MATH 4370/7370 – Linear Algebra and Matrix Analysis

Norms and Matrix Norms

Julien Arino

Fall 2023

University of Manitoba

Outline

Definition 5.1 (Norm)

Let V be a vector space over a field $\mathbb{F}$. A function $\|\cdot\| : V \to \mathbb{R}_+$ is a **norm** if for all $\mathbf{x}, \mathbf{y} \in V$ and for all $c \in \mathbb{F}$

1. $\|\mathbf{x}\| \geq 0$ [**Nonnegativity**]
2. $\|\mathbf{x}\| = 0 \iff \mathbf{x} = \mathbf{0}$ [**Positivity**]
3. $\|c\mathbf{x}\| = |c| \, \|\mathbf{x}\|$ [**Homogeneity**]
4. $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ [**Triangle Inequality**]

Remark 5.2

If we have 1, 3, and 4 but not 2, then we have a **seminorm**

Definition 5.3 (Inner product)

Let V be a vector space over $\mathbb{F}$. A function $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{F}$ is an **inner product** if for all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$ and all $c \in \mathbb{F}$

1. $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$
2. $\langle \mathbf{x}, \mathbf{x} \rangle = 0 \iff \mathbf{x} = 0$
3. $\langle \mathbf{x} + \mathbf{y}, \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle$
4. $\langle c\mathbf{x}, \mathbf{y} \rangle = c \langle \mathbf{x}, \mathbf{y} \rangle$
5. $\langle \mathbf{x}, \mathbf{y} \rangle = \overline{\langle \mathbf{y}, \mathbf{x} \rangle}$

Theorem 5.4 (Cauchy-Schwartz)

Let $\langle \cdot, \cdot \rangle$ be an inner product on a vector space V over $\mathbb{F}$, then

$$|\langle \mathbf{x}, \mathbf{y} \rangle|^2 \leq \langle \mathbf{x}, \mathbf{x} \rangle \langle \mathbf{y}, \mathbf{y} \rangle$$

Corollary 5.5

If $\langle \cdot, \cdot \rangle$ is an inner product on a real or complex vector space V, then $\|\cdot\| : V \to \mathbb{R}_+$ defined by $\|\mathbf{x}\| = \langle \mathbf{x}, \mathbf{x} \rangle^{1/2}$ is a norm on V

Remark 5.6

If $\langle \cdot, \cdot \rangle$ is a semi-inner product, then the resulting $\|\mathbf{x}\| = \langle \mathbf{x}, \mathbf{x} \rangle^{1/2}$ is a seminorm

Theorem 5.7

Consider the norm $\|\cdot\|$. Then $\|\cdot\|$ is derived from an inner product if and only if it satisfies the parallelogram identity

$$\frac{1}{2}\left(\|\mathbf{x}+\mathbf{y}\|^2+\|\mathbf{x}-\mathbf{y}\|^2\right)=\|\mathbf{x}\|^2+\|\mathbf{y}\|^2$$

Theorem 5.8

If $\|\cdot\|$ is a nom on $\mathbb{C}^n$ and a matrix $T \in \mathcal{M}_n$ which is non-singular. Then

$$\|\mathbf{x}\|_T = \|T\mathbf{x}\|$$

is also a norm on $\mathbb{C}^n$

Definition 5.9

Let V be a vector space over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. Take a norm $\|\cdot\|$ on V. The sequence $\{\mathbf{x}^{(k)}\}$ of vectors in V converges to $\mathbf{x} \in V$ with respect to the norm $\|\cdot\|$ if and only if $\|\mathbf{x}^{(k)} - \mathbf{x}\| \to 0$ as $k \to \infty$

We write $\lim\limits_{k\to\infty} \mathbf{x}^{(k)} = \mathbf{x}$ with respect to $\|\cdot\|$ or

$$\mathbf{x}^{(k)} \xrightarrow{\|\cdot\|} \mathbf{x}$$

Theorem 5.10

Every (vector) norm in $\mathbb{C}^n$ is uniformly continuous

Corollary 5.11

Let $\|\cdot\|_\alpha$ and $\|\cdot\|_\beta$ be any two norms on a finite-dimensional vector space V. Then there exist $C_m, C_r > 0$ such that

$$C_m\|\mathbf{x}\|_\alpha \leq \|\mathbf{x}\|_\beta \leq C_r\|\mathbf{x}\|_\alpha, \forall \mathbf{x} \in V$$

Corollary 5.12

Let $\|\cdot\|_\alpha$ and $\|\cdot\|_\beta$ norms on a finite-dimensional vector space V over $\mathbb{R}$ or $\mathbb{C}$, $\{\mathbf{x}^{(k)}\}$ a given sequence in V, then

$$\mathbf{x}^{(k)} \xrightarrow{\|\cdot\|_\alpha} \mathbf{x} \iff \mathbf{x}^{(k)} \xrightarrow{\|\cdot\|_\beta} \mathbf{x}$$

Definition 5.13 (Equivalent norms)

Two norms are **equivalent** if whenever a sequence $\{\mathbf{x}^{(k)}\}$ converges to $\mathbf{x}$ with respect to one of the norm, it converges to $\mathbf{x}$ in the other norm

Theorem 5.14

In finite-dimensional vector spaces, all norm are equivalent

Definition 5.15 (Dual norm)

Let f be a pre-norm on $V = \mathbb{R}^n$ or $\mathbb{C}^n$. The function

$$f_d = (\mathbf{y}) \max_{f(\mathbf{x})=1} \operatorname{Re} \mathbf{y}^* \mathbf{x}$$

is the **dual norm** of f

Remark 5.16

The dual norm is well defined. $\operatorname{Re} \mathbf{y}^* \mathbf{x}$ *is a continuous function for all* $\mathbf{y} \in V$ *fixed. The set* $\{f(\mathbf{x}) = 1\}$ *is compact*

Equivalent definition for dual norm: $f^D(\mathbf{y}) = \max_{f(\mathbf{x})=1} |\mathbf{y}^* \mathbf{x}|$

Lemma 5.17 (Extension of Cauchy-Schwartz)

Let f be a prenorm on $V = \mathbb{R}^n$ or $\mathbb{C}^n$ for all $\mathbf{x}, \mathbf{y} \in V$. Then

$$|\mathbf{y}^*\mathbf{x}\,| \leq f(\mathbf{x}) f^D(\mathbf{y})$$
$$|\mathbf{y}^*\mathbf{x}| \leq f^D(\mathbf{x}) f(\mathbf{x})$$

Remark 5.18

- *The dual norm of a pre-norm is a norm*
- *The only norm that equals its dual norm is the Euclidean norm*

Theorem 5.19

Let $\|\cdot\|$ be a norm on $\mathbb{C}^n$ or $\mathbb{R}^n$, and $\|\cdot\|^D$ its dual, $c > 0$ given. Then for all $\mathbf{x} \in V$, $\|\mathbf{x}\| = c\|\mathbf{x}\|^d \iff \|\cdot\| = \sqrt{c}\|\cdot\|^d$. In particular, $\|\cdot\| = \|\cdot\|^2 \iff \|\cdot\| = \|\cdot\|_2$

Definition 5.20

Let $x \in \mathbb{F}^n$. Denote $|x| = [|x_i|]$ ($|\cdot|$ entry-wise), and write that $|x| \leq |y|$ if $|x_i| \leq |y_i|$ for all $i = 1, \ldots, n$. Assume $\|\cdot\|$ is

1. monotone if $|\mathbf{x}| \leq |\mathbf{y}| \implies \|\mathbf{x}\| \leq \|\mathbf{y}\|$ for all $\mathbf{x}, \mathbf{y}$
2. absloute if $\||\mathbf{x}|\|$ for all $\mathbf{x} \in V$

Theorem 5.21

Let $\|\cdot\|$ be a norm on $\mathbb{F}^n$. Then

1. *If $\|\cdot\|$ is absolute, then*

$$\|\mathbf{y}\|^D = \max_{\mathbf{x}\neq 0} = \frac{|\mathbf{y}|^T|\mathbf{x}|}{\|\mathbf{x}\|}$$

   for all $\mathbf{y} \in V$

2. *If $\|\cdot\|$ absolute, then $\|\cdot\|^D$ is absolute and monotone*
3. *$\|\cdot\|$ absolute if and only if $\|\cdot\|$*

Definition 5.22 (Matrix norm)

Let $\|\cdot\|$ be a function from $\mathcal{M}_n \to \mathbb{R}$. $\|\cdot\|$ is a **matrix norm** if for all $A, B \in \mathcal{M}_n$ and $c \in \mathbb{C}$, it satisfies the following

1. $\|A\| \geq 0$ [**nonnegativity**]
2. $\|A\| = 0 \iff A = 0$ [**positivity**]
3. $\|cA\| = |c| \; \|A\|$ [**homogeneity**]
4. $\|A + B\| \leq \|A\| + \|B\|$ [**triangle inequality**]
5. $\|AB\| \leq \|A\|\|B\|$ [**submultiplicativity**]

Remark 5.23

As with vector norms, if property 2 does not hold, $\|\cdot\|$ is a **matrix semi-norm**

Remark 5.24

$\|A^2\| = \|AA\| \leq \|A\|^2$ *[for any matrix norm].*
If $A^2 = A$, *then*

$$\|A^2\| = \|A\| \leq \|A\|^2 \implies \|A\| \geq 1.$$

In particular, $\|I\| \geq 1$ *for any matrix norm.*
Assume that A *is invertible, then* $AA^{-1} = I$, *thus*

$$\|I\| = \|AA^{-1}\| \leq \|A\|\|A^{-1}\| \tag{1}$$

$$\|A^{-1}\| \geq \frac{\|I\|}{\|A\|} \tag{2}$$

Definition 5.25 (Induced matrix norm)

Let $\|\cdot\|$ be a norm on $\mathbb{C}^n$. Define $|||\cdot|||$ on $\mathcal{M}_n(\mathbb{C})$ by

$$|||A||| = \max_{\|\mathbf{x}\|=1} \|A\mathbf{x}\|$$

Then $|||\cdot|||$ is the **matrix norm induced** by $\|\cdot\|$

Theorem 5.26

The function $|||\cdot|||$ defined in Definition 5.25 has the following properties

1. $|||\mathbb{I}||| = 1$
2. $\|A\mathbf{y}\| \leq |||A|||\|\mathbf{y}\|$ *for all* $A \in \mathcal{M}_n(\mathbb{C})$ *and all* $\mathbf{y} \in \mathbb{C}^n$
3. $|||\cdot|||$ *is a matrix norm on* $\mathcal{M}_n(\mathbb{C})$.
4. $|||A||| = \max_{\|\mathbf{x}\|=\|\mathbf{y}\|^D} |\mathbf{y}^* A\mathbf{x}|$

Definition 5.27 (Induced norm/Operator norm)

$\|\cdot\|$ defined from $\| \cdot \|$ by any of the previous methods is the matrix norm induced by $\| \cdot \|$. It is also called the **operator norm**

Definition 5.28 (Unital norm)

A norm such that $\|\mathbb{I}\| = 1$ is **unital**

Remark 5.29

Every induced matrix norm is unital. Every induced norm is a matrix norm

Proposition 5.30

For all U, V unitary matrices, we have $\|UAV\|_2 = \|A\|_2$

Theorem 5.31

Let $\|\cdot\|$ be a matrix norm in $\mathcal{M}_n$ and let $S \in \mathcal{M}_n$ be nonsingular. Then for all $A \in \mathcal{M}_n$, $\|A\|_S = \|SAS^{-1}\|$ is a matrix norm. Furthermore, if $\|\cdot\|$ on $\mathbb{C}^n$, then $\|\mathbf{x}\|_S = \|S\mathbf{x}\|$ induces $\|\cdot\|_S$ on $\mathcal{M}_n$

Theorem 5.32

Let $\|\cdot\|$ be a matrix norm on $\mathcal{M}_n$, $A \in \mathcal{M}_n$ and $\lambda \in \sigma(A)$. Then

1. $|\lambda| \leq \rho(A) \leq \|A\|$
2. *If A is nonsingular, then*
$$\rho(A) \geq |\lambda| \geq \frac{1}{\|A^{-1}\|}$$

Lemma 5.33

Let $A \in \mathcal{M}_n$. If there exists a norm $\|\cdot\|$ on $\mathcal{M}_n$ such that $\|A\| < 1$, then $\lim_{k\to\infty} A^k = 0$ entry-wise

Remark 5.34

When $\|A\| < 1$ for some norm, we say that A is **convergent**

Theorem 5.35

Let $A \in \mathcal{M}_n$, then

$$\lim_{k\to\infty} A^k = 0 \iff \rho(A) < 1$$

Theorem 5.36 (Gelfand Formula)

Let $\|\cdot\|$ be a matrix norm on $\mathcal{M}_n$, let $A \in \mathcal{M}_n$. Then

$$\rho(A) = \lim_{k\to\infty} \|A^k\|^{1/k}$$

Theorem 5.37

Let R be the radius of convergence of the (scalar) power series $\sum_{k=0}^{\infty} a_k z^k$ and $A \in \mathcal{M}_n$.

Then the matrix power series $\sum_{k=1}^{\infty} a_k A^k$ converges if $\rho(A) < R$

Remark 5.38

The convergence condition for the matrix power series is "there exists a matrix norm $\|\cdot\|$ such that $\|A\| < R$"

Corollary 5.39

Let $A \in \mathcal{M}_n$ be nonsingular, if there $\|\cdot\|$ matrix norm such that $\|\mathbb{I} - A\| \leq 1$

Corollary 5.40

Let $A \in \mathcal{M}_n$ is such that $|a_{ii}| > \sum_{j \neq i} |a_{ij}|$ for all $i = 1, \ldots, n$. Then A is invertible

Let $V = \mathcal{M}_{mn}(\mathbb{C})$ with Frobenius inner product

$$\langle A, B \rangle_F = \mathrm{tr}(B^* A)$$

The norm derived from the Frobenius inner product is

$$\|A\|_2 = (\mathrm{tr}(A^* A))^{1/2}$$

is the ℓ-2 norm (or Frobenius norm)

The spectral norm $|||\cdot|||$ defined on $\mathcal{M}_n$ by

$$|||A|||_2 = \sigma_1(A),$$

where $\sigma_1(A)$ is the largest singular value of A is induced by the ℓ-2 norm on $\mathbb{C}^n$. Inded, from the singular value decomposition theorem, let

$$A = V\Sigma W^*$$

be a singular value decomposition of A, where V, W unitary, $\Sigma = \sigma(\sigma_1, \ldots, \sigma_n)$ and $\sigma_1 \geq \cdots \geq \sigma_n \geq 0$ are the non-increasingly ordered singular values of A

From unitary invariance and monotonicity of the Euclidean norm, we say that

$$
\begin{aligned}
\max \|Ax\|_1 &= \max_{\|x\|_1} \|V\Sigma W^*\|_2 \\
&= \max_{\|x\|_2} \|\Sigma W^* x\|_2 \\
&= \max_{\|Wy\|_2=1} \|\Sigma y\|_2 \\
&= \max_{\|y\|_2} \|\Sigma y\|_2 \\
&\leq \max_{\|y\|_2} \|\sigma_1 y\|_2 \\
&= \sigma_1 \max_{\|y\|_2} \|y\|_2 \\
&= \sigma_1
\end{aligned}
$$

Since $\|\Sigma y\|_2 = \sigma_1$ for $y = e_1$,

$$
\max_{\|x\|_2=1} \|Ax\|_2 = \sigma_1(A)
$$

We could have used

$$\max_{\|x\|_2=1} = \|Ax\|_2^2 = \max_{\|x\|_2=1} x^*A^*AX$$
$$= \lambda_{max}(A^*A)$$
$$= \sigma_1(A)$$

Remark 5.41

For all U, V unitary $\mathcal{M}_n$ matrices, for all $A \in \mathcal{M}_n$, $|||UAV|||_2 = |||A|||_2$

References I