

1 A Complete Solution Guide to Introduction to Nonlinear  
2 Optimization Theory, Algorithms, and Applications with  
3 MATLAB

4 Youming Zhao  
Email: [youming0.zhao@gmail.com](mailto:youming0.zhao@gmail.com)

5 First draft: May 24, 2022 Last update: September 9, 2023

6 **Contents**

|    |   |           |
|----|---|-----------|
| 7  | <b>1 Chapter 1 Mathematical Preliminaries</b>                           | <b>2</b>  |
| 8  | 1.1 Some important concepts . . . . .                                   | 2         |
| 9  | 1.1.1 Induced matrix norm and several equivalent definitions . . . . .  | 2         |
| 10 | 1.1.2 Accumulation point . . . . .                                      | 3         |
| 11 | 1.1.3 Closed set . . . . .  | 3         |
| 12 | 1.1.4 Boundary point . . . . .  | 3         |
| 13 | 1.1.5 Closure . . . . .   | 3         |
| 14 | 1.1.6 Interior point and interior of a set . . . . .                    | 3         |
| 15 | 1.1.7 De Morgan's Law/Theorem . . . . .                                 | 4         |
| 16 | 1.2 Exercises . . . . .   | 4         |
| 17 | <b>2 Chapter 2 Optimality Conditions for Unconstrained Optimization</b> | <b>15</b> |
| 18 | <b>3 Chapter 3 Least Squares</b>  | <b>17</b> |
| 19 | <b>4 Chapter 4 The Gradient Method</b>                                  | <b>17</b> |
| 20 | <b>5 Chapter 5 Newton's Method</b>                                      | <b>18</b> |
| 21 | <b>6 Chapter 6 Convex Sets</b>  | <b>18</b> |
| 22 | <b>7 Chapter 7 Convex Functions</b>                                     | <b>18</b> |
| 23 | <b>8 Chapter 8 Convex Optimization</b>                                  | <b>22</b> |
| 24 | <b>9 Chapter 9 Optimization over a Convex Set</b>                       | <b>23</b> |
| 25 | <b>Bibliography</b>   | <b>24</b> |

# Chapter 1 Mathematical Preliminaries

## 1.1 Some important concepts

### 1.1.1 Induced matrix norm and several equivalent definitions

Here we introduce the definition of the induced matrix norm from the textbook. That is, the induced matrix norm  $\|\mathbf{A}\|_{a,b}$  is defined by

$$\|\mathbf{A}\|_{a,b} = \max_{\mathbf{x}} \{\|\mathbf{Ax}\|_b : \|\mathbf{x}\|_a \leq 1\}. \quad (1)$$

$\|\mathbf{A}\|_{a,b}$  can also be computed in the following alternative ways (Horn and Johnson, 2013, p. 343, Definition 5.6.1):

$$\|\mathbf{A}\|_{a,b} = \max_{\mathbf{x}} \{\|\mathbf{Ax}\|_b : \|\mathbf{x}\|_a = 1\} = \max_{\|\mathbf{x}\|_a \neq 0} \frac{\|\mathbf{Ax}\|_b}{\|\mathbf{x}\|_a}. \quad (2)$$

Now we show that they are valid alternatives of (1) by proving two lemmas. The first alternative is exactly the following lemma.

**Lemma 1.1.** *The maximum points  $\mathbf{x}^*$  of the RHS of (1) must satisfy  $\|\mathbf{x}^*\|_a = 1$ .*

*Proof.* We will prove it by contradiction. Given  $\mathbf{A} \neq \mathbf{0}$ , it is obvious that  $\mathbf{x}^* \neq \mathbf{0}$  must hold, otherwise  $\|\mathbf{Ax}^*\|_b = 0$  which is the minimum value and it is easy to find an  $\mathbf{x}$  such that  $\|\mathbf{Ax}\|_b > 0$ . Suppose that the maximum points satisfy  $\|\mathbf{x}^*\|_a < 1$ , then there exists real numbers  $k$  such that  $\|k\mathbf{x}^*\|_a = 1$  in which  $|k| = 1/\|\mathbf{x}^*\|_a > 1$ . Let  $\mathbf{y} = k\mathbf{x}^*$ , then we get

$$\|\mathbf{Ay}\|_b = \|\mathbf{A}(k\mathbf{x}^*)\|_b = |k|\|\mathbf{Ax}^*\|_b > \|\mathbf{Ax}^*\|_b \quad (3)$$

which contradicts that  $\mathbf{x}^*$  are the maximum points. Thus,  $\|\mathbf{x}^*\|_a = 1$  holds.  $\square$

We directly present the second alternative as a lemma as follows and prove it through Lemma 1.1.

**Lemma 1.2.** *For any  $\mathbf{x} \in \mathbb{R}^n$ ,*

$$\|\mathbf{A}\|_{a,b} = \max_{\|\mathbf{x}\|_a \neq 0} \frac{\|\mathbf{Ax}\|_b}{\|\mathbf{x}\|_a}. \quad (4)$$

*Proof.* An equivalent form of Lemma 1.1 is

$$\|\mathbf{A}\|_{a,b} = \max_{\mathbf{y}} \left\{ \frac{\|\mathbf{Ay}\|_b}{\|\mathbf{y}\|_a} : \|\mathbf{y}\|_a = 1 \right\} = \max_{\|\mathbf{y}\|_a = 1} \frac{\|\mathbf{Ay}\|_b}{\|\mathbf{y}\|_a}. \quad (5)$$

By letting  $\mathbf{y} = k\mathbf{x}$  where  $k \in \mathbb{R} \setminus \{0\}$ , we have

$$\|\mathbf{A}\|_{a,b} = \max_{|k|\|\mathbf{x}\|_a = 1} \frac{|k|\|\mathbf{Ax}\|_b}{|k|\|\mathbf{x}\|_a} = \max_{\|\mathbf{x}\|_a = 1/|k|} \frac{\|\mathbf{Ax}\|_b}{\|\mathbf{x}\|_a} = \max_{\|\mathbf{x}\|_a \neq 0} \frac{\|\mathbf{Ax}\|_b}{\|\mathbf{x}\|_a} \quad (6)$$

where the last equality follows from that  $k$  is an arbitrary nonnegative real number. This completes our proof.  $\square$

The textbook gives a result about the induced matrix norm without a proof right after its definition. Here, we will present it as a proposition with a proof. The proof is an immediate result of Lemma 4.

**Proposition 1.3.** *For any  $\mathbf{x} \in \mathbb{R}^n$  the inequality*

$$\|\mathbf{Ax}\|_b \leq \|\mathbf{A}\|_{a,b} \|\mathbf{x}\|_a \quad (7)$$

*holds.*

*Proof.* According to Lemma 4, for any  $\mathbf{x} \neq \mathbf{0}$ , it follows that

$$\frac{\|\mathbf{Ax}\|_b}{\|\mathbf{x}\|_a} \leq \|\mathbf{A}\|_{a,b} \iff \|\mathbf{Ax}\|_b \leq \|\mathbf{A}\|_{a,b} \|\mathbf{x}\|_a \quad (8)$$

40 completing the proof. □

### 41 1.1.2 Accumulation point

**Definition 1.4 (accumulation points).** *If any open ball of a point  $x$  contains infinitely many points of a set  $S$ , then  $x$  is called an accumulation point of  $S$ . The set of all accumulation points of  $S$  is denoted by  $S'$ .*

42

### 43 1.1.3 Closed set

44 We describe the definition of closed sets in a slightly different way than the textbook. However, in  
45 essence, they are the same thing.

**Definition 1.5 (closed sets).** *If a set  $S$  contains all of its accumulation points, then we call  $S$  a closed set.*

46

### 47 1.1.4 Boundary point

**Definition 1.6 (boundary points).** *Given a set  $U \subseteq \mathbb{R}^n$ , a **boundary point** of  $U$  is a point  $\mathbf{x} \in \mathbb{R}^n$  satisfying the following: any neighborhood of  $\mathbf{x}$  contains at least one point in  $U$  and at least one point in its complement  $U^c$ . The set of all boundary points of a set is denoted by  $\text{bd}(U)$  or  $\partial U$  and is called the boundary of  $U$ .*

48

### 49 1.1.5 Closure

**Definition 1.7 (closure of a set).** *The closure of a set  $U \subseteq \mathbb{R}^n$  is the smallest closed set containing  $U$ :*

$$\text{cl}(U) = \bigcap \{T : U \subseteq T, T \text{ is closed}\}. \quad (9)$$

*Another equivalent definition of  $\text{cl}(U)$  is given by*

$$\text{cl}(U) = U \cup \text{bd}(U). \quad (10)$$

50

51 The closure set is indeed a closed set as an intersection of closed sets (see Exercise 1.16(ii)).

### 52 1.1.6 Interior point and interior of a set

**Definition 1.8 (interior points).** *Given a set  $U \subseteq \mathbb{R}^n$ , a point  $\mathbf{c} \in U$  is an interior point of  $U$  if there exists  $r > 0$  for which  $B(\mathbf{c}, r) \subseteq U$ .*

53

**Definition 1.9 (interior of a set).** *The set of all interior points of a given set  $U$  is called the interior of a set and is denoted by  $\text{int}(U)$ :*

$$\text{int}(U) = \{\mathbf{x} \in U : B(\mathbf{x}, r) \subseteq U \text{ for some } r > 0\}. \quad (11)$$

### 1.1.7 De Morgan's Law/Theorem

Here we present a generalized form of De Morgan's Law which is also known as De Morgan's Theorem from Wikipedia<sup>1</sup>.

**Theorem 1.10 (De Morgan's Law/Theorem).**

$$\left( \bigcup_{i \in I} A_i \right)^c = \bigcap_{i \in I} A_i^c \quad (12)$$

$$\left( \bigcap_{i \in I} A_i \right)^c = \bigcup_{i \in I} A_i^c \quad (13)$$

where  $I$  is some, possibly countably or uncountably infinite, indexing set.

## 1.2 Exercises

### Exercise 1.1

Show that  $\|\cdot\|_{1/2}$  is not a norm.

*Proof.* To show that a function is not a norm, it suffices to find a counterexample which does not satisfy at least one of the three properties of a norm. For  $\|\cdot\|_{1/2}$ , we let

$$\mathbf{x} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Then we have

$$\|\mathbf{x} + \mathbf{y}\|_{1/2} = \left\| \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\|_{1/2} = (\sqrt{1} + \sqrt{1})^2 = 4$$

$$\|\mathbf{x}\|_{1/2} = (\sqrt{1} + \sqrt{0})^2 = 1$$

$$\|\mathbf{y}\|_{1/2} = (\sqrt{0} + \sqrt{1})^2 = 1$$

However,

$$\|\mathbf{x} + \mathbf{y}\|_{1/2} = 4 > \|\mathbf{x}\|_{1/2} + \|\mathbf{y}\|_{1/2} = 1 + 1 = 2.$$

Hence,  $\|\cdot\|_{1/2}$  does not satisfy the triangle inequality. This completes the proof.  $\square$

In fact, when  $0 < p < 1$ ,  $\|\cdot\|_p$  satisfies the reverse of Minkowski's inequality within the domain of  $\mathbb{R}_+^n$ . Formally, we have the following theorem.

<sup>1</sup>[https://en.wikipedia.org/wiki/De\\_Morgan%27s\\_laws](https://en.wikipedia.org/wiki/De_Morgan%27s_laws)

**Theorem 1.11 (reversed Minkowski's inequality).** For any  $\mathbf{x}, \mathbf{y} \in \mathbb{R}_+^n$  and  $0 < p < 1$ , the following inequality

$$\|\mathbf{x} + \mathbf{y}\|_p \geq \|\mathbf{x}\|_p + \|\mathbf{y}\|_p$$

holds.

The following proof largely follows [Jax \(2016\)](#) but in greater detail.

*Proof.* Obviously, the claim holds when either  $\mathbf{x} = 0$  or  $\mathbf{y} = 0$ . We only need to consider the case when  $\mathbf{x} \neq 0$  and  $\mathbf{y} \neq 0$ , which guarantees  $\|\mathbf{x} + \mathbf{y}\|_p \neq 0$ . Let  $f(x) = x^p$  with  $x > 0$  and  $0 < p < 1$ . Since  $f''(x) = p(p-1)x^{p-2} < 0$  for any  $x > 0$ ,  $f(x)$  is concave. Thus, we have

$$\begin{aligned} (x_i + y_i)^p &= \left( t \cdot \frac{x_i}{t} + (1-t) \cdot \frac{y_i}{1-t} \right)^p, \quad 0 < t < 1, i \in \{1, 2, \dots, n\} \\ &\geq t \cdot \frac{x_i^p}{t^p} + (1-t) \cdot \frac{y_i^p}{(1-t)^p}. \end{aligned}$$

Taking summation over  $i$  gives

$$\begin{aligned} \sum_{i=1}^n (x_i + y_i)^p &\geq t \sum_{i=1}^n \frac{x_i^p}{t^p} + \frac{y_i^p}{(1-t)^p} \\ \|\mathbf{x} + \mathbf{y}\|_p^p &\geq t \frac{\|\mathbf{x}\|_p^p}{t^p} + (1-t) \frac{\|\mathbf{y}\|_p^p}{(1-t)^p} \end{aligned}$$

Letting  $t = \frac{\|\mathbf{x}\|_p}{\|\mathbf{x}\|_p + \|\mathbf{y}\|_p}$  yields

$$\begin{aligned} \|\mathbf{x} + \mathbf{y}\|_p^p &\geq t \frac{\|\mathbf{x}\|_p^p}{(\|\mathbf{x}\|_p + \|\mathbf{y}\|_p)^p} + (1-t) \frac{\|\mathbf{y}\|_p^p}{(\|\mathbf{x}\|_p + \|\mathbf{y}\|_p)^p} \\ &= t(\|\mathbf{x}\|_p + \|\mathbf{y}\|_p)^p + (1-t)(\|\mathbf{x}\|_p + \|\mathbf{y}\|_p)^p \\ &= (\|\mathbf{x}\|_p + \|\mathbf{y}\|_p)^p \\ \implies \|\mathbf{x} + \mathbf{y}\|_p &\geq \|\mathbf{x}\|_p + \|\mathbf{y}\|_p, \end{aligned}$$

as desired. □

*Remark 1.12.* You may observe that the reversed Minkowski's inequality does not hold when  $\mathbf{x} = -\mathbf{y} \neq 0$ . The reason is that in the above proof, the condition  $x_i, y_i \geq 0, \forall i$  is required to ensure that  $f(x)$  is concave and well defined. Concretely speaking,  $\sqrt[3]{x}$  is convex on  $\mathbb{R}_-$  and  $\sqrt[4]{x}$  is not well defined on  $\mathbb{R}_-$ . Hence, the reversed Minkowski's inequality only works for both vectors with nonnegative entries. Note that Minkowski's inequality works not only for  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$  but also for  $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$ .

## Extensions

Since  $\|\cdot\|_0$  does not satisfy the positive homogeneity, it is not a true norm.

## Exercise 1.2

Prove that for any  $\mathbf{x} \in \mathbb{R}^n$  one has

$$\|\mathbf{x}\|_\infty = \lim_{p \rightarrow \infty} \|\mathbf{x}\|_p.$$

*Proof.* Since the definitions  $\|\mathbf{x}\|_\infty \equiv \max_{i=1,2,\dots,n} |x_i|$  and  $\|\mathbf{x}\|_p \equiv \sqrt[p]{\sum_{i=1}^n |x_i|^p}$ , we only need to show  $\lim_{p \rightarrow \infty} \|\mathbf{x}\|_p = \max_{i=1,2,\dots,n} |x_i|$ . Given any  $\mathbf{x} \in \mathbb{R}^n$  where  $n$  is a finite positive integer, we have

$$\begin{aligned}
\lim_{p \rightarrow \infty} \sqrt[p]{\left(\max_{i=1,2,\dots,n} |x_i|\right)^p} &\leq \lim_{p \rightarrow \infty} \sqrt[p]{\sum_{i=1}^n |x_i|^p} \leq \lim_{p \rightarrow \infty} \sqrt[p]{\left(n \cdot \max_{i=1,2,\dots,n} |x_i|\right)^p} \\
&\Downarrow \\
\max_{i=1,2,\dots,n} |x_i| &\leq \lim_{p \rightarrow \infty} \sqrt[p]{\sum_{i=1}^n |x_i|^p} \leq \lim_{p \rightarrow \infty} \underbrace{\sqrt[p]{n}}_{=1} \cdot \max_{i=1,2,\dots,n} |x_i| \\
&\Downarrow \\
\max_{i=1,2,\dots,n} |x_i| &\leq \lim_{p \rightarrow \infty} \sqrt[p]{\sum_{i=1}^n |x_i|^p} \leq \max_{i=1,2,\dots,n} |x_i| \\
&\Downarrow \\
\lim_{p \rightarrow \infty} \sqrt[p]{\sum_{i=1}^n |x_i|^p} &= \max_{i=1,2,\dots,n} |x_i|.
\end{aligned}$$

71

□

72 This completes our proof.

### Exercise 1.3

Show that for any  $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$

$$\|\mathbf{x} - \mathbf{z}\| \leq \|\mathbf{x} - \mathbf{y}\| + \|\mathbf{y} - \mathbf{z}\|.$$

73

*Proof.* Here,  $\|\cdot\|$  refers to the vector norm  $\|\cdot\|_2$  whose subscript is frequently omitted for brevity. By the definition of the vector norm,  $\|\cdot\|_2$  satisfies the triangle inequality as follows.

$$\begin{aligned}
\|\mathbf{x} - \mathbf{z}\|_2 &= \|\mathbf{x} - \mathbf{y} + \mathbf{y} - \mathbf{z}\|_2 \\
&\leq \|\mathbf{x} - \mathbf{y}\|_2 + \|\mathbf{y} - \mathbf{z}\|_2
\end{aligned}$$

74 as desired.

□

### Exercise 1.4

Prove the Cauchy-Schwarz inequality (Lemma 1.5)

$$|\mathbf{x}^T \mathbf{y}| \leq \|\mathbf{x}\|_2 \cdot \|\mathbf{y}\|_2, \quad \forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n. \quad (14)$$

Show that equality holds if and only if the vectors  $\mathbf{x}$  and  $\mathbf{y}$  are linearly dependent.

75

*Proof.* This lemma can be concisely proved via the following formula from geometry.

$$\mathbf{x}^T \mathbf{y} = \|\mathbf{x}\|_2 \cdot \|\mathbf{y}\|_2 \cdot \cos \theta \quad (15)$$

where  $\theta$  denotes the angle between  $\mathbf{x}$  and  $\mathbf{y}$ . Since  $|\cos \theta| \leq 1$ , it follows that

$$-\|\mathbf{x}\|_2 \cdot \|\mathbf{y}\|_2 \leq \mathbf{x}^T \mathbf{y} = \|\mathbf{x}\|_2 \cdot \|\mathbf{y}\|_2 \cdot \cos \theta \leq \|\mathbf{x}\|_2 \cdot \|\mathbf{y}\|_2 \quad (16)$$

where the equality holds if and only if  $|\cos \theta| = 1$  which geometrically implies that  $\mathbf{x}$  and  $\mathbf{y}$  are parallel to each other, in other words,  $\mathbf{x}$  and  $\mathbf{y}$  are linearly dependent. If we express (16) in a compact way, then we get

$$|\mathbf{x}^T \mathbf{y}| \leq \|\mathbf{x}\|_2 \cdot \|\mathbf{y}\|_2, \quad \forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n. \quad (17)$$

76 This completes the proof.  $\square$

### Exercise 1.5

Suppose that  $\mathbb{R}^m$  and  $\mathbb{R}^n$  are equipped with norms  $\|\cdot\|_b$  and  $\|\cdot\|_a$ , respectively. Show that the induced matrix norm  $\|\cdot\|_{a,b}$  satisfies the triangle inequality. That is, for any  $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m \times n}$  the inequality

$$\|\mathbf{A} + \mathbf{B}\|_{a,b} \leq \|\mathbf{A}\|_{a,b} + \|\mathbf{B}\|_{a,b} \quad (18)$$

holds.

77

*Proof.* By the definition of the induced norm, namely (1),

$$\|\mathbf{A} + \mathbf{B}\|_{a,b} = \max_{\mathbf{x}} \{ \|(\mathbf{A} + \mathbf{B})\mathbf{x}\|_b : \|\mathbf{x}\|_a \leq 1 \} \quad (19)$$

$$= \max_{\mathbf{x}} \{ \|\mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{x}\|_b : \|\mathbf{x}\|_a \leq 1 \} \quad (20)$$

$$\leq \max_{\mathbf{x}} \{ \|\mathbf{A}\mathbf{x}\|_b + \|\mathbf{B}\mathbf{x}\|_b : \|\mathbf{x}\|_a \leq 1 \} \quad (21)$$

$$\leq \max_{\mathbf{x}} \{ \|\mathbf{A}\mathbf{x}\|_b : \|\mathbf{x}\|_a \leq 1 \} + \max_{\mathbf{x}} \{ \|\mathbf{B}\mathbf{x}\|_b : \|\mathbf{x}\|_a \leq 1 \} \quad (22)$$

$$= \|\mathbf{A}\|_{a,b} + \|\mathbf{B}\|_{a,b} \quad (23)$$

78 where the first inequality follows from the triangle inequality. This completes the proof.  $\square$

### Exercise 1.6

Let  $\|\cdot\|$  be a norm on  $\mathbb{R}^n$ . Show that the norm function  $f(\mathbf{x}) = \|\mathbf{x}\|$  is a continuous function over  $\mathbb{R}^n$ .

79

*Proof.* As we know, the continuity of  $f(\mathbf{x})$  at a point  $\mathbf{x}_0$  requires that, for any  $\epsilon > 0$  and the point  $\mathbf{x}_0$  in the domain  $\mathcal{D}$  of  $f$ , there always exists a  $\delta$  such that  $|f(\mathbf{x}) - f(\mathbf{x}_0)| < \epsilon$  whenever  $\mathbf{x} \in \mathcal{D}$  and  $\|\mathbf{x} - \mathbf{x}_0\| < \delta$ . Here, any nonnegative  $\delta < \epsilon$  will satisfy this requirement. To see this, we need to analyze two cases. For the case when  $\|\mathbf{x}\| > \|\mathbf{x}_0\|$ ,

$$|f(\mathbf{x}) - f(\mathbf{x}_0)| = \|\mathbf{x}\| - \|\mathbf{x}_0\| \quad (24)$$

$$= \|\mathbf{x} - \mathbf{x}_0 + \mathbf{x}_0\| - \|\mathbf{x}_0\| \quad (25)$$

$$\leq \|\mathbf{x} - \mathbf{x}_0\| + \|\mathbf{x}_0\| - \|\mathbf{x}_0\| \quad (26)$$

$$= \|\mathbf{x} - \mathbf{x}_0\| < \delta < \epsilon. \quad (27)$$

The case of  $\|\mathbf{x}\| = \|\mathbf{x}_0\|$  is trivial. For the case when  $\|\mathbf{x}\| < \|\mathbf{x}_0\|$ ,

$$|f(\mathbf{x}) - f(\mathbf{x}_0)| = \|\mathbf{x}_0\| - \|\mathbf{x}\| \quad (28)$$

$$= \|\mathbf{x}_0 - \mathbf{x} + \mathbf{x}\| - \|\mathbf{x}\| \quad (29)$$

$$\leq \|\mathbf{x} - \mathbf{x}_0\| + \|\mathbf{x}\| - \|\mathbf{x}\| \quad (30)$$

$$= \|\mathbf{x} - \mathbf{x}_0\| < \delta < \epsilon. \quad (31)$$

80 Since the above argument holds for any  $\mathbf{x}_0 \in \mathbb{R}^n$ , it follows that  $f(\mathbf{x}) = \|\mathbf{x}\|$  is continuous over  $\mathbb{R}^n$ .

81 This completes the proof.  $\square$

### Exercise 1.7

(attainment of the maximum in the induced norm definition) Suppose that  $\mathbb{R}^m$  and  $\mathbb{R}^n$  are equipped with norms  $\|\cdot\|_b$  and  $\|\cdot\|_a$ , respectively, and let  $\mathbf{A} \in \mathbb{R}^{m \times n}$ . Show that there exists  $\mathbf{x} \in \mathbb{R}^n$  such that  $\|\mathbf{x}\|_a \leq 1$  and  $\|\mathbf{Ax}\|_b = \|\mathbf{A}\|_{a,b}$ .

82

83 *Proof.* Define the set  $C = \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\|_a \leq 1\}$ . It is easy to see that  $C$  contains all the limits of  
 84 convergent sequences of points in  $C$ , so  $C$  is closed. We can find a positive number  $M$ , say 2, such that  
 85  $C \subset B(\mathbf{0}, M)$ , so  $C$  is bounded. Since  $\mathbf{0} \in C$ ,  $C$  is nonempty. Thus,  $C$  is a nonempty and compact  
 86 set. From Exercise 1.6, since  $\|\cdot\|_b$  is a norm,  $\|\mathbf{Ax}\|_b$  is continuous. According to Weierstrass theorem  
 87 (see Theorem 2.30 in the textbook), there exists a global minimum of  $f$  and a global maximum of  $f$   
 88 over  $C$ . By the definition of the induced norm, the maximum is denoted  $\|\mathbf{A}\|_{a,b}$ . This completes our  
 89 proof.  $\square$

### Exercise 1.8

Suppose that  $\mathbb{R}^m$  and  $\mathbb{R}^n$  are equipped with norms  $\|\cdot\|_b$  and  $\|\cdot\|_a$ , respectively. Show that the induced matrix norm  $\|\cdot\|_{a,b}$  can be computed by the formula

$$\|\mathbf{A}\|_{a,b} = \max_{\mathbf{x}} \{\|\mathbf{Ax}\|_b : \|\mathbf{x}\|_a = 1\}. \quad (32)$$

90

91 *Proof.* By the definition of the induced norm, the claim is equivalent to proving that the maxima are  
 92 achieved at  $\mathbf{x}^*$  satisfying  $\|\mathbf{x}^*\|_a = 1$ , which has been shown in Lemma 1.1.  $\square$

### Exercise 1.9

Suppose that  $\mathbb{R}^m$  and  $\mathbb{R}^n$  are equipped with norms  $\|\cdot\|_b$  and  $\|\cdot\|_a$ , respectively. Show that the induced matrix norm  $\|\cdot\|_{a,b}$  can be computed by the formula

$$\|\mathbf{A}\|_{a,b} = \max_{\mathbf{x} \neq \mathbf{0}} \frac{\|\mathbf{Ax}\|_b}{\|\mathbf{x}\|_a}. \quad (33)$$

93

94 *Proof.* This is exactly Lemma 2 which includes a proof.  $\square$

### Exercise 1.10

Let  $\mathbf{A} \in \mathbb{R}^{m \times n}$ ,  $\mathbf{B} \in \mathbb{R}^{n \times k}$  and assume that  $\mathbb{R}^m, \mathbb{R}^n, \mathbb{R}^k$  are equipped with the norms  $\|\cdot\|_c$ ,  $\|\cdot\|_b$ , and  $\|\cdot\|_a$ , respectively. Prove that

$$\|\mathbf{AB}\|_{a,c} \leq \|\mathbf{A}\|_{b,c} \|\mathbf{B}\|_{a,b}. \quad (34)$$

95

*Proof.* From Exercise 1.9, we have

$$\|\mathbf{AB}\|_{a,c} \leq \frac{\|\mathbf{ABx}\|_c}{\|\mathbf{x}\|_a} \quad (35)$$

where  $\mathbf{x} \neq \mathbf{0}$ . For every  $\mathbf{x} \neq \mathbf{0}$ , if  $\mathbf{Bx} = \mathbf{0}$ , then  $\mathbf{B} = \mathbf{0}$  must hold, in which case the claim is obviously true. When  $\mathbf{Bx} \neq \mathbf{0}$ , let  $\mathbf{y} = \mathbf{Bx}$  and then,

$$\|\mathbf{AB}\|_{a,c} \leq \frac{\|\mathbf{Ay}\|_c}{\|\mathbf{y}\|_b} \frac{\|\mathbf{Bx}\|_b}{\|\mathbf{x}\|_a} \leq \|\mathbf{A}\|_{b,c} \|\mathbf{B}\|_{a,b}. \quad (36)$$

96

This completes the proof.  $\square$



### Exercise 1.11

Prove the formula of the  $\infty$ -matrix norm given in Example 1.9 of the textbook. Specifically, given  $\mathbf{A} \in \mathbb{R}^{m \times n}$ ,

$$\|\mathbf{A}\|_\infty = \max_{i=1,2,\dots,m} \sum_{j=1}^n |A_{i,j}|. \quad (37)$$

97

*Proof.* From Exercise 1.8, the induced norm  $\|\mathbf{A}\|_\infty$  can also be computed by

$$\|\mathbf{A}\|_\infty = \max_{\mathbf{x}} \{\|\mathbf{Ax}\|_\infty : \|\mathbf{x}\|_\infty = 1\} \quad (38)$$

$$= \max_{\mathbf{x}} \left\{ \max_{i=1,\dots,m} \left| \sum_{j=1}^n A_{i,j} x_j \right| : \max_{j=1,\dots,n} |x_j| = 1 \right\} \quad (39)$$

$$= \max_{\mathbf{x}} \left\{ \max_{i=1,\dots,m} \sum_{j=1}^n |A_{i,j} x_j| : \max_{j=1,\dots,n} |x_j| = 1 \right\} \quad (40)$$

$$= \max_{i=1,\dots,m} \sum_{j=1}^n |A_{i,j} \text{sign}(A_{i,j})| = \max_{i=1,\dots,m} \sum_{j=1}^n |A_{i,j}| \quad (41)$$

98 where  $\text{sign}(A_{i,j}) = 1$  if  $A_{i,j} \geq 0$  otherwise  $\text{sign}(A_{i,j}) = -1$ . Note that, besides the last line, (40) also  
99 makes use of the constraint  $|x_j| \leq 1$  for every  $j \in \{1, \dots, n\}$ .  $\square$

### Exercise 1.12

Let  $\mathbf{A} \in \mathbb{R}^{m \times n}$ . Prove that

$$(i) \quad \frac{1}{\sqrt{n}} \|\mathbf{A}\|_\infty \leq \|\mathbf{A}\|_2 \leq \sqrt{m} \|\mathbf{A}\|_\infty,$$

$$(ii) \quad \frac{1}{\sqrt{m}} \|\mathbf{A}\|_1 \leq \|\mathbf{A}\|_2 \leq \sqrt{n} \|\mathbf{A}\|_1.$$

100

*Proof.* Before we prove the claimed 4 inequalities, we have

$$\|\mathbf{A}\|_2 = \max_{\|\mathbf{x}\|_2=1} \|\mathbf{Ax}\|_2 \quad (\text{Definition of } \|\mathbf{A}\|_2) \quad (42)$$

$$= \max_{\|\mathbf{x}\|_2=1} \sqrt{\sum_{i=1}^m \left( \sum_{j=1}^n A_{i,j} x_j \right)^2} \quad (\text{Definition of } \|\mathbf{A}\|_2) \quad (43)$$

$$= \max_{\|\mathbf{x}\|_2=1} \sqrt{\sum_{i=1}^m \left( \sum_{j=1}^n |A_{i,j}| |x_j| \right)^2} \quad (\forall j, \text{sgn}(x_j) \text{ does not change } \|\mathbf{x}\|_2) \quad (44)$$

Given this, for Part (i), we first show the second inequality.

$$\max_{\|\mathbf{x}\|_2=1} \sqrt{\sum_{i=1}^m \left( \sum_{j=1}^n |A_{i,j}| |x_j| \right)^2} \leq \max_{\|\mathbf{x}\|_\infty=1} \sqrt{\sum_{i=1}^m \left( \sum_{j=1}^n |A_{i,j}| |x_j| \right)^2} \quad (\{\mathbf{x} \mid \|\mathbf{x}\|_2=1\} \subset \{\mathbf{x} \mid \|\mathbf{x}\|_\infty=1\}) \quad (45)$$

$$= \sqrt{\sum_{i=1}^m \left( \sum_{j=1}^n |A_{ij}| \right)^2} \quad (\text{Maximum is attained at } |x_i| = 1 \ \forall i) \quad (46)$$

$$\leq \sqrt{\sum_{i=1}^m \left( \max_{j=1, \dots, n} \sum_{i=1}^m |A_{ij}| \right)^2} \quad (u_i \leq \max_i |u_i|, \ \forall i) \quad (47)$$

$$= \sqrt{\sum_{i=1}^m (\|\mathbf{A}\|_\infty)^2} = \sqrt{m} \|\mathbf{A}\|_\infty \quad (\text{Definition of } \|\mathbf{A}\|_\infty) \quad (48)$$

as desired. Now we prove the first inequality of Part (i).

$$\max_{\|\mathbf{x}\|_2=1} \sqrt{\sum_{i=1}^m \left( \sum_{j=1}^n |A_{ij}| |x_j| \right)^2} \geq \sqrt{\sum_{i=1}^m \left( \sum_{j=1}^n |A_{ij}| \cdot \frac{1}{\sqrt{n}} \right)^2} \quad \left( \sum_{j=1}^n \left( \frac{1}{\sqrt{n}} \right)^2 = 1 \right) \quad (49)$$

$$= \sqrt{\frac{1}{n} \sum_{i=1}^m \left( \sum_{j=1}^n |A_{ij}| \right)^2} \quad \left( \left( \frac{1}{\sqrt{n}} \right)^2 = \frac{1}{n} \right) \quad (50)$$

$$\geq \sqrt{\max_{i=1, \dots, m} \frac{1}{n} \left( \sum_{j=1}^n |A_{ij}| \right)^2} \quad \left( \sum_i |u_i| \geq \max_i |u_i| \ \forall i \right) \quad (51)$$

$$= \frac{1}{\sqrt{n}} \max_{i=1, \dots, m} \sum_{j=1}^n |A_{ij}| = \frac{1}{\sqrt{n}} \|\mathbf{A}\|_\infty \quad (\text{Definition of } \|\mathbf{A}\|_\infty) \quad (52)$$

For part (ii), we first consider the left inequality.

$$\max_{\|\mathbf{x}\|_2=1} \sqrt{\sum_{i=1}^m \left( \sum_{j=1}^n |A_{ij}| |x_j| \right)^2} = \sqrt{m} \cdot \max_{\|\mathbf{x}\|_2=1} \frac{\sum_{i=1}^m \sum_{j=1}^n |A_{ij}| |x_j|}{m} \quad (\text{AM-QM inequality}) \quad (53)$$

$$= \frac{1}{\sqrt{m}} \cdot \max_{\|\mathbf{x}\|_2=1} \sum_{j=1}^n |x_j| \left( \sum_{i=1}^m |A_{ij}| \right) \quad \left( \forall m, n < \infty, \sum_{i=1}^m \sum_{j=1}^n = \sum_{j=1}^n \sum_{i=1}^m \right) \quad (54)$$

$$= \frac{1}{\sqrt{m}} \cdot \max_{\|\mathbf{x}\|_2=1} \sqrt{\sum_{j=1}^n |x_j|^2} \sqrt{\sum_{j=1}^n \left( \sum_{i=1}^m |A_{ij}| \right)^2} \quad (\text{Cauchy-Schwarz inequality}) \quad (55)$$

$$= \frac{1}{\sqrt{m}} \sqrt{\sum_{j=1}^n \left( \sum_{i=1}^m |A_{ij}| \right)^2} \quad (\|\mathbf{A}\|_2 = 1) \quad (56)$$

$$\geq \frac{1}{\sqrt{m}} \sqrt{\max_{j=1, \dots, n} \left( \sum_{i=1}^m |A_{ij}| \right)^2} \quad \left( \sum_i |u_i| \geq \max_i |u_i| \ \forall i \right) \quad (57)$$

$$= \frac{1}{\sqrt{m}} \max_{j=1, \dots, n} \sum_{i=1}^m |A_{ij}| = \frac{1}{\sqrt{m}} \|\mathbf{A}\|_1 \quad (\text{Definition of } \|\mathbf{A}\|_1) \quad (58)$$

101 When applying the AM-GM inequality, the equality holds if and only if  $\sum_{j=1}^n |A_{1j}x_j| = \dots =$   
 102  $\sum_{j=1}^n |A_{mj}x_j|$ , which is attainable. For Cauchy-Schwarz inequality, the equality holds if and only if  
 103  $\sum_{i=1}^m |A_{ij}| = k|x_j|$  for all  $j = 1, \dots, n$  where  $k$  is a constant, which is attainable too.

Now we show the inequality on the right hand side.

$$\max_{\|\mathbf{x}\|_2=1} \sqrt{\sum_{i=1}^m \left( \sum_{j=1}^n |A_{ij}| |x_j| \right)^2} \leq \max_{\|\mathbf{x}\|_2=1} \sqrt{\sum_{i=1}^m \sum_{j=1}^n A_{ij}^2 \sum_{j=1}^n x_j^2} \quad (\text{Cauchy-Schwarz inequality}) \quad (59)$$

$$= \sqrt{\sum_{i=1}^m \sum_{j=1}^n A_{ij}^2} \quad (\|\mathbf{x}\|_2 = 1) \quad (60)$$

$$= \sqrt{\sum_{j=1}^n \sum_{i=1}^m |A_{ij}|^2} \quad \left( \forall m, n < \infty, \sum_{i=1}^m \sum_{j=1}^n = \sum_{j=1}^n \sum_{i=1}^m \right) \quad (61)$$

$$\leq \sqrt{\sum_{j=1}^n \left( \sum_{i=1}^m |A_{ij}| \right)^2} \quad \left( \forall a_i \geq 0, \sum_{i=1}^m a_i^2 \leq \left( \sum_{i=1}^m a_i \right)^2 \right) \quad (62)$$

$$\leq \sqrt{\sum_{j=1}^n \left( \max_{i=1, \dots, m} \sum_{i=1}^m |A_{ij}| \right)^2} \quad (u_i \leq \max_i |u_i|, \forall i) \quad (63)$$

$$= \sqrt{n} \cdot \max_{j=1, \dots, n} \sum_{i=1}^m |A_{ij}| \quad \left( \sum_{j=1}^n c = nc \right) \quad (64)$$

$$= \sqrt{n} \|\mathbf{A}\|_1 \quad (\text{Definition of } \|\mathbf{A}\|_1) \quad (65)$$

104 where in the first line the equality holds if and only if  $|A_{ij}| = k_i |x_j|$  for all  $i = 1, \dots, m$  and  
 105  $j = 1, \dots, n$ , and  $k_i$  is a constant, which is not necessarily attainable. This completes the proof.  $\square$

### Exercise 1.13

Let  $\mathbf{A} \in \mathbb{R}^{m \times n}$ . Show that

(i)  $\|\mathbf{A}\| = \|\mathbf{A}^T\|$  (here  $\|\cdot\|$  is the spectral norm),

(ii)  $\|\mathbf{A}\|_F^2 = \sum_{i=1}^n \lambda_i(\mathbf{A}^T \mathbf{A})$ .

*Proof.* For part (i), the spectral norm is defined by

$$\|\mathbf{A}\|_2 = \sqrt{\lambda_{\max}(\mathbf{A}^T \mathbf{A})} = \sigma_{\max}(\mathbf{A}) \quad (66)$$

where  $\lambda_{\max}(\mathbf{A}^T \mathbf{A})$  is the maximum eigenvalue of  $\mathbf{A}^T \mathbf{A}$ , and  $\sigma_{\max}(\mathbf{A})$  is the largest singular values of  $\mathbf{A}$ . Similarly,

$$\|\mathbf{A}^T\|_2 = \sqrt{\lambda_{\max}(\mathbf{A} \mathbf{A}^T)} = \sigma_{\max}(\mathbf{A}^T) \quad (67)$$

By the Theorem 2.6.3(a) in [Horn and Johnson \(2013\)](#), the singular values are supposed to be nonnegative. And by the Theorem 2.6.3(b) in [Horn and Johnson \(2013\)](#), the nonzero eigenvalues of  $\mathbf{A}\mathbf{A}^T$  and  $\mathbf{A}^T\mathbf{A}$  are identical. Thus,

$$\|\mathbf{A}\|_2 = \sqrt{\lambda_{\max}(\mathbf{A}^T\mathbf{A})} = \sqrt{\lambda_{\max}(\mathbf{A}\mathbf{A}^T)} = \|\mathbf{A}^T\|_2 \quad (68)$$

107 as desired.

Now we consider part (ii).

$$\|\mathbf{A}\|_F^2 = \sum_{i=1}^m \sum_{j=1}^n A_{ij}^2 \quad (\text{Definition of Frobenius norm}) \quad (69)$$

$$= \text{Tr}(\mathbf{A}^T\mathbf{A}) \quad (\text{Definition of trace}) \quad (70)$$

$$= \sum_{i=1}^n \lambda_i(\mathbf{A}^T\mathbf{A}) \quad (71)$$

where the last line follows from the following argument<sup>2</sup>. By definition, the characteristic polynomial of  $\mathbf{A}^T\mathbf{A}$  is given by

$$p(t) = \det(t\mathbf{I} - \mathbf{A}^T\mathbf{A}) \quad (72)$$

$$= t^n - \text{Tr}(\mathbf{A}^T\mathbf{A})t^{n-1} + \cdots + (-1)^n \det(\mathbf{A}^T\mathbf{A}) \quad (\text{Definition of determinant}) \quad (73)$$

Also, by the definition, eigenvalues are the roots of  $p(t)$ . Hence,

$$p(t) = (t - \lambda_1)(t - \lambda_2) \cdots (t - \lambda_n) \quad (74)$$

By comparing coefficients, we have

$$\text{Tr}(\mathbf{A}^T\mathbf{A}) = \sum_{i=1}^n \lambda_i(\mathbf{A}^T\mathbf{A}) \quad (75)$$

108 which completes the proof. □

#### Exercise 1.14

Let  $\mathbf{A} \in \mathbb{R}^{n \times n}$  be a symmetric matrix. Show that

$$\max_{\mathbf{x}} \{\mathbf{x}^T \mathbf{A} \mathbf{x} : \|\mathbf{x}\|^2 = 1\} = \lambda_{\max}(\mathbf{A}). \quad (76)$$

109

110 The inspiration of the following proof is from the proof of Lemma 1.11 in the textbook.

*Proof.* According to the spectral decomposition theorem there exists an orthogonal matrix  $\mathbf{U} \in \mathbb{R}^{n \times n}$  and a diagonal matrix  $\mathbf{D} = \text{diag}(d_1, d_2, \dots, d_n)$  such that  $\mathbf{U}^T \mathbf{A} \mathbf{U} = \mathbf{D}$ . Without the loss of generality, we can assume that the diagonal elements of  $\mathbf{D}$ , which are the eigenvalues of  $\mathbf{A}$ , are ordered nonincreasingly:  $d_1 \geq d_2 \geq \cdots \geq d_n$ , where  $d_1 = \lambda_{\max}(\mathbf{A})$ . Since  $\mathbf{U}$  is an orthogonal matrix, we can make the change of variables  $\mathbf{x} = \mathbf{U}\mathbf{y}$ .

$$\max_{\|\mathbf{x}\|_2^2=1} \mathbf{x}^T \mathbf{A} \mathbf{x} = \max_{\|\mathbf{U}\mathbf{y}\|_2^2=1} (\mathbf{U}\mathbf{y})^T \mathbf{A} \mathbf{U} \mathbf{y} \quad (77)$$

$$= \max_{\|\mathbf{y}\|_2^2=1} \mathbf{y}^T \mathbf{U}^T \mathbf{A} \mathbf{U} \mathbf{y} \quad (\|\mathbf{U}\mathbf{y}\|_2^2 = \|\mathbf{y}\|_2^2) \quad (78)$$

<sup>2</sup><https://math.stackexchange.com/questions/546155/proof-that-the-trace-of-a-matrix-is-the-sum-of-its-eigenvalues>

$$= \max_{\|\mathbf{y}\|_2=1} \mathbf{y}^T \mathbf{D} \mathbf{y} \quad (\mathbf{U}^T \mathbf{A} \mathbf{U} = \mathbf{D}) \quad (79)$$

$$= \max_{\|\mathbf{y}\|_2=1} \sum_{i=1}^n d_i y_i^2 \leq d_1 \max_{\|\mathbf{y}\|_2=1} \sum_{i=1}^n y_i^2 \quad (d_1 \geq d_2 \geq \dots \geq d_n) \quad (80)$$

$$= d_1 = \lambda_{\max}(\mathbf{A}) \quad (81)$$

111

□

### Exercise 1.15

Prove that a set  $U \subseteq \mathbb{R}^n$  is closed if and only if its complement  $U^c$  is open.

112

113 *Proof.* We first prove the sufficiency. Given  $U^c$  is open, we suppose that  $U$  is not closed. Then there  
 114 must exist at least one accumulation point of  $U$ , say  $x$ , such that  $x \notin U$ , i.e.,  $x \in U^c$ . Since  $U^c$  is  
 115 open, then there exists an open ball  $B(x, r) \subseteq U^c$  with  $r > 0$ , which contradicts  $x \in U'$  where  $U'$   
 116 denotes the set of accumulation points of  $U$ . Specifically, since  $x \in U'$ , by Definition 1.4, there are  
 117 infinitely many points of  $B(x, r)$  belonging to  $U$ , which is impossible for  $B(x, r) \subseteq U^c$ .

118 Now we show the necessity. Given any point  $x \in U^c$ , it suffices to show that  $x$  is an interior point  
 119 of  $U^c$ . Obviously,  $x \notin U$ . Since  $U$  is closed,  $x$  is not an accumulation point of  $U$ . By Definition 1.5,  
 120 this implies that there exists an open ball  $B(x, r)$  such that  $B(x, r) \cap U = \emptyset$ . Thus,  $B(x, r) \subseteq U^c$ .  
 121 This completes our proof. □

### Exercise 1.16

1. Let  $\{A_i\}_{i \in I}$  be a collection of open sets where  $I$  is a given index set. Show that  $\bigcup_{i \in I} A_i$  is an open Set. Show that if  $I$  is finite, then  $\bigcap_{i \in I} A_i$  is open.
2. Let  $\{A_i\}_{i \in I}$  be a collection of closed sets where  $I$  is a given index set. Show that  $\bigcap_{i \in I} A_i$  is a closed Set. Show that if  $I$  is finite, then  $\bigcup_{i \in I} A_i$  is closed.

122

123 The following proof is taken from the proof of Theorem 11.1.5 in [Chen et al. \(2019\)](#).

124 *Proof.*

- 125 1. For any  $\mathbf{x} \in \bigcup_{i \in I} A_i$ , then there exists at least an  $i \in I$  such that  $\mathbf{x} \in A_i$ . Since  $A_i$  is an open set,  
 126 then  $\mathbf{x}$  is an interior point of  $A_i$ . Also,  $\mathbf{x}$  is an interior point of  $\bigcup_{i \in I} A_i$ . Thus,  $\bigcup_{i \in I} A_i$  is an open  
 127 set.

128 Since  $I$  is finite, suppose there are  $k$  sets in total. For any  $\mathbf{x} \in \bigcap_{i \in I} A_i$ ,  $x \in A_i$  for arbitrary  
 129  $i = 1, \dots, k$ . Thus, for any  $i \in I$ , there exists an  $r_i > 0$  such that  $B(\mathbf{x}, r_i) \subset A_i$ . Let  $r = \min_{i \in I} r_i$ ,  
 130 then  $B(\mathbf{x}, r) \subset \bigcap_{i \in I} A_i$ . Therefore,  $\bigcap_{i \in I} A_i$  is open.

- 131 2. By De Morgan's Theorem (see Theorem 1.10),  $(\bigcap_{i \in I} A_i)^c = \bigcup_{i \in I} A_i^c$ . Since  $A_i$  is closed, its  
 132 complement  $A_i^c$  is open. From the first part of this proof,  $\bigcup_{i \in I} A_i^c$  is open. Thus, its complement  
 133  $\bigcap_{i \in I} A_i$  is closed.

134 If each  $A_i$  is closed, then  $A_i^c$  is open. If  $I$  is finite, by the first part of this proof,  $\bigcap_{i \in I} A_i^c$  is open.  
 135 According to De Morgan's Theorem, its complement is  $\bigcup_{i \in I} A_i$  which is closed. This completes  
 136 the proof.

137

□

### Exercise 1.17

Give an example of open sets  $A_i$ ,  $i \in I$  for which  $\bigcap_{i \in I} A_i$  is not open.

The following solution is from Mathematics Stack Exchange<sup>3</sup>.

**Solution:** Let  $\mathbb{Z}_+$  denote the set of positive integers. When  $A_i$  is defined as

$$A_i = \left(-\frac{1}{i}, \frac{1}{i}\right), \quad i \in \mathbb{Z}_+,$$

the intersection

$$\bigcap_{i \in \mathbb{Z}_+} A_i = [0]$$

is not open. However, it is a closed set. □

### Extensions

Likewise, we can construct an example of closed sets  $A_i$ ,  $i \in \mathbb{Z}_+$  for which  $\bigcup_{i \in \mathbb{Z}_+} A_i$  is not closed. For example, the union of the closed sets  $A_i = [\frac{1}{i}, 2 - \frac{1}{i}]$ ,  $\forall i \in \mathbb{Z}_+$  is  $(0, 2)$  which is an open set.

### Exercise 1.18

Let  $A, B \subseteq \mathbb{R}^n$ . Prove that  $\text{cl}(A \cap B) \subseteq \text{cl}(A) \cap \text{cl}(B)$ . Give an example in which the inclusion is proper.

This proof is from Mathematics Stack Exchange<sup>4</sup>.

*Proof.* By the definition of closure, i.e. Definition 1.7,  $\text{cl}(U) = U \cup \text{bd}(U)$ . since  $A \cap B \subseteq A$ , it follows that  $\text{cl}(A \cap B) \subseteq \text{cl}(A)$ . Likewise,  $\text{cl}(A \cap B) \subseteq \text{cl}(B)$ . Thus,  $\text{cl}(A \cap B) \subseteq \text{cl}(A) \cap \text{cl}(B)$  as desired.

Given  $A = (0, 1)$  and  $B = (1, 2)$ , then  $A \cap B = \emptyset$  and  $\text{cl}(A \cap B) = \emptyset$ . On the other hand,  $\text{cl}(A) = [0, 1]$  and  $\text{cl}(B) = [1, 2]$ . Thus,  $\text{cl}(A) \cap \text{cl}(B) = \{1\}$ . Obviously,  $\emptyset \neq \{1\}$ . Hence, the inclusion is proper in this case. □

### Exercise 1.19

Let  $A, B \subseteq \mathbb{R}^n$ . Prove that  $\text{int}(A \cap B) = \text{int}(A) \cap \text{int}(B)$  and that  $\text{int}(A) \cup \text{int}(B) \subseteq \text{int}(A \cup B)$ . Show an example in which the latter inclusion is proper.

*Proof.* The first part of the following proof is from a YouTube video<sup>5</sup>.

1.  $\text{int}(A \cap B) \subseteq \text{int}(A) \cap \text{int}(B)$  follows from

$$A \cap B \subseteq A \Rightarrow \text{int}(A \cap B) \subseteq \text{int}(A) \tag{82}$$

$$A \cap B \subseteq B \Rightarrow \text{int}(A \cap B) \subseteq \text{int}(B) \tag{83}$$

$\Downarrow$

$$\text{int}(A \cap B) \subseteq \text{int}(A) \cap \text{int}(B). \tag{84}$$

<sup>3</sup><https://math.stackexchange.com/questions/1460853/infinite-intersection-of-open-sets>

<sup>4</sup><https://math.stackexchange.com/questions/1485869/closure-of-intersection-of-two-sets>

<sup>5</sup><https://www.youtube.com/watch?v=uZZkMloQbd0>

$\text{int}(A) \cap \text{int}(B) \subseteq \text{int}(A \cap B)$  follows from

$$\text{int}(A) \subseteq A, \quad \text{int}(B) \subseteq B \quad (85)$$

$\Downarrow$

$$\text{int}(A) \cap \text{int}(B) \subseteq A \cap B. \quad (86)$$

151 Since the finite intersection of open sets is an open set (see Exercise 1.16(i)), then  $\text{int}(A) \cap \text{int}(B)$   
 152 is open. By definition, the interior of a set is the largest open subset of that set, so  $\text{int}(A \cap B)$   
 153 contains  $\text{int}(A) \cap \text{int}(B)$ . In other words,  $\text{int}(A) \cap \text{int}(B) \subseteq \text{int}(A \cap B)$ . Therefore,  $\text{int}(A \cap B) =$   
 154  $\text{int}(A) \cap \text{int}(B)$ .

2.  $\text{int}(A) \cup \text{int}(B) \subseteq \text{int}(A \cup B)$  follows from

$$\text{int}(A) \subseteq A, \quad \text{int}(B) \subseteq B \quad (87)$$

$\Downarrow$

$$\text{int}(A) \cup \text{int}(B) \subseteq A \cup B. \quad (88)$$

155 In Exercise 1.16(i), we have shown that the union of open sets is open, so  $\text{int}(A) \cup \text{int}(B)$  is an  
 156 open set. By definition, the interior of  $A \cup B$  is the largest open set of  $A \cup B$ . Thus,  $\text{int}(A \cup B)$   
 157 contains  $\text{int}(A) \cup \text{int}(B)$ . Hence,  $\text{int}(A) \cup \text{int}(B) \subseteq \text{int}(A \cup B)$ .

158 For example,  $A = (0, 1)$  and  $B = [1, 2)$ . It is easy to see that  $\text{int}(A) \cup \text{int}(B) = (0, 1) \cup (1, 2)$ ,  
 159 but  $\text{int}(A \cup B) = (1, 2)$ . This inclusion is proper.

160 □

## 161 2 Chapter 2 Optimality Conditions for Unconstrained Opti- 162 mization

### Exercise 2.1

Find the global minimum and maximum points of the function  $f(x, y) = x^2 + y^2 + 2x - 3y$  over the unit ball  $S = B[0, 1] = \{(x, y) : x^2 + y^2 \leq 1\}$ .

163

**Solution:** By applying Cauchy-Swcharz inequality on  $2x - 3y$ , we get

$$|2x - 3y| = \left| \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 2 \\ -3 \end{pmatrix} \right| \leq \sqrt{2^2 + (-3)^2} \sqrt{x^2 + y^2} = \sqrt{13} \sqrt{x^2 + y^2}$$

$\Downarrow$

$$-\sqrt{13} \sqrt{x^2 + y^2} \leq 2x - 3y \leq \sqrt{13} \sqrt{x^2 + y^2}$$

where the equalities hold when  $-3x = 2y$ . Thus,

$$x^2 + y^2 - \sqrt{13} \sqrt{x^2 + y^2} \leq x^2 + y^2 + 2x - 3y \leq x^2 + y^2 + \sqrt{13} \sqrt{x^2 + y^2}$$

Let  $t = \sqrt{x^2 + y^2}$ , then the right hand side can be written as

$$f_{\text{RHS}}(t) = t^2 + \sqrt{13}t, \quad \text{with } 0 \leq t \leq 1. \quad (89)$$

164 Since  $f'_{\text{RHS}}(t) = 2t + \sqrt{13} \geq 0$ , then  $f_{\text{RHS}}(t)$  is increasing on  $[0, 1]$ . So, the maximum can be  
 165 attained at  $t = 1$ . Thus, solving  $x^2 + y^2 = 1$  and  $-3x = 2y$  gives  $x = 2/\sqrt{13}$  and  $y = -3/\sqrt{13}$  and  
 166  $f(2/\sqrt{13}, -3/\sqrt{13}) = 1 + \sqrt{13}$ , which is equal to  $f_{\text{RHS}}(1) = 1 + \sqrt{13}$ .

The left hand side is

$$f_{\text{LHS}}(t) = t^2 - \sqrt{13}t, \text{ with } 0 \leq t \leq 1 \quad (90)$$

167 Its derivative with respect to  $t$  is  $f'_{\text{LHS}}(t) = 2t - \sqrt{13} < 0$  on  $[0, 1]$ , which means  $f_{\text{LHS}}(t)$  is  
 168 strictly decreasing on  $[0, 1]$ . The minimum can be achieved at  $t = 1$ , i.e.  $x^2 + y^2 = 1$  and  
 169  $f_{\text{LHS}}(1) = 1 - \sqrt{13}$ . Given  $-3x = 2y$ , we obtain  $x = -2/\sqrt{13}$  and  $y = 3/\sqrt{13}$ , which gives the desired  
 170  $f(-2/\sqrt{13}, 3/\sqrt{13}) = 1 - \sqrt{13}$ .

171 To sum up, the global minimum and maximum points are  $(x, y) = (2/\sqrt{13}, -3/\sqrt{13})$  and  
 172  $(x, y) = (-2/\sqrt{13}, 3/\sqrt{13})$ , respectively.  $\square$

### Exercise 2.2

Let  $\mathbf{a} \in \mathbb{R}^n$  be a nonzero vector. Show that the maximum of  $\mathbf{a}^T \mathbf{x}$  over  $B[0, 1] = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\| \leq 1\}$  is attained at  $\mathbf{x}^* = \frac{\mathbf{a}}{\|\mathbf{a}\|}$  and that the maximal value is  $\|\mathbf{a}\|$ .

173

*Proof.* According to Cauchy-Schwarz inequality, we have

$$\mathbf{a}^T \mathbf{x} \leq \|\mathbf{a}\| \|\mathbf{x}\| \quad (91)$$

174 the equality holds if and only if  $\mathbf{x} = \lambda \mathbf{a}$  where  $0 \neq \lambda \in \mathbb{R}$ . Since  $\|\mathbf{x}\| \leq 1$ , the maximum of the right  
 175 hand side can be achieved when  $\|\mathbf{x}\| = 1$ . Combining this with  $\mathbf{x} = \lambda \mathbf{a}$ , we get  $\|\lambda \mathbf{a}\| = 1$  and  $\lambda = \frac{1}{\|\mathbf{a}\|}$ .  
 176 Thus,  $\mathbf{x}^* = \lambda \mathbf{a} = \frac{\mathbf{a}}{\|\mathbf{a}\|}$  and the maximum value is  $\|\mathbf{a}\| \|\mathbf{x}\| = \|\mathbf{a}\|$ .  $\square$

### Exercise 2.3

Find the global minimum and maximum points of the function  $f(x, y) = 2x - 3y$  over the set  $S = \{(x, y) : 2x^2 + 5y^2 \leq 1\}$ .

177

178 **Solution:** We can make use of the result in Exercise 2.2. To do this, we need to perform a change of  
 179 variables. Specifically, let  $u = \sqrt{2}x$  and  $v = \sqrt{5}y$ . By doing this, the original problem is equivalently  
 180 reformulated as finding the global minimum and maximum points of  $\tilde{f}(u, v) = \sqrt{2}u - \frac{3\sqrt{5}}{5}v$  over the  
 181 set  $\tilde{S} = \{(u, v) : u^2 + v^2 \leq 1\}$ . In this case,  $\mathbf{a} = (\sqrt{2}, -\frac{3\sqrt{5}}{5})^T$ . It follows from that the maximum  
 182 point is  $\frac{\mathbf{a}}{\|\mathbf{a}\|} = (\frac{5\sqrt{2}}{19}, -\frac{3\sqrt{5}}{19})^T$ . Changing back to the original variables gives  $x = 5/19$  and  $-3/19$ .  
 183 Similarly, the minimum point is  $x = -5/19$  and  $3/19$ .  $\square$

### Exercise 2.4

Show that if  $\mathbf{A}, \mathbf{B}$  are  $n \times n$  positive semidefinite matrices, then their sum  $\mathbf{A} + \mathbf{B}$  is also positive semidefinite.

184

*Proof.* Since  $\mathbf{A}, \mathbf{B}$  are semidefinite matrices, then  $\mathbf{x}^T \mathbf{A} \mathbf{x} \geq 0$  and  $\mathbf{x}^T \mathbf{B} \mathbf{x} \geq 0$  for every  $\mathbf{x} \in \mathbb{R}^n$ . It follows that

$$\mathbf{x}^T (\mathbf{A} + \mathbf{B}) \mathbf{x} = \mathbf{x}^T \mathbf{A} \mathbf{x} + \mathbf{x}^T \mathbf{B} \mathbf{x} \geq 0 \quad (92)$$

185 for every  $\mathbf{x} \in \mathbb{R}^n$ . Hence,  $\mathbf{A} + \mathbf{B}$  is also positive semidefinite. This completes the proof.  $\square$



### Exercise 2.5

Let  $\mathbf{A} \in \mathbb{R}^{n \times n}$  and  $\mathbf{B} \in \mathbb{R}^{m \times m}$  be two symmetric matrices. Prove that the following two claims are equivalent:

- i  $\mathbf{A}$  and  $\mathbf{B}$  are positive semidefinite.
- ii  $\begin{pmatrix} \mathbf{A} & \mathbf{0}_{n \times m} \\ \mathbf{0}_{m \times n} & \mathbf{B} \end{pmatrix}$  is positive semidefinite.

186

## 3 Chapter 3 Least Squares

187

### Exercise 3.1

Let  $\mathbf{A} \in \mathbb{R}^{m \times n}$ ,  $\mathbf{b} \in \mathbb{R}^m$ ,  $\mathbf{L} \in \mathbb{R}^{p \times n}$ , and  $\lambda \in \mathbb{R}_{++}$ . Consider the regularized least squares problem

$$\min_{\mathbf{x} \in \mathbb{R}^n} \|\mathbf{Ax} - \mathbf{b}\|^2 + \lambda \|\mathbf{Lx}\|^2. \quad (\text{RLS})$$

Show that (RLS) has a unique solution if and only if  $\text{Null}(\mathbf{A}) \cap \text{Null}(\mathbf{L}) = \{\mathbf{0}\}$ , where here for a matrix  $\mathbf{B}$ ,  $\text{Null}(\mathbf{B})$  is the null space of  $\mathbf{B}$  given by  $\{\mathbf{x} : \mathbf{Bx} = \mathbf{0}\}$ .

188

189 Note that it is supposed to be  $\mathbf{b} \in \mathbb{R}^m$  instead of  $\mathbf{b} \in \mathbb{R}^n$ . In the textbook, this is a typo which is  
190 not yet mentioned at [http://www.siam.org/books/mo19/mo19\\_err.pdf](http://www.siam.org/books/mo19/mo19_err.pdf).

*Proof.* Since the Hessian of the objective function is  $2(\mathbf{A}^T \mathbf{A} + \lambda \mathbf{L}^T \mathbf{L}) \succeq \mathbf{0}$ , it follows by Lemma 2.41 of the textbook that any stationary point is a global minimum point. Then, we have

$$\begin{aligned} (\text{RLS}) \text{ has a unique solution} &\iff \mathbf{A}^T \mathbf{A} + \lambda \mathbf{L}^T \mathbf{L} \succ \mathbf{0} \\ &\iff \\ \mathbf{x}^T (\mathbf{A}^T \mathbf{A} + \lambda \mathbf{L}^T \mathbf{L}) \mathbf{x} > 0, \forall \mathbf{x} \neq \mathbf{0} &\iff \|\mathbf{Ax}\|^2 + \lambda \|\mathbf{Lx}\|^2 > 0, \forall \mathbf{x} \neq \mathbf{0} \\ &\iff \\ \text{There exists no nonzero } \mathbf{x} \text{ such that } \mathbf{Ax} = \mathbf{0} \text{ and } \mathbf{Lx} = \mathbf{0} &\text{ hold simultaneously.} \\ &\iff \\ \text{Null}(\mathbf{A}) \cap \text{Null}(\mathbf{L}) = \{\mathbf{0}\}. & \end{aligned}$$

191 This completes the proof. □

## 4 Chapter 4 The Gradient Method

192

193 Before working on the exercises of Chapter 4, we first introduce the notation of  $f \in C_L^{k,p}(D)$ . We  
194 write  $f \in C_L^{k,p}(D)$  if

- 195 1.  $f^{(k)}$  exists and is continuous on  $D$ .
- 2.  $f^{(p)}$  is Lipschitz continuous with a constant  $L$ , namely,

$$\|f^{(p)}(y_1) - f^{(p)}(y_2)\| \leq L \|y_1 - y_2\|, \quad \forall y_1, y_2 \in D.$$

### Exercise 4.1

Let  $f \in C_L^{1,1}(\mathbb{R}^n)$  and let  $\{\mathbf{x}^k\}_{k \geq 0}$  be the sequence generated by the gradient method with a constant stepsize  $t_k = \frac{1}{L}$ . Assume that  $\mathbf{x}_k \rightarrow \mathbf{x}^*$ . Show that if  $\nabla f(\mathbf{x}_k) \neq \mathbf{0}$  for all  $k \geq 0$ , then  $\mathbf{x}^*$  is not a local maximum point.

*Proof.* Suppose  $\mathbf{x}^*$  is a local maximum point, then there exists a ball  $B(\mathbf{x}^*, r)$  with any  $r > 0$  such that

$$f(\mathbf{x}^*) \geq f(\mathbf{x}_k), \quad \forall \mathbf{x}_k \in B(\mathbf{x}^*, r)$$

Since  $t_k = \frac{1}{L}$ , by the descent lemma (Lemma 4.22 in the textbook), we have

$$\begin{aligned} f(\mathbf{x}^*) &\leq f(\mathbf{x}_k) + \nabla f(\mathbf{x}_k)^T (\mathbf{x}^* - \mathbf{x}_k) + \frac{L}{2} \|\mathbf{x}^* - \mathbf{x}_k\|^2 \\ &= f(\mathbf{x}_k) + \nabla f(\mathbf{x}_k)^T \left(-\frac{1}{L} \nabla f(\mathbf{x}_k)\right) + \frac{L}{2} \left\|-\frac{1}{L} \nabla f(\mathbf{x}_k)\right\|^2 \\ &= f(\mathbf{x}_k) - \frac{1}{2L} \|\nabla f(\mathbf{x}_k)\|^2 \\ &< f(\mathbf{x}_k) \end{aligned}$$

where the last line follows from that  $\nabla f(\mathbf{x}_k) \neq \mathbf{0}$  for all  $k \geq 0$ . This contradicts the supposition, which implies that  $\mathbf{x}^*$  is not a local maximum point. This completes the proof.  $\square$

## 5 Chapter 5 Newton's Method

## 6 Chapter 6 Convex Sets

## 7 Chapter 7 Convex Functions

### Exercise 7.36

Prove that for any  $x_1, x_2, \dots, x_n \in \mathbb{R}_+$  the following inequality holds:

$$\frac{\sum_{i=1}^n x_i}{n} \leq \sqrt{\frac{\sum_{i=1}^n x_i^2}{n}}$$

*Proof.* According to Cauchy-Schwartz inequality which says that given two vectors  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ ,  $\|\mathbf{x}\|_2 \|\mathbf{y}\|_2 \geq |\mathbf{x}^T \mathbf{y}|$ , we have

$$\begin{aligned} \sqrt{\frac{\sum_{i=1}^n x_i^2}{n}} &= \sqrt{\sum_{i=1}^n \left(\frac{|x_i|}{\sqrt{n}}\right)^2} \cdot \sqrt{\sum_{i=1}^n \left(\frac{1}{\sqrt{n}}\right)^2} \\ &\geq \frac{\sum_{i=1}^n |x_i|}{n} \geq \frac{\sum_{i=1}^n x_i}{n}, \end{aligned}$$

where the equalities in the first and second inequalities hold if and only if  $|x_1| = |x_2| = \dots = |x_n|$  and  $x_1 = x_2 = \dots = x_n$ , respectively. This completes the proof.  $\square$

### Exercise 7.37

Prove that for any  $x_1, x_2, \dots, x_n \in \mathbb{R}_{++}$  the following inequality holds:

$$\frac{\sum_{i=1}^n x_i^2}{\sum_{i=1}^n x_i} \leq \sqrt{\frac{\sum_{i=1}^n x_i^3}{\sum_{i=1}^n x_i}}$$

*Proof.* Let  $f(x) = x^2$  and then  $f''(x) = 2 > 0$  implying that  $f$  is convex. Furthermore, given  $\lambda_1, \lambda_2, \dots, \lambda_n \in [0, 1]$  satisfying  $\sum_{i=1}^n \lambda_i = 1$ , we have

$$\left( \sum_{i=1}^n \lambda_i x_i \right)^2 \leq \sum_{i=1}^n \lambda_i x_i^2$$

By letting  $\lambda_i = \frac{x_i}{\sum_{i=1}^n x_i}$ , we have

$$\left( \sum_{i=1}^n \frac{x_i}{\sum_{i=1}^n x_i} x_i \right)^2 \leq \sum_{i=1}^n \frac{x_i}{\sum_{i=1}^n x_i} x_i^2 \iff \left( \frac{\sum_{i=1}^n x_i^2}{\sum_{i=1}^n x_i} \right)^2 \leq \frac{\sum_{i=1}^n x_i^3}{\sum_{i=1}^n x_i} \iff \frac{\sum_{i=1}^n x_i^2}{\sum_{i=1}^n x_i} \leq \sqrt{\frac{\sum_{i=1}^n x_i^3}{\sum_{i=1}^n x_i}}.$$

Note that the condition  $\lambda_i \in [0, 1]$  is satisfied automatically since  $x_i > 0, \forall i = 1, 2, \dots, n$ . This completes our proof.  $\square$

### Exercise 7.38

Let  $x_1, x_2, \dots, x_n > 0$  satisfy  $\sum_{i=1}^n x_i = 1$ . Prove that

$$\sum_{i=1}^n \frac{x_i}{\sqrt{1-x_i}} \geq \sqrt{\frac{n}{n-1}}.$$

*Proof.* Define  $f(x) = 1/\sqrt{1-x}$  and then  $f''(x) = \frac{3}{4}(1-x)^{-5/2} > 0$ . So  $f(x)$  is convex. Since  $\sum_{i=1}^n x_i = 1$ , then we have

$$\begin{aligned} \sum_{i=1}^n x_i f(x_i) &\geq f\left(\sum_{i=1}^n x_i \cdot x_i\right) = f\left(\sum_{i=1}^n x_i^2\right) \\ &= 1/\sqrt{1 - \sum_{i=1}^n x_i^2} \\ &\geq 1/\sqrt{1 - \frac{(\sum_{i=1}^n x_i)^2}{n}} \\ &= 1/\sqrt{1 - \frac{1}{n}} = 1/\sqrt{\frac{n-1}{n}} \\ &= \sqrt{\frac{n}{n-1}} \end{aligned}$$

where the second inequality follows from the result given in Exercise 7.36.  $\square$

### Exercise 7.39

Prove that for any  $a, b, c > 0$  the following inequality holds:

$$\frac{9}{a+b+c} \leq 2 \left( \frac{1}{a+b} + \frac{1}{b+c} + \frac{1}{c+a} \right)$$

To simplify the proof of Exercise 7.39, we introduce the following theorem which says that the **harmonic mean** (HM) is less than or equal to the **geometric mean** (GM).

**Theorem 7.1 (HM $\leq$ GM).** *For any  $x_1, x_2, \dots, x_n > 0$  the following inequality holds:*

$$\frac{n}{\sum_{i=1}^n \frac{1}{x_i}} \leq \sqrt[n]{\prod_{i=1}^n x_i}$$

*Proof.* According to AGM inequality, for any  $a_1, a_2, \dots, a_n \geq 0$ , we have

$$\frac{1}{n} \sum_{i=1}^n a_i \geq \sqrt[n]{\prod_{i=1}^n a_i}.$$

Replacing  $a_i$  with  $\frac{1}{x_i}$  where  $x_i > 0$  for  $i \in \{1, 2, \dots, n\}$ , we get

$$\frac{1}{n} \sum_{i=1}^n \frac{1}{x_i} \geq \sqrt[n]{\prod_{i=1}^n \frac{1}{x_i}}.$$

Since both sides are positive, taking reciprocals and reversing the inequality yield

$$\begin{aligned} \frac{1}{\frac{1}{n} \sum_{i=1}^n \frac{1}{x_i}} &\leq \frac{1}{\sqrt[n]{\prod_{i=1}^n \frac{1}{x_i}}} \\ \frac{n}{\sum_{i=1}^n \frac{1}{x_i}} &\leq \sqrt[n]{\prod_{i=1}^n x_i}, \end{aligned}$$

as desired.  $\square$

Naturally, we get the following corollary in which AM is short for the arithmetic mean.

**Corollary 7.2 (HM $\leq$ GM $\leq$ AM).** *For any  $x_1, x_2, \dots, x_n > 0$  the following inequality holds:*

$$\frac{n}{\sum_{i=1}^n \frac{1}{x_i}} \leq \sqrt[n]{\prod_{i=1}^n x_i} \leq \frac{1}{n} \sum_{i=1}^n x_i$$

*Proof.* The first inequality and the second inequality are exactly Theorem 7.1 and AGM inequality, respectively.  $\square$

Now we prove Exercise 7.39 using Corollary 7.2.

*Proof.* Since  $\text{HM} \leq \text{AM}$ , letting  $x_1 = \frac{2}{a+b}$ ,  $x_2 = \frac{2}{b+c}$  and  $x_3 = \frac{2}{c+a}$  yields

$$\begin{aligned} \frac{3}{\frac{1}{\frac{2}{a+b}} + \frac{1}{\frac{2}{b+c}} + \frac{1}{\frac{2}{c+a}}} &\leq \frac{\frac{2}{a+b} + \frac{2}{b+c} + \frac{2}{c+a}}{3} \\ \frac{3}{a+b+c} &\leq \frac{2}{3} \left( \frac{1}{a+b} + \frac{1}{b+c} + \frac{1}{c+a} \right) \\ \frac{9}{a+b+c} &\leq 2 \left( \frac{1}{a+b} + \frac{1}{b+c} + \frac{1}{c+a} \right), \end{aligned}$$

as desired.  $\square$

#### Exercise 7.40

- (i) Prove that the function  $f(x) = \frac{1}{1+e^x}$  is strictly convex over  $[0, \infty)$ .  
(ii) Prove that for any  $a_1, a_2, \dots, a_n \geq 1$  the equality

$$\sum_{i=1}^n \frac{1}{1+a_i} \geq \frac{n}{1 + \sqrt[n]{a_1 a_2 \cdots a_n}}$$

holds.

*Proof.* (i) The second derivative is given by

$$f''(x) = \frac{e^x(e^x - 1)}{(1 + e^x)^3} > 0, \quad x > 0$$

Thus,  $f(x)$  is strictly convex on  $(0, +\infty)$ . By Theorem 7.13 in the textbook,  $f''(x) > 0$  is a sufficient, not necessary, condition for strict convexity. Even though  $f''(x) = 0$  at the unique boundary point  $x = 0$ , this does not alter the strict convexity of  $f(x)$ . To see this, recall the definition of strict convexity, i.e. Definition 7.2, that is, for any  $x \neq y \in C, \lambda \in (0, 1)$ ,

$$f(\lambda x + (1 - \lambda)y) < \lambda f(x) + (1 - \lambda)f(y).$$

It is easy to see that for any  $y > x = 0$ , the above always holds for any  $\lambda \in (0, 1)$ . Thus,  $\frac{1}{1+e^x}$  is strictly convex over  $[0, +\infty)$ .

- (ii) Let  $a_i = e^{x_i}, i = 1, \dots, n$ . Then for any  $a_i \geq 1, x_i \geq 0$ . Since  $f(x) = \frac{1}{1+e^x}$  is strictly convex, then

$$\begin{aligned} \sum_{i=1}^n \frac{1}{n} \cdot \frac{1}{1+a_i} &= \sum_{i=1}^n \frac{1}{n} \cdot \frac{1}{1+e^{x_i}} \geq \frac{1}{1+e^{1/n \sum_{i=1}^n x_i}} \\ &= \frac{1}{1+(e^{\sum_{i=1}^n x_i})^{1/n}} \\ &= \frac{1}{1+(\prod_{i=1}^n e^{x_i})^{1/n}} \\ &= \frac{1}{1+(\prod_{i=1}^n a_i)^{1/n}} = \frac{1}{1+\sqrt[n]{a_1 a_2 \cdots a_n}} \end{aligned}$$

Multiplying both sides by  $n$  gives the claim, namely,

$$\sum_{i=1}^n \frac{1}{1+a_i} \geq \frac{n}{1+\sqrt[n]{a_1 a_2 \cdots a_n}}$$

224 Since  $\frac{1}{1+e^x}$  is strictly convex, the equality holds if and only if  $a_1 = a_2 = \dots = a_n = 1$ . This  
 225 completes our proof.  
 226 □

## 227 8 Chapter 8 Convex Optimization

### Exercise 8.1

Consider the problem

$$\begin{aligned} \min \quad & f(\mathbf{x}) \\ \text{s. t.} \quad & g(\mathbf{x}) \leq 0 \\ & \mathbf{x} \in X \end{aligned} \quad (\text{P})$$

where  $f$  and  $g$  are convex functions over  $\mathbb{R}^n$  and  $X \subseteq \mathbb{R}^n$  is a convex set. Suppose that  $\mathbf{x}^*$  is an optimal solution of (P) that satisfies  $g(\mathbf{x}^*) < 0$ . Show that  $\mathbf{x}^*$  is also an optimal solution of the problem

$$\begin{aligned} \min \quad & f(\mathbf{x}) \\ \text{s. t.} \quad & \mathbf{x} \in X \end{aligned}$$

228 *Proof.* We denote the feasible sets of (P) and the second problem by  $C_p$  and  $C$ , respectively. Since  
 229  $f(\mathbf{x}), g(\mathbf{x})$  and  $X$  are convex, both  $C_p$  and  $C$  are convex sets with  $C_p \subseteq C$ . Since  $g(\mathbf{x}^*) < 0$ ,  
 230  $\mathbf{x}^* \in \text{int}(C_p)$ . This indicates that the second problem has a local optimal solution on  $C_p$ , i.e.  $\mathbf{x}^*$ . By  
 231 Theorem 8.1, we know that a local minimum is also a global minimum in terms of convex optimization.  
 232 Hence,  $\mathbf{x}^*$  is also an optimal solution of the problem without the constraint of  $g(\mathbf{x}) \leq 0$ . □

### Exercise 8.2

Let  $C = B[\mathbf{x}_0, r]$ , where  $\mathbf{x}_0 \in \mathbb{R}^n$  and  $r > 0$  are given. Find a formula for the orthogonal projection operator  $P_C$ .

234

**Solution:** Given  $\mathbf{x} \in \mathbb{R}^n$ , we want to find its projection onto the closed ball  $B[\mathbf{x}_0, r]$ . Then the optimization problem associated with the computation of  $P_C(\mathbf{x})$  is given by

$$\min_{\mathbf{y}} \{ \|\mathbf{y} - \mathbf{x}\|^2 \mid \|\mathbf{y} - \mathbf{x}_0\|^2 \leq r^2 \}.$$

If  $\|\mathbf{x} - \mathbf{x}_0\| \leq r$ , then obviously  $\mathbf{y} = \mathbf{x}$  since it corresponds to the optimal value 0. When  $\|\mathbf{x} - \mathbf{x}_0\| > r$ , then the optimal solution must belong to the boundary of the ball due to Theorem 2.6 in the textbook. Specifically, Theorem 2.6 says that for a differentiable function  $f(\mathbf{x})$ , if  $\mathbf{x}^*$  is a local optimum point, then  $\nabla f(\mathbf{x}^*) = 0$ . Accordingly,

$$2(\mathbf{y} - \mathbf{x}) = 0 \iff \mathbf{y} = \mathbf{x},$$

which is impossible since  $\mathbf{x} \notin C$ . Thus, we conclude that in the case of  $\|\mathbf{x} - \mathbf{x}_0\| > r$ , the projection problem is equivalent to

$$\begin{aligned} & \min_{\mathbf{y}} \{ \|\mathbf{y} - \mathbf{x}\|^2 \mid \|\mathbf{y} - \mathbf{x}_0\|^2 = r^2 \} \\ \iff & \min_{\mathbf{y}} \{ \|\mathbf{y} - \mathbf{x}_0 + \mathbf{x}_0 - \mathbf{x}\|^2 \mid \|\mathbf{y} - \mathbf{x}_0\|^2 = r^2 \} \\ \iff & \min_{\mathbf{y}} \{ \|\mathbf{y} - \mathbf{x}_0\|^2 + 2\langle \mathbf{y} - \mathbf{x}_0, \mathbf{x}_0 - \mathbf{x} \rangle + \|\mathbf{x}_0 - \mathbf{x}\|^2 \mid \|\mathbf{y} - \mathbf{x}_0\|^2 = r^2 \} \\ \iff & \min_{\mathbf{y}} \{ r^2 + 2\langle \mathbf{y} - \mathbf{x}_0, \mathbf{x}_0 - \mathbf{x} \rangle + \|\mathbf{x}_0 - \mathbf{x}\|^2 \mid \|\mathbf{y} - \mathbf{x}_0\|^2 = r^2 \}. \end{aligned}$$

After dropping those terms that are not depend on  $\mathbf{y}$ , we get the equivalent form as follows.

$$\operatorname{argmin}_{\mathbf{y}} \{ \langle \mathbf{y}, \mathbf{x}_0 - \mathbf{x} \rangle \mid \|\mathbf{y} - \mathbf{x}_0\|^2 = r^2 \}$$

By the Cauchy-Schwarz inequality, the objective function can be lower bounded by

$$\langle \mathbf{y}, \mathbf{x}_0 - \mathbf{x} \rangle \geq -\|\mathbf{y}\| \|\mathbf{x}_0 - \mathbf{x}\| = -r \|\mathbf{x}_0 - \mathbf{x}\|,$$

and this lower bound can be attained at  $\mathbf{y} = r \frac{\mathbf{x} - \mathbf{x}_0}{\|\mathbf{x} - \mathbf{x}_0\|}$ . Therefore, the orthogonal projection operator  $P_C$  is

$$P_{B[\mathbf{x}_0, r]} = \begin{cases} \mathbf{x}, & \text{if } \|\mathbf{x}\| \leq r \\ r \frac{\mathbf{x} - \mathbf{x}_0}{\|\mathbf{x} - \mathbf{x}_0\|}, & \text{if } \|\mathbf{x}\| > r. \end{cases}$$

235

□

## 236 9 Chapter 9 Optimization over a Convex Set

### Exercise 9.1

Let  $f$  be a continuously differentiable convex function over a closed and convex set  $C \subseteq \mathbb{R}^n$ . Show that  $x^* \in C$  is an optimal solution of the problem

$$\min \{f(\mathbf{x}) : \mathbf{x} \in C\} \quad (\text{P})$$

if and only if

$$\langle \nabla f(\mathbf{x}), \mathbf{x}^* - \mathbf{x} \rangle \leq 0 \text{ for all } \mathbf{x} \in C.$$

237

238 The necessity is easy to show, but proving the sufficiency is hard. On Math StackExchange,  
239 Parasseux Nguyen provides a beautiful proof for the sufficiency<sup>6</sup>.

*Proof.* We first show the necessity. Since  $x^* \in C$  is an optimal solution of (P), then we have

$$f(\mathbf{x}^*) - f(\mathbf{x}) \leq 0.$$

By the convexity of  $f$ , we have

$$f(\mathbf{x}) + \langle \nabla f(\mathbf{x}), \mathbf{x}^* - \mathbf{x} \rangle \leq f(\mathbf{x}^*) \iff \langle \nabla f(\mathbf{x}), \mathbf{x}^* - \mathbf{x} \rangle \leq f(\mathbf{x}^*) - f(\mathbf{x}) \leq 0.$$

Proving the sufficiency is not trivial. For all  $\mathbf{x} \in C$ , let  $\mathbf{v} = \mathbf{x} - \mathbf{x}^*$  and then  $\mathbf{x}^* + t\mathbf{v} = (1-t)\mathbf{x}^* + t\mathbf{x} \in C$ . Define  $g(t) = f(\mathbf{x}^* + t\mathbf{v})$  on  $t \in [0, 1]$ . Since  $f$  is continuously differentiable over  $C$ , then  $g(t)$  is also continuously differentiable on  $[0, 1]$ . Furthermore,

$$\begin{aligned} g'(t) &= \langle \nabla f(\mathbf{x}^* + t\mathbf{v}), \mathbf{v} \rangle \\ &= \frac{1}{t} \langle \nabla f(\mathbf{x}^* + t\mathbf{v}), t\mathbf{v} \rangle \\ &= \frac{1}{t} \langle \nabla f(\mathbf{x}^* + t\mathbf{v}), (\mathbf{x}^* + t\mathbf{v}) - \mathbf{x}^* \rangle \\ &= -\frac{1}{t} \langle \nabla f(\mathbf{x}^* + t\mathbf{v}), \mathbf{x}^* - (\mathbf{x}^* + t\mathbf{v}) \rangle \\ &\geq 0 \end{aligned}$$

240

where the inequality follows from the premise of  $\langle \nabla f(\mathbf{x}), \mathbf{x}^* - \mathbf{x} \rangle \leq 0$  for all  $\mathbf{x} \in C$ . □

241

*Note.* It is interesting to note that from the above proof, we can see that the convexity of  $f$  is not required for the sufficiency and we only used the convexity of  $C$ .

242

<sup>6</sup><https://math.stackexchange.com/questions/4178673/if-nabla-fxt-x-x-leq-0-for-all-x-in-c-then-x-is-optimal-so?>

## 243 Bibliography

244 Chen, J., Yu, C., and Jin, L. (2019). *Mathematical Analysis, third edition*.

245 Horn, R. A. and Johnson, C. R. (2013). *Matrix Analysis, Second Edition*.

246 Jax (2016). Minkowski inequality for  $0 < p < 1$ . [https://math.stackexchange.com/questions/](https://math.stackexchange.com/questions/73294/minkowski-inequality-for-p-le-1)  
247 [73294/minkowski-inequality-for-p-le-1](https://math.stackexchange.com/questions/73294/minkowski-inequality-for-p-le-1). [Online; accessed 25-May-2022].