Coursework (2) for *Introductory Lectures on Optimization*

Zhou Nan
3220102535

October 29, 2024

Excercise 1. For the function $f(x) : \mathbb{R}^n \to \mathbb{R}^m$, please write down the zeroth-order Taylor expansion with an integral remainder term.

Solution of Excercise 1: The zeroth-order Taylor expansion with an integral remainder term is

$$f(x) = f(x_0) + \int_0^1 \nabla f(x_0 + \tau(x - x_0)) \cdot (x - x_0) d\tau$$

$\square$

Excercise 2. Please write down the definition of the p-norm for a n-dimensional real vector.

Solution of Excercise 2: Suppose a vector $\mathbf{x}$ is a n-dimensional real vector. $\mathbf{x} = (x_1, x_2, \cdots, x_n)$

$$\|\mathbf{x}\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p}.$$

$\square$

Excercise 3. Please write down the definition of the matrix norms induced by vector p-norms.

Solution of Excercise 3: For a given matrix $A \in \mathbb{R}^{m \times n}$, and a vector p-norm $\|\mathbf{x}\|_p$

$$\|A\|_p = \sup_{\mathbf{x} \neq 0} \frac{\|A\mathbf{x}\|_p}{\|\mathbf{x}\|_p}.$$

$\square$

Excercise 4. Let A be an $n \times n$ symmetric matrix. Proof that A is positive semidefinite if and only if all eigenvalues of A are nonnegative. Moreover, A is positive definite if and only if all eigenvalues of A are positive.

Proof of Excercise 4: According to the spectural theorem,

$$A = Q\Lambda Q^T$$

- Matrix Q:

  Q is an orthogonal matrix (or unitary matrix in the complex case), meaning that its columns are orthonormal vectors (i.e., $Q^T Q = I$, where I is the identity matrix). The columns of Q are the eigenvectors of the matrix A.

- Matrix Λ:

  Λ is a diagonal matrix that contains the eigenvalues of A on its diagonal. The eigenvalues correspond to the scaling factors associated with their respective eigenvectors.

1. If A is positive semidefinite, then all eigenvalues of A are nonnegative.

   Since A is positive semidefinite, then for any n-dimensional vector $\mathbf{x}$, we have

   $$\mathbf{x}^T A\mathbf{x} \geq 0$$
   $$\mathbf{x}^T Q\Lambda Q^T \mathbf{x} \geq 0$$

   Suppose there is a eigenvalue λ_i is negative, then we construct a n-dimensional vector $\mathbf{y}$ such that $\mathbf{y} = (0, 0, \cdots, 0, 1, 0, \cdots, 0)$, $y_i = 1$.

   Let $\mathbf{x} = Q\mathbf{y}$, then

   $$\mathbf{x}^T A\mathbf{x} = (Q\mathbf{y})^T A(Q\mathbf{y}) = (Q\mathbf{y})^T Q\Lambda Q^T (Q\mathbf{y}) = \mathbf{y}^T Q^T Q\Lambda Q^T Q\mathbf{y} = \mathbf{y}^T \Lambda\mathbf{y} = \lambda_i < 0$$

   which is a contradiction.

2. If all eigenvalues of A are nonnegative, then A is positive semidefinite.

   For any n-dimensional vector $\mathbf{x}$, we have

   $$\mathbf{x}^T A\mathbf{x} = \mathbf{x}^T Q\Lambda Q^T \mathbf{x} = (Q\mathbf{x})^T \Lambda(Q\mathbf{x}) = \sum_{i=1}^{n} \lambda_i (Q^T \mathbf{x})_i^2 \geq 0$$

   Therefore, A is positive semidefinite.

3. If A is positive definite, then all eigenvalues of A are positive.

   Since A is positive definite, then for any n-dimensional vector $\mathbf{x}$, we have

   $$\mathbf{x}^T A\mathbf{x} > 0$$

   Suppose there is a eigenvalue λ_i is not positive, then we construct a n-dimensional vector $\mathbf{y}$ such that $\mathbf{y} = (0, 0, \cdots, 0, 1, 0, \cdots, 0)$, $y_i = 1$.

   Let $\mathbf{x} = Q\mathbf{y}$, then

   $$\mathbf{x}^T A\mathbf{x} = (Q\mathbf{y})^T A(Q\mathbf{y}) = \mathbf{y}^T \Lambda\mathbf{y} = \lambda_i \leq 0$$

   which is a contradiction.

4. If all eigenvalues of A are positive, then A is positive definite.

   For any n-dimensional vector $\mathbf{x}$, we have

$$\mathbf{x}^T A\mathbf{x} = \mathbf{x}^T Q\Lambda Q^T \mathbf{x} = (Q\mathbf{x})^T \Lambda (Q\mathbf{x}) = \sum_{i=1}^{n} \lambda_i (Q^T \mathbf{x})_i^2 > 0$$

   Therefore, A is positive definite.

$\square$

Excercise 5. Suppose that $f : \mathbb{R}^n \to \mathbb{R}$ is convex and upper bounded. Show that f must be a constant function.

Proof of Excercise 5: Suppose that f is not a constant function. Then we have: $\exists \mathbf{x}_1 \in \mathbb{R}^n, \mathbf{x}_2 \in \mathbb{R}^n$ such that $f(\mathbf{x}_1) > f(\mathbf{x}_2)$

Since $f : \mathbb{R}^n \to \mathbb{R}$ is convex, we have:

$$f(\mathbf{x}_2) \geq f(\mathbf{x}_1) + \langle \nabla f(\mathbf{x}_1), \mathbf{x}_2 - \mathbf{x}_1 \rangle$$

Because $f(\mathbf{x}_1) > f(\mathbf{x}_2)$, then $\langle \nabla f(\mathbf{x}_1), \mathbf{x}_2 - \mathbf{x}_1 \rangle < 0$

Since $f : \mathbb{R}^n \to \mathbb{R}$ is upper bounded, we have:

$$\exists M, \forall \mathbf{x} \in \mathbb{R}^n, f(\mathbf{x}) \leq M$$

Let G be $\dfrac{M + 1 - f(\mathbf{x}_1)}{\langle \nabla f(\mathbf{x}_1), \mathbf{x}_2 - \mathbf{x}_1 \rangle}$, let $\mathbf{y} = (1 - G)\mathbf{x}_1 + G\mathbf{x}_2$, then

$$\begin{aligned}
f(\mathbf{y}) &\geq f(\mathbf{x}_1) + \langle \nabla f(\mathbf{x}_1), \mathbf{y} - \mathbf{x}_1 \rangle \\
&= f(\mathbf{x}_1) + \langle \nabla f(\mathbf{x}_1), (1 - G)\mathbf{x}_1 + G\mathbf{x}_2 - \mathbf{x}_1 \rangle \\
&= f(\mathbf{x}_1) + \langle \nabla f(\mathbf{x}_1), G(\mathbf{x}_2 - \mathbf{x}_1) \rangle \\
&= f(\mathbf{x}_1) + G\langle \nabla f(\mathbf{x}_1), \mathbf{x}_2 - \mathbf{x}_1 \rangle \\
&= f(\mathbf{x}_1) + M + 1 - f(\mathbf{x}_1) \\
&= M + 1
\end{aligned}$$

Therefore there exists $\mathbf{y} \in \mathbb{R}^n$ such that $f(\mathbf{y}) \geq M + 1$, which is a contradiction.

Therefore f is a constant function. $\square$