$, then $C' = [0,1]$ and $x = 2$ is an isolated point of C.

   (b) The set C' is closed.

   (c) The closure of C is the union of C' and C; that is $\mathbf{cl}\, C = C' \cup C$. Moreover, $C' \subset C$ if and only if C is closed.

Solution: ■

Exercise 2: Open and Closed Sets 10pts

The norm ball $\{\mathbf{y} \in \mathbb{R}^n : \|\mathbf{y} - \mathbf{x}\|_2 < r, \mathbf{x} \in \mathbb{R}^n\}$ is denoted by $B_r(\mathbf{x})$.

1. Given a set $C \subset \mathbb{R}^n$, please show the following are equivalent.
   (a) The set C is closed; that is $\mathbf{cl}\, C = C$.
   (b) The complement of C is open.
   (c) If $B_\epsilon(\mathbf{x}) \cap C \neq \emptyset$ for every $\epsilon > 0$, then $\mathbf{x} \in C$.
2. Given $A \subset \mathbb{R}^n$, a set $C \subset A$ is called open in A if

$$C = \{\mathbf{x} \in C : B_\epsilon(\mathbf{x}) \cap A \subset C \,\text{for some}\, \epsilon > 0\}.$$

   A set C is said to be closed in A if $A \setminus C$ is open in A.
   (a) Let $B = [0, 1] \cup \{2\}$. Please show that $[0, 1]$ is not an open set in $\mathbb{R}$, while it is both open and closed in B.
   (b) Please show that a set $C \subset A$ is open in A if and only if $C = A \cap U$, where U is open in $\mathbb{R}^n$.

Solution: ■

Exercise 3: Bolzano-Weierstrass Theorem 10pts

The Least Upper Bound Axiom

Any nonempty set of real numbers with an upper bound has a least upper bound. That is, $\sup C$ *always exists for a nonempty bounded above set* $C \subset \mathbb{R}$*.*

Please show the following statements from **the least upper bound axiom**.

1. Let C be a nonempty subset of $\mathbb{R}$ that is bounded above. Prove that $u = \sup C$ if and only if u is an upper bound of C and

$$\forall\, \epsilon > 0, \exists\, a \in C \text{ such that } a > u - \epsilon.$$

2. Every bounded sequence in $\mathbb{R}$ has at least one limit point.

3. Every bounded sequence in $\mathbb{R}^n$ has at least one limit point.

Solution: ■

Exercise 4: Extreme Value Theorem 15pts

1. Show that a set $C \subset \mathbb{R}^n$ is compact if and only if every sequence in C has a subsequence that converges to a point in C.
2. Let C be a compact subset of $\mathbb{R}^n$ and $f : C \to \mathbb{R}$ be continuous. Please show that there exist $\mathbf{a}, \mathbf{b} \in C$ such that

$$f(\mathbf{a}) \leq f(\mathbf{x}) \leq f(\mathbf{b}), \forall\, \mathbf{x} \in C.$$

3. Let $f : [a, b] \to \mathbb{R}$ be continuous. Show that the range of f is a compact interval $[c, d]$ for some $c, d \in \mathbb{R}$.

Solution: ■

Exercise 5: Convex Sets 10pts

Let $C \subset \mathbb{R}^n$ be a convex set. Please show the following statements.

1. The intersection $\cap_{i \in I} C_i$ of any collection $\{C_i : i \in I\}$ of convex sets is convex.
2. Both **cl** C and **int** C are convex.
3. The set$\{\mathbf{y} \in \mathbb{R}^m : \mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{a}, \mathbf{x} \in C\}$ is convex, where $\mathbf{A} \in \mathbb{R}^{m \times n}$ and $\mathbf{a} \in \mathbb{R}^m$.
4. The set $\{\mathbf{y} \in \mathbb{R}^m : \mathbf{x} = \mathbf{B}\mathbf{y} + \mathbf{b}, \mathbf{x} \in C\}$ is convex, where $\mathbf{B} \in \mathbb{R}^{n \times m}$ and $\mathbf{b} \in \mathbb{R}^n$.

Solution: ■

Exercise 6: Carathéodory's Lemma 5pts

Suppose that $S \subset \mathbb{R}^n$. Show that every element of **conv** S is a convex combination of at most $n+1$ points of S.

Solution: ■

Exercise 7: Strictly Convex Functions 10pts

1. Suppose that $f : \mathbb{R}^n \to \mathbb{R}$ is continuously differentiable. Please show that f is strictly convex if and only if

$$f(\mathbf{y}) > f(\mathbf{x}) + \langle \nabla f(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle, \forall\, \mathbf{x}, \mathbf{y} \in \mathbb{R}^n, \mathbf{x} \neq \mathbf{y}.$$

2. Suppose that $f : \mathbb{R}^n \to \mathbb{R}$ is twice continuously differentiable. Please show that f is strictly convex if

$$\nabla^2 f(\mathbf{x}) \succ 0, \forall\, \mathbf{x} \in \mathbb{R}^n.$$

   Is the converse true? Please show your statement.

Solution: ■

Exercise 8: Strongly Convex Functions 15pts

1. Suppose that f is continuously differentiable. Show that a continuously differentiable function f is strongly convex with parameter $\mu > 0$ if and only if

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

2. Suppose that f is twice continuously differentiable and strongly convex with parameter $\mu > 0$. Please give an interpretation of μ in terms of the eigenvalues of $\nabla^2 f(\mathbf{x})$.

3. (**Lipschitz Continuity**). Suppose that $f : \mathbb{R}^n \to \mathbb{R}$ is twice continuously differentiable, and the gradient of f is Lipschitz continuous, i.e.,

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

where $L > 0$ is the Lipschitz constant. Please give an interpretation of L in terms of the eigenvalues of $\nabla^2 f(\mathbf{x})$.

Solution: ■

Exercise 9: Epigraph 10pts

1. Given a real-valued function $f : \mathbb{R}^n \to (-\infty, +\infty]$. Show that the following are equivalent:
   (a) The epigraph **epi** f is closed.
   (b) The α-level set C_α is closed for any value of α.
2. Let the function $f : \mathbb{R}^n \times \mathbb{S}^n_{++} \to \mathbb{R}$ be defined as

$$f(\mathbf{x}, \mathbf{Y}) = \mathbf{x}^\top \mathbf{Y}^{-1} \mathbf{x},$$

   where $\mathbb{S}^n_{++}$ denotes the set of symmetric positive-definte $n \times n$ matrices. Please show that f is convex.

Solution: ∎

Exercise 10: Operations that Preserve Convexity 25pts

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

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

Show that F is convex.

2. Let $f_i : \mathbb{R}^n \to (-\infty, +\infty]$, $i = 1, \ldots, m$ be given convex functions. 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. (a) Let $f_i : \mathbb{R}^n \to (-\infty, +\infty]$ be given convex functions for $i \in I$, where I is an arbitrary index set. Show that

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

is convex.

(b) Consider the function $f(\mathbf{X}) = \lambda_{\max}(\mathbf{X})$, with $\mathbf{dom}\, f = \mathbb{S}^n$, where $\lambda_{\max}(\mathbf{X})$ is the largest eigenvalue of $\mathbf{X}$ and $\mathbb{S}^n$ is a set of $n \times n$ real symmetric matrices. Please show that f is a convex function.

4. Suppose that the training set is $\{(\mathbf{x}_i, y_i)\}_{i=1}^m$, where $\mathbf{x}_i \in \mathbb{R}^d$ is the i^{th} data instance and $y_i \in \mathbb{R}$ is the corresponding label.

(a) Lasso refers to the regression problem as follows:

$$\min_{\boldsymbol{\beta}} \frac{1}{2m} \|\mathbf{X}\boldsymbol{\beta} - \mathbf{y}\|_2^2 + \lambda \|\boldsymbol{\beta}\|_1,$$

where $\mathbf{X} \in \mathbb{R}^{m \times d}$ with its i^{th} row being $\mathbf{x}_i^\top$, $\boldsymbol{\beta} \in \mathbb{R}^d$, and $\lambda > 0$ is the regularization parameter.

(b) Logistic regression refers to the problem as follows:

$$\min_{\mathbf{w}, b} \frac{1}{m} \sum_{i=1}^{m} \log\left(1 + \exp(-y_i(\langle \mathbf{w}, \mathbf{x}_i \rangle - b))\right),$$

where $y_i \in \{1, -1\}$, $\mathbf{w} \in \mathbb{R}^d$, and $b \in \mathbb{R}$.

Show that the objective functions in the above problems are convex.

Solution: ■