Coursework (1) for *Introductory Lectures on Optimization*

Hongjue Zhao
3190104515

Oct. 1, 2022

Excercise 1. 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 1: Suppose that

$$Q = \{x \in \mathbb{R}^n \mid x \neq 0\},$$

and the eigenvalues of A are $\lambda_1, \lambda_2, \dots, \lambda_n$. Because A is a symmetric matrix, A is also orthogonally diagnoalizable, which means

$$\exists\ P \in \mathbb{R}^{n\times n},\ \text{s.t. } P^\top AP = \text{diag}(\lambda_1, \dots, \lambda_n) \text{ and } P^\top P = E. \tag{1}$$

Here we define $\Lambda \overset{\text{def}}{=} \text{diag}(\lambda_1, \dots, \lambda_n)$. In this case, for any real vector $x \in Q$, we can get

$$\begin{aligned} x^\top Ax &= x^\top P^\top \Lambda P x = (Px)^\top \Lambda (Px) \\ &= y^\top \Lambda y = \sum_{i=1}^{n} \lambda_i y_i^2. \end{aligned} \tag{2}$$

where $y = Px$. According to Eq. 2 and related definition[3], $x^\top Ax > 0$ if and only if $\lambda_i > 0$ ($i = 1, \dots, n$), which also means A is positive definite. $x^\top Ax \geq 0$ if and only if $\lambda_i \geq 0$ ($i = 1, \dots, n$), which also means A is positive semidefinite. $\square$

Excercise 2. For the performance analysis of the Uniform Grid Method, Proof that

$$\left(\left\lfloor \frac{L}{2\epsilon} \right\rfloor + 2\right)^n, \text{ and } \left(\left\lfloor \frac{L}{2\epsilon} \right\rfloor\right)^n,$$

coincide up to an absolute constant multiplicative factor if $\epsilon \leq O(\frac{L}{n})$.

Proof of Excercise 2: Since $\epsilon \leq O\left(\frac{L}{n}\right)$, there exists $M > 0$ that satisfies

$$\epsilon \leq M\left|\frac{L}{n}\right| = M\frac{L}{n} \Leftrightarrow \frac{L}{2\epsilon} \geq \frac{n}{2M}. \tag{3}$$

Obviously, according to Eq. 3, we can get

$$1 \leq \frac{\left(\left\lfloor \frac{L}{2\epsilon} \right\rfloor + 2\right)^n}{\left(\left\lfloor \frac{L}{2\epsilon} \right\rfloor\right)^n} = 1 + \sum_{k=1}^{n} \frac{c_k}{\left(\left\lfloor \frac{L}{2\epsilon} \right\rfloor\right)^k} \leq 1 + \sum_{k=1}^{n} \frac{c_k}{\left(\left\lfloor \frac{n}{2M} \right\rfloor\right)^k}, \tag{4}$$

in which $c_k = C_n^k 2^k$. As $n \to \infty$, we can also get

$$\lim_{n\to\infty} \left[1 + \sum_{k=1}^{n} \frac{c_k}{\left(\left\lfloor \frac{n}{2M} \right\rfloor\right)^k} \right] = 1. \tag{5}$$

Above all, according to Eq. 4 and Eq. 5, we can get the result based on Sandwich theorem:

$$\lim_{n\to\infty} \frac{\left(\left\lfloor \frac{L}{2\epsilon} \right\rfloor + 1\right)^n}{\left(\left\lfloor \frac{L}{2\epsilon} \right\rfloor\right)^n} = 1.$$

Therefore, the statement is proofed. $\square$

Excercise 3. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a strictly convex function. Let $x_i \in \mathbb{R}^n$ and $\lambda_i > 0$ for $i = 1, 2, \ldots, k$ such that $\sum_{i=1}^{k} \lambda_i = 1$. If

$$f(\lambda_1 x_1 + \cdots + \lambda_k x_k) = \lambda_1 f(x_1) + \cdots + \lambda_k f(x_k),$$

then show that $x_1 = x_2 = \cdots = x_k$.

Proof of Excercise 3: According to related definition about convex function[1], for strictly convex function $f : \mathbb{R}^n \to \mathbb{R}$

$$f(\lambda_1 x_1 + \lambda_2 x_2) \le \lambda_1 f(x_1) + \lambda_2 f(x_2), \tag{6}$$

where $x_1,\ x_2 \in \mathbb{R}^n$ and $\lambda_1 + \lambda_2 = 1$. Equality in Eq. 6 holds if and only if $x_1 = x_2$.

Suppose the following inequality is true:

$$f(\lambda_1 x_1 + \cdots + \lambda_k x_k) \le \lambda_1 f(x_1) + \cdots + \lambda_k f(x_k) \tag{7}$$

where $x_1, \ldots, x_k \in \mathbb{R}^n$ and $\lambda_1 + \cdots + \lambda_k = 1$. Therefore, according to Eq. 6, we can get

$$\begin{aligned} f(\lambda_1 x_1 + \cdots + \lambda_k x_k + \lambda_{k+1} x_{k+1}) &= f\left((1-\lambda_{k+1}) \sum_{i=1}^{k} \frac{\lambda_i}{1-\lambda_{k+1}} x_i + \lambda_{k+1} x_{k+1} \right) \\ &\le (1-\lambda_{k+1}) f\left(\sum_{i=1}^{k} \frac{\lambda_i}{1-\lambda_{k+1}} x_i \right) + \lambda_{k+1} f(x_{k+1}). \end{aligned} \tag{8}$$

According to Eq. 7, we can get

$$(1-\lambda_{k+1}) f\left(\sum_{i=1}^{k} \frac{\lambda_i}{1-\lambda_{k+1}} x_i \right) \le (1-\lambda_{k+1}) \sum_{i=1}^{k} \frac{\lambda_i}{1-\lambda_{k+1}} f(x_i) = \sum_{i=1}^{k} \lambda_i f(x_i). \tag{9}$$

Based on Eq. 8, 9, finally we can get

$$f(\lambda_1 x_1 + \cdots + \lambda_k x_k + \lambda_{k+1} x_{k+1}) \le \lambda_1 f(x_1) + \cdots + \lambda_k f(x_k) + \lambda_{k+1} f(x_{k+1}). \tag{10}$$

Above all, Eq. 7 holds.

Since f is strictly convex, in Eq. 7:

1. When $k = 2$, the inequality is strict since it's the definition of strictly convex functions.
2. When $k = m + 1$, if $x_1, \ldots, x_{m+1}$ are not all equal, then the inequality in Eq. 9 should be strict. If $x_1 = x_2 = \cdots = x_m \ne x_{m+1}$, which also means $x_{m+1} \ne \sum_{i=1}^{m} \frac{\lambda_i}{1-\lambda_{m+1}} x_i$, then equality in Eq. 8 should be strict.

Above all, inequality in Eq. 7 should be strict when $x_1, \dots, x_k$ are not all equal. Thus, equality holds if and only if $x_1 = x_2 = \cdots = x_k$. $\square$

Excercise 4. Proof that the following univariate functions are in the set of $\mathcal{F}^1(\mathbb{R})$:

$$\begin{aligned} f(x) &= e^x, \\ f(x) &= |x|^p,\ p > 1, \\ f(x) &= \frac{x^2}{1+|x|}, \\ f(x) &= |x| - \ln(1+|x|). \end{aligned}$$

Proof of Excercise 4: Since all these functions are univariate functions, so $\nabla f(x) = f'(x)$, $\langle \nabla f(x) - \nabla f(y),\ x - y\rangle = (f'(x) - f'(y))(x-y)$

For $f(x) = e^x$, $f'(x) = f''(x) = e^x \geq 0$. Therefore, $f(x) = e^x$ is convex and in the set $\mathcal{F}^1(\mathbb{R})$.

For $f(x) = |x|^p,\ p > 1$, obviously, it's continuous in $\mathbb{R}$. In the meanwhile,

$$f'(x) = \begin{cases} px^{p-1}, & x \geq 0, \\ -p(-x)^{p-1}, & x < 0, \end{cases},\ f''(x) = \begin{cases} p(p-1)x^{p-2}, & x \geq 0, \\ p(p-1)(-x)^{p-2}, & x < 0, \end{cases}.$$

We can find that

$$\lim_{x\to 0^+} f'(x) = \lim_{x\to 0^-} f'(x) = 0,\ \lim_{x\to 0^+} f''(x) = \lim_{x\to 0^-} f''(x) = 0.$$

So it's twice differentiable in $\mathbb{R}$. Since $\forall x \in \mathbb{R},\ f''(x) > 0$, $f(x) = |x|^p,\ p > 1$ is convex and in the set $\mathcal{F}^1(\mathbb{R})$.

For $f(x) = \frac{x^2}{1+|x|}$, obviously, it's continuous in $\mathbb{R}$. In the meanwhile,

$$f'(x) = \begin{cases} \frac{2x}{1+x} + \left(\frac{x}{1+x}\right)^2, & x > 0 \\ \frac{2x}{1-x} + \left(\frac{x}{1-x}\right)^2, & x < 0 \end{cases},\ f''(x) = \begin{cases} 2\left(1 + \frac{x}{1+x}\right)\frac{1}{(1+x)^2}, & x > 0 \\ 2\left(1 - \frac{x}{1-x}\right)\frac{1}{(1-x)^2}, & x < 0 \end{cases}.$$

We can find that

$$\lim_{x\to 0^+} f'(x) = \lim_{x\to 0^-} f'(x) = 0,\ \lim_{x\to 0^+} f''(x) = \lim_{x\to 0^-} f''(x) = 2.$$

So it's twice differentiable in $\mathbb{R}$. Since $\forall x \in \mathbb{R},\ f''(x) > 0$, $f(x) = \frac{x^2}{1+|x|}$ is convex and in the set $\mathcal{F}^1(\mathbb{R})$.

For $f(x) = |x| - \ln(1+|x|)$, obviously it's continuous in $\mathbb{R}$. In the meanwhile:

$$f'(x) = \begin{cases} 1 - \frac{1}{1+x}, & x > 0 \\ -1 + \frac{1}{1-x}. & x < 0 \end{cases},f''(x) = \begin{cases} \frac{1}{(1+x)^2}, & x > 0 \\ \frac{1}{(1-x)^2}, & x < 0 \end{cases}.$$

We can find that

$$\lim_{x\to 0^+} f'(x) = \lim_{x\to 0^-} f'(x) = 0,\ \lim_{x\to 0^+} f''(x) = \lim_{x\to 0^-} f''(x) = 1.$$

So it's twice differentiable in $\mathbb{R}$. Since $\forall x \in \mathbb{R},\ f''(x) > 0$, $f(x) = |x| - \ln(1+|x|)$ is convex and in the set $\mathcal{F}^1(\mathbb{R})$. $\square$

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

Proof of Excercise 5: For any $x, y \in \mathbb{R}^n$ and $\alpha \in (0, 1)$:

1. f is convex $\rightarrow f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha)f(y)$;
2. f is concave $\rightarrow f(\alpha x + (1-\alpha)y) \geq \alpha f(x) + (1-\alpha)f(y)$;

Above all, we can get

$$f(\alpha x + (1-\alpha)y) = \alpha f(x) + (1-\alpha)f(y). \tag{11}$$

According to Eq. 11, f must be a affine function. □

Excercise 6. Suppose that $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is convex and upper bounded. Show that f must be a constant function.

Proof of Excercise 6: Suppose that f is not a constant function, i.e., $\exists\, x, y \in \mathbb{R}^n$ s.t. $f(y) > f(x)$. Without loss of generality, here we can assume that $x < y$.

In the meanwhile, there must be z that satisfies $z > y$. Since the convexity of f, we can get

$$\frac{f(z) - f(x)}{z - x} \geq \frac{f(y) - f(x)}{y - x} \Rightarrow f(z) \geq f(x) + \frac{f(y) - f(x)}{y - x}(z - x). \tag{12}$$

Let $z \rightarrow +\infty$ in Eq. 12. In this case, $f(z) \rightarrow +\infty$, and f is not upper bounded. Thus, if f is convex and upper bounded, f must be a constant function[2]. □

References

[1] Stephen Boyd, Stephen P Boyd, and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

[2] hmakholm left over Monica (https://math.stackexchange.com/users/14366/hmakholm-left-over monica). Show bounded and convex function on $\setminus$ is constant. Mathematics Stack Exchange. URL:https://math.stackexchange.com/q/518103 (version: 2013-10-07).

[3] Gilbert Strang. *Linear algebra and its applications*. Belmont, CA: Thomson, Brooks/Cole, 2006.