Lecture 3 Norms

Fang Zhu

February 17, 2023

1 Prerequisite

Definition 1.1 (Vector Norm)**.** *A norm is a function* $\|\cdot\| : \mathbb{C}^m \to \mathbb{R}$ *that assigns a real-valued length to each vector. In order to conform to a reasonable notion of length, a norm must satisfy the following three conditions. For all vectors* x *and* y *and for all scalars* $\alpha \in \mathbb{C}$*,*

(1) $\|x\| \geq 0$ *and* $\|x\| = 0$ *only if* $x = 0$*,*

(2) $\|x + y\| \leq \|x\| + \|y\|$*,*

(3) $\|\alpha x\| = |\alpha|\|x\|$*.*

Lemma 1.1. *Given a permutation matrix* $\mathbf{P} \in \mathcal{M}_{\mathbf{m}\times\mathbf{n}}$*, and a vector* $\boldsymbol{x} \in \mathbb{C}^n$*, then*

$$\|\boldsymbol{P}\boldsymbol{x}\|_p = \|\boldsymbol{x}\|_p.$$

Proof. By the definition of vector norm, we can get that

$$\|\boldsymbol{x}\|_p = \begin{cases} \left(\sum_{i=1}^m |x_i|^p)^{\frac{1}{p}}\right), & 1 \leq p < \infty; \\ \max_i\{x_i\}, & p = \infty. \end{cases}$$

It is clear that $\|\boldsymbol{x}\|_p$ won't be changed after permutation of entries. Therefore, for any permutation matrix $\boldsymbol{P}$,

$$\|\boldsymbol{P}\boldsymbol{x}\|_p = \|\boldsymbol{x}\|_p,$$

which is what we need to prove. □

Corollary 1.1. *Given matrix* $\boldsymbol{A} \in \mathcal{M}_{m\times n}$ *and two permutation matrix* $\boldsymbol{P} \in \mathcal{M}_{m\times m}, \boldsymbol{Q} \in \mathcal{M}_{n\times n}$*. Then,*

$$\|\boldsymbol{P}\boldsymbol{A}\boldsymbol{Q}\|_p = \|\boldsymbol{A}\|_p.$$

Proof. By the definition of induced norm, we can get that the LHS equals

$$LHS = \|\boldsymbol{A}\|_p = \sup_x \frac{\|\boldsymbol{A}x\|_p}{\|x\|_p} = \sup_x \frac{\|\boldsymbol{AQ}x\|_p}{\|\boldsymbol{Q}x\|_p} = \sup_x \frac{\|\boldsymbol{AQ}x\|_p}{\|x\|_p} = \|\boldsymbol{AQ}\|_p,$$

Futhermore, the RHS equals

$$RHS = \|\boldsymbol{PAQ}\|_p = \sup_x \frac{\|\boldsymbol{PAQ}x\|_p}{\|x\|_p} = \sup_x \frac{\|\boldsymbol{AQ}x\|_p}{\|x\|_p} = \|\boldsymbol{AQ}\|_p = LHS,$$

which is we need to prove. □

2 Solutions

2.1 Exercise 3.1

By the definition of *weighted p-norms*, we can get that

$$\|\boldsymbol{x}\|_{\boldsymbol{W}} = \|\boldsymbol{W}\boldsymbol{x}\|,$$

where $\|\cdot\|$ is a vector norm. It is clear that $\|\cdot\|_{\boldsymbol{W}}$ meets (2), (3) of the vector norm's definition. Furthermore, we assume that

$$\boldsymbol{W}\boldsymbol{x} = \boldsymbol{0}. \tag{$\star$}$$

Since $\boldsymbol{W}$ is non-singular, ($\star$) is true iif $\boldsymbol{x} = \boldsymbol{0}$. Then $\|x\|_{\boldsymbol{W}} = \|\boldsymbol{W}\boldsymbol{x}\| \geq 0$, and $= \|\boldsymbol{W}\boldsymbol{x}\| \geq 0$, and $\|\boldsymbol{x}\| = 0$ iif $\boldsymbol{x} = 0$, which meets condition (1) of vector norm's definition. Hence, we can conclude that $\|\cdot\|_{\boldsymbol{W}}$ is a vector norm.

2.2 Exercise 3.2

Let λ be the eigenvalue of $\boldsymbol{A}$ and $\boldsymbol{x}$ be the eigenvector w.r.t. λ, then

$$\boldsymbol{A}x = \lambda x, \tag{$\star$}$$

where $x \neq 0$. We can concentrate m's x as a matrix $\boldsymbol{X}$ with m columns, which is

$$(x, x, \cdots, x) = \boldsymbol{X} \in \mathbb{R}^{m \times n}.$$

By ($\star$), we can get that

$$\boldsymbol{A}(x, x, \cdots, x) = \boldsymbol{A}\boldsymbol{X} = \lambda \boldsymbol{X}, \tag{$\spadesuit$}$$

implement the norm on both sides of ($\spadesuit$), then

$$\|\lambda \boldsymbol{X}\| = |\lambda|\|\boldsymbol{X}\| = \|\boldsymbol{A}\boldsymbol{X}\| \leq \|\boldsymbol{A}\|\|\boldsymbol{X}\|,$$

Note that $\boldsymbol{X} \neq \boldsymbol{0}$ and hence $|\lambda| \leq \|\boldsymbol{A}\|$ for any eigenvalue λ of $\boldsymbol{A}$. Therefore

$$\rho(\boldsymbol{A}) = \max\{\lambda\} \leq \|\boldsymbol{A}\|.$$

which is exactly we want to prove.

2.3 Exercise 3.3(a)

Assume that $\boldsymbol{x} = \{x_1, \cdots, x_n\}$, then

$$\|\boldsymbol{x}\| = \max_i \{x_i\} = \sqrt{x_{\max}^2} \le \sqrt{x_1^2 + \cdots + x_m^2} = \|\boldsymbol{x}\|_2.$$

2.4 Exercise 3.3(b)

$$\|\boldsymbol{x}\|_2 = \sqrt{x_1^2 + \cdots + x_m^2} \le \sqrt{x_{\max}^2 + \cdots + x_{\max}^2} = \sqrt{m}\|\boldsymbol{x}\|_\infty.$$

2.5 Exercise 3.3(c)

$$\|\boldsymbol{A}\|_\infty = \sup \frac{\|\boldsymbol{A}\boldsymbol{x}\|_\infty}{\|\boldsymbol{x}\|_\infty} \le \sup \frac{\|\boldsymbol{A}\boldsymbol{x}\|_2}{\|\boldsymbol{x}\|_\infty} \le \sup \frac{\|\boldsymbol{A}\boldsymbol{x}\|_2}{\|\boldsymbol{x}\|_2/\sqrt{n}} = \sqrt{n}\|\boldsymbol{A}\|_2.$$

2.6 Exercise 3.4

We can divide the procedure into 2 steps. Firstly, we get the certain rows: W.L.O.G., assume that we need the first m rows of matrix $\boldsymbol{A}$, that is

$$\begin{bmatrix} \boldsymbol{e}_1^T \\ \boldsymbol{e}_2^T \\ \vdots \\ \boldsymbol{e}_m^T \end{bmatrix} \boldsymbol{A} = \boldsymbol{P},$$

where $\boldsymbol{e}_i$ is the unit vector whose i-th entry is 1, and $\boldsymbol{B}$ is the submatrix contains the first m rows of $\boldsymbol{A}$. Then we get the certain columns from $\boldsymbol{B}$. W.L.O.G. , assume that we need the first n columns of matrix $\boldsymbol{A}$, that is

$$\boldsymbol{P}[\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n] = \boldsymbol{B},$$

Hence,

$$\boldsymbol{B} = \begin{bmatrix} \boldsymbol{e}_1^T \\ \boldsymbol{e}_2^T \\ \vdots \\ \boldsymbol{e}_m^T \end{bmatrix} A[\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n].$$

Denote $[\boldsymbol{e}_1^T, \boldsymbol{e}_2^T, \cdots, \boldsymbol{e}_m^T]^T$ by $\boldsymbol{E}$, $[\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n]$ by $\boldsymbol{F}$. Then we have that

$$\|\boldsymbol{B}\| \le \|\boldsymbol{E}\boldsymbol{A}\boldsymbol{F}\| \le \|\boldsymbol{E}\|\|\boldsymbol{A}\|\|\boldsymbol{F}\|.$$

It is clear that $\|\boldsymbol{E}\| = \|\boldsymbol{F}\| = 1$, then we can get that

$$\|\boldsymbol{B}\| \le \|\boldsymbol{A}\|.$$

Remark. *We give another (non-rigorous enough for p is odd mumber) proof here, assume that*

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{B} & \boldsymbol{T} \\ \boldsymbol{Y} & \boldsymbol{Z} \end{bmatrix}$$

then we can get that

$$\begin{aligned} \|\boldsymbol{B}\|_p = \sup_{\|\boldsymbol{x}\|_p=1} \|\boldsymbol{B}\boldsymbol{x}\|_p \leq \sup_{\|\boldsymbol{x}\|_p=1} \begin{bmatrix} \boldsymbol{B} \\ \boldsymbol{Y} \end{bmatrix} \boldsymbol{x}\|_p \\ = \sup_{\|\boldsymbol{x}\|=1} \|\boldsymbol{A} \begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{0} \end{bmatrix} \|_p \quad \leq \sup_{\|\boldsymbol{u}\|_p=1} \|\boldsymbol{A}\boldsymbol{u}\|_p = \|\boldsymbol{A}\|_p. \end{aligned}$$

2.7 Exercise 3.5

Since

$$\|\boldsymbol{u}\boldsymbol{v}^\star\|_F = tr(\boldsymbol{v}\boldsymbol{u}^\star\boldsymbol{u}\boldsymbol{v}^\star) = \|u\|_2 tr(\boldsymbol{v}\boldsymbol{v}^\star) = \|\boldsymbol{u}\|_2\|\boldsymbol{v}\|_F.$$

we can get that $\|\boldsymbol{u}\boldsymbol{v}^\star\|_F = \|\boldsymbol{u}\|_F\|\boldsymbol{v}\|_F$ is not true.

2.8 Exercise 3.6(a)

We can prove by verifying these properties:

(i) $\|\boldsymbol{x}\|' = 0 \Leftrightarrow |\boldsymbol{y}^\star\boldsymbol{x}| = 0(\forall\|\boldsymbol{y}\| = 1) \Leftrightarrow \boldsymbol{x} = \boldsymbol{0}$.

(ii) $\|\alpha\boldsymbol{x}\|' = \sup_{\|\boldsymbol{y}\|=1} |\boldsymbol{y}^\star\alpha\boldsymbol{x}| = |\alpha| \sup_{\|\boldsymbol{y}\|=1} |\boldsymbol{y}^\star\boldsymbol{x}| = |\alpha|\|\boldsymbol{x}\|'$.

(iii)

$$\begin{aligned} \|\boldsymbol{x} + \boldsymbol{z}\|' &= \sup_{\|\boldsymbol{y}\|=1} |\boldsymbol{y}^\star\boldsymbol{x} + \boldsymbol{y}^\star\boldsymbol{z}| \\ &\leq \sup_{\|\boldsymbol{y}\|=1} (|\boldsymbol{y}^\star\boldsymbol{x}| + |\boldsymbol{y}^\star\boldsymbol{z}|) \\ &\leq \sup_{\|\boldsymbol{y}_1\|=1} |\boldsymbol{y}_1^\star\boldsymbol{x}| + \sup_{\|\boldsymbol{y}_2\|=1} \sup |\boldsymbol{y}_2^\star\boldsymbol{z}| \\ &= \|\boldsymbol{x}\|' + \|\boldsymbol{z}\|'. \end{aligned}$$

then we can get that $\|\cdot\|'$ is a norm.

2.9 Exercise 3.6(b)

Assume that $\boldsymbol{B} = \boldsymbol{y}\boldsymbol{z}^\star$ such that

$$\begin{cases} \boldsymbol{B}\boldsymbol{x} = \boldsymbol{y} \\ \|\boldsymbol{B}\| = 1. \end{cases}$$

which is equivalent to $\exists\boldsymbol{z} \in \mathbb{C}^m$ such that

$$\begin{cases} \boldsymbol{z}^\star\boldsymbol{x} = 1 \\ \|\boldsymbol{z}\|' = 1. \end{cases}$$

by lemma, we can get that $\exists \boldsymbol{u} \in \mathbb{C}^m$ such that

$$|\boldsymbol{u}^\star \boldsymbol{x}| = \|\boldsymbol{u}\|' \|\boldsymbol{x}\| = \|\boldsymbol{u}\|',$$

let $\alpha = \boldsymbol{u}^\star \boldsymbol{x}$, we can denote $\boldsymbol{z}$ by $\boldsymbol{z} = (1/\alpha)\boldsymbol{u}$, then

- $\|\boldsymbol{z}\|' = \|\frac{1}{\alpha}\boldsymbol{u}\|' = \frac{1}{|\alpha|}\|\boldsymbol{u}\|' = 1, \ \checkmark$
- $\boldsymbol{z}^\star \boldsymbol{x} = \frac{1}{\alpha}\boldsymbol{u}^\star \boldsymbol{x} = 1, \ \checkmark$

Hence, $\boldsymbol{z}$ is exactly what we need.