

# Lecture 1: Inner product, Orthogonality, Vector/Matrix norm



School of Mathematical Sciences, Xiamen University

## 1. Inner product on a linear space $\mathbb{V}$ over a number field $\mathbb{F}$ ( $\mathbb{C}$ or $\mathbb{R}$ )

- **Definition:** A function  $\langle \cdot, \cdot \rangle : \mathbb{V} \times \mathbb{V} \rightarrow \mathbb{F}$  is called an *inner product*, if it satisfies the following three conditions ( $\forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{V}, \forall \alpha \in \mathbb{F}$ ):

(1) Conjugate symmetry:

$$\langle \mathbf{x}, \mathbf{y} \rangle = \overline{\langle \mathbf{y}, \mathbf{x} \rangle}$$

(2) Positive definiteness:

$$\langle \mathbf{x}, \mathbf{x} \rangle \geq 0, \quad \langle \mathbf{x}, \mathbf{x} \rangle = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$$

(3) Linearity in the first variable:

$$\langle \mathbf{x} + \mathbf{y}, \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle, \quad \langle \alpha \mathbf{x}, \mathbf{y} \rangle = \alpha \langle \mathbf{x}, \mathbf{y} \rangle$$

**Example:** the standard inner product on the space  $\mathbb{V} = \mathbb{C}^m$ :

$$\forall \mathbf{x}, \mathbf{y} \in \mathbb{C}^m, \quad \langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{y}^* \mathbf{x} = \sum_{i=1}^m x_i \bar{y}_i.$$

**Example:** the  $\mathbf{A}$ -inner product on the space  $\mathbb{V} = \mathbb{C}^m$ :

$$\forall \mathbf{x}, \mathbf{y} \in \mathbb{C}^m, \quad \langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{y}^* \mathbf{A} \mathbf{x},$$

where  $\mathbf{A}$  is a given Hermitian positive definite matrix.

## 2. Orthogonality

- Orthogonality is a mathematical concept with respect to a given inner product  $\langle \cdot, \cdot \rangle$ .
  - (1) Two vectors  $\mathbf{x}$  and  $\mathbf{y}$  are called *orthogonal*: if  $\langle \mathbf{x}, \mathbf{y} \rangle = 0$
  - (2) Two sets of vectors  $\mathcal{X}$  and  $\mathcal{Y}$  are called orthogonal: if  $\forall \mathbf{x} \in \mathcal{X}$  and  $\mathbf{y} \in \mathcal{Y}$ ,  $\langle \mathbf{x}, \mathbf{y} \rangle = 0$
  - (3) A set of nonzero vectors  $\mathcal{S}$  is orthogonal if  $\forall \mathbf{x}, \mathbf{y} \in \mathcal{S}$  and  $\mathbf{x} \neq \mathbf{y}$ ,  $\langle \mathbf{x}, \mathbf{y} \rangle = 0$ ; if further  $\forall \mathbf{x} \in \mathcal{S}$ ,  $\langle \mathbf{x}, \mathbf{x} \rangle = 1$ ,  $\mathcal{S}$  is *orthonormal*.

### Proposition 1

*The vectors in an orthogonal set  $\mathcal{S}$  are linearly independent.*

## 2.1. Orthogonal components of a vector

- Inner products can be used to decompose arbitrary vectors into orthogonal components. Given an *orthonormal* set  $\{\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_n\}$  and an arbitrary vector  $\mathbf{v}$ , let

$$\mathbf{r} = \mathbf{v} - \langle \mathbf{v}, \mathbf{q}_1 \rangle \mathbf{q}_1 - \langle \mathbf{v}, \mathbf{q}_2 \rangle \mathbf{q}_2 - \cdots - \langle \mathbf{v}, \mathbf{q}_n \rangle \mathbf{q}_n.$$

Obviously,

$$\mathbf{r} \in \text{span}\{\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_n\}^\perp.$$

Thus we see that  $\mathbf{v}$  can be decomposed into  $n + 1$  orthogonal components:

$$\mathbf{v} = \mathbf{r} + \langle \mathbf{v}, \mathbf{q}_1 \rangle \mathbf{q}_1 + \langle \mathbf{v}, \mathbf{q}_2 \rangle \mathbf{q}_2 + \cdots + \langle \mathbf{v}, \mathbf{q}_n \rangle \mathbf{q}_n.$$

We call  $\langle \mathbf{v}, \mathbf{q}_i \rangle \mathbf{q}_i$  the part of  $\mathbf{v}$  in the direction of  $\mathbf{q}_i$ , and  $\mathbf{r}$  the part of  $\mathbf{v}$  orthogonal to the subspace  $\text{span}\{\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_n\}$ .

**Exercise:** Write the expression for  $\mathbf{v}$  when the set  $\{\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_n\}$  is only orthogonal.

- **Cauchy–Schwarz inequality:** For any given inner product  $\langle \cdot, \cdot \rangle$ ,

$$|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} \sqrt{\langle \mathbf{y}, \mathbf{y} \rangle}.$$

The equality holds if and only if  $\mathbf{x}$  and  $\mathbf{y}$  are linearly dependent.

**Exercise:** Prove the inequality. Hint: write

$$\mathbf{x} = \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\langle \mathbf{y}, \mathbf{y} \rangle} \mathbf{y} + \mathbf{z}.$$

Then  $\langle \mathbf{z}, \mathbf{y} \rangle = 0$ . Consider  $\langle \mathbf{x}, \mathbf{x} \rangle$ .

**Application:** For any Hermitian positive definite matrix  $\mathbf{A}$ ,

$$|\mathbf{y}^* \mathbf{A} \mathbf{x}|^2 \leq (\mathbf{x}^* \mathbf{A} \mathbf{x})(\mathbf{y}^* \mathbf{A} \mathbf{y}).$$

### 3. Norm on a linear space $\mathbb{V}$ over a number field $\mathbb{F}$ ( $\mathbb{C}$ or $\mathbb{R}$ )

- **Definition:** A function  $\|\cdot\| : \mathbb{V} \rightarrow \mathbb{R}$  is called a *norm* if it satisfies the following three conditions ( $\forall \mathbf{x}, \mathbf{y} \in \mathbb{V}$  and  $\forall \alpha \in \mathbb{F}$ ):

(1) Nonnegativity:

$$\|\mathbf{x}\| \geq 0, \quad \|\mathbf{x}\| = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$$

(2) Positive homogeneity:

$$\|\alpha\mathbf{x}\| = |\alpha|\|\mathbf{x}\|$$

(3) Triangle inequality:

$$\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$$

**Exercise:** Show that any norm is continuous.

- **More on metric, norm, and inner product**

(click!)



**Exercise:** For any given inner product  $\langle \cdot, \cdot \rangle$ , let  $\| \cdot \| = \sqrt{\langle \cdot, \cdot \rangle}$ .

(1) Prove that the function  $\| \cdot \| = \sqrt{\langle \cdot, \cdot \rangle}$  is a norm.

(2) Prove the parallelogram law

$$\|\mathbf{u} + \mathbf{v}\|^2 + \|\mathbf{u} - \mathbf{v}\|^2 = 2\|\mathbf{u}\|^2 + 2\|\mathbf{v}\|^2.$$

(3) For a set of  $n$  orthogonal (with respect to the inner product  $\langle \cdot, \cdot \rangle$ ) vectors  $\{\mathbf{x}_i\}$ , prove that

$$\left\| \sum_{i=1}^n \mathbf{x}_i \right\|^2 = \sum_{i=1}^n \|\mathbf{x}_i\|^2.$$

The function  $\| \cdot \| = \sqrt{\langle \cdot, \cdot \rangle}$  is called the norm *induced* by the inner product  $\langle \cdot, \cdot \rangle$ . Using this norm, we can write the Cauchy–Schwarz inequality as

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

## Theorem 2 (Equivalence of norms)

For each pair of norms  $\|\cdot\|_\alpha$  and  $\|\cdot\|_\beta$  on a finite-dimensional linear space  $\mathbb{V}$ , there exist positive constants  $a > 0$  and  $b > 0$  (depending only on the norms) such that

$$a \leq \frac{\|\mathbf{x}\|_\alpha}{\|\mathbf{x}\|_\beta} \leq b, \quad \forall \mathbf{x} (\neq \mathbf{0}) \in \mathbb{V}.$$

### Proof.

$\forall \mathbf{x} = \sum_i x_i \mathbf{v}_i$ , where  $\{\mathbf{v}_i\}$  is a basis of  $\mathbb{V}$ ,  $\|\mathbf{x}\| = \sum_i |x_i|$  is a norm on  $\mathbb{V}$ . By  $\|\mathbf{x}\|_\alpha = \|\sum_i x_i \mathbf{v}_i\|_\alpha \leq \sum_i |x_i| \|\mathbf{v}_i\|_\alpha \leq \|\mathbf{x}\| \cdot \max_i \|\mathbf{v}_i\|_\alpha$ , we know  $\|\cdot\|_\alpha$  is a continuous function with respect to  $\|\cdot\|$ , which attains its minimum  $c$  and maximum  $C$  on the unit sphere  $\{\mathbf{x} \in \mathbb{V}, \|\mathbf{x}\| = 1\}$  (because it is a compact set). Then,  $\forall \mathbf{x} \in \mathbb{V}$ ,  $c\|\mathbf{x}\| \leq \|\mathbf{x}\|_\alpha \leq C\|\mathbf{x}\|$ .  $\square$

- Convergence of a sequence  $\{\mathbf{x}_k\} \subset \mathbb{V}$ :  $\mathbf{x}_k \rightarrow \mathbf{x}$

We say  $\mathbf{x}_k$  converges to  $\mathbf{x}$  if  $\lim_{k \rightarrow \infty} \|\mathbf{x}_k - \mathbf{x}\| = 0$ .



### 3.1. Vector norms on $\mathbb{C}^m$

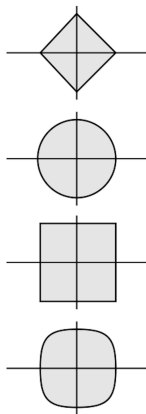
- $\ell_p$ -norm

$$\|\mathbf{x}\|_1 = \sum_{i=1}^m |x_i|,$$

$$\|\mathbf{x}\|_2 = \left( \sum_{i=1}^m |x_i|^2 \right)^{1/2} = \sqrt{\mathbf{x}^* \mathbf{x}},$$

$$\|\mathbf{x}\|_\infty = \max_{1 \leq i \leq m} |x_i|,$$

$$\|\mathbf{x}\|_p = \left( \sum_{i=1}^m |x_i|^p \right)^{1/p}, \quad (1 \leq p < \infty)$$



- Inequalities:  $\|\mathbf{x}\|_2 \leq \|\mathbf{x}\|_1 \leq \sqrt{m} \|\mathbf{x}\|_2,$

$$\|\mathbf{x}\|_\infty \leq \|\mathbf{x}\|_2 \leq \sqrt{m} \|\mathbf{x}\|_\infty, \quad \|\mathbf{x}\|_\infty \leq \|\mathbf{x}\|_1 \leq m \|\mathbf{x}\|_\infty.$$

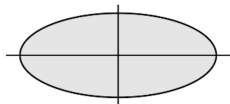
- Weighted norm

Let  $\|\cdot\|$  denote any norm on  $\mathbb{C}^m$ . Suppose a diagonal matrix  $\mathbf{W} = \text{diag}\{w_1, \dots, w_m\}$ ,  $w_i \neq 0$ . Then

$$\|\mathbf{x}\|_{\mathbf{W}} = \|\mathbf{W}\mathbf{x}\|$$

is a norm, called *weighted norm*. For example, weighted 2-norm

$$\|\mathbf{x}\|_{\mathbf{W}} = \|\mathbf{W}\mathbf{x}\|_2 = \left( \sum_{i=1}^m |w_i x_i|^2 \right)^{1/2}.$$



- Dual norm

Let  $\|\cdot\|$  denote any norm on  $\mathbb{C}^m$ . The corresponding *dual norm*  $\|\cdot\|'$  (with respect to an inner product  $\langle \cdot, \cdot \rangle$ ) is defined by

$$\|\mathbf{x}\|' = \sup_{\mathbf{y} \in \mathbb{C}^m, \|\mathbf{y}\|=1} |\langle \mathbf{x}, \mathbf{y} \rangle|.$$

### 3.2. Matrix norms on $\mathbb{C}^{m \times n}$

- Frobenius norm:  $\forall \mathbf{A} = [\mathbf{a}_1 \quad \mathbf{a}_2 \quad \cdots \quad \mathbf{a}_n] \in \mathbb{C}^{m \times n}$ , define

$$\|\mathbf{A}\|_F := \left( \sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2 \right)^{1/2} = \left( \sum_{j=1}^n \|\mathbf{a}_j\|_2^2 \right)^{1/2}$$

or

$$\|\mathbf{A}\|_F = \sqrt{\text{tr}(\mathbf{A}^* \mathbf{A})} = \sqrt{\text{tr}(\mathbf{A} \mathbf{A}^*)}.$$

- Max norm:

$$\|\mathbf{A}\|_{\max} := \max_{i,j} |a_{ij}|.$$

- Induced matrix norm (operator norm):  $\forall \mathbf{A} \in \mathbb{C}^{m \times n}$ , define

$$\|\mathbf{A}\|_{\alpha, \beta} := \sup_{\substack{\mathbf{x} \in \mathbb{C}^n \\ \mathbf{x} \neq \mathbf{0}}} \frac{\|\mathbf{A}\mathbf{x}\|_{\alpha}}{\|\mathbf{x}\|_{\beta}} = \sup_{\substack{\mathbf{x} \in \mathbb{C}^n \\ \|\mathbf{x}\|_{\beta}=1}} \|\mathbf{A}\mathbf{x}\|_{\alpha} = \sup_{\substack{\mathbf{x} \in \mathbb{C}^n \\ \|\mathbf{x}\|_{\beta} \leq 1}} \|\mathbf{A}\mathbf{x}\|_{\alpha},$$

where  $\|\cdot\|_{\alpha}$  is a norm on  $\mathbb{C}^m$  and  $\|\cdot\|_{\beta}$  is a norm on  $\mathbb{C}^n$ . We say that  $\|\cdot\|_{\alpha, \beta}$  is the matrix norm induced by  $\|\cdot\|_{\alpha}$  and  $\|\cdot\|_{\beta}$ .

**Exercise:**  $\forall \mathbf{x} \in \mathbb{C}^n$ , prove that

$$\|\mathbf{Ax}\|_\alpha \leq \|\mathbf{A}\|_{\alpha,\beta} \|\mathbf{x}\|_\beta.$$

**Exercise:** Let  $\mathbf{A} \in \mathbb{C}^{m \times n}$ ,  $\mathbf{B} \in \mathbb{C}^{n \times r}$  and let  $\|\cdot\|_\alpha$ ,  $\|\cdot\|_\beta$ , and  $\|\cdot\|_\gamma$  be norms on  $\mathbb{C}^m$ ,  $\mathbb{C}^n$ , and  $\mathbb{C}^r$ , respectively. Prove the induced matrix norms  $\|\cdot\|_{\alpha,\gamma}$ ,  $\|\cdot\|_{\alpha,\beta}$ , and  $\|\cdot\|_{\beta,\gamma}$  satisfy

$$\|\mathbf{AB}\|_{\alpha,\gamma} \leq \|\mathbf{A}\|_{\alpha,\beta} \|\mathbf{B}\|_{\beta,\gamma}.$$

**Exercise:** Prove that

$$\|\mathbf{A}\|_{\infty,1} = \max_{i,j} |a_{ij}|,$$

i.e.,  $\|\mathbf{A}\|_{\max}$  is the matrix norm induced by  $\|\cdot\|_\infty$  and  $\|\cdot\|_1$ .

- The Frobenius norm  $\|\cdot\|_F$  on  $\mathbb{C}^{m \times n}$  is not induced by norms on  $\mathbb{C}^m$  and  $\mathbb{C}^n$ . (See [Ref. 1](#) and [Ref. 2](#))

- Induced matrix  $p$ -norm of  $\mathbf{A} \in \mathbb{C}^{m \times n}$ : For  $p \in [1, +\infty]$ ,

$$\|\mathbf{A}\|_p := \|\mathbf{A}\|_{p,p} = \sup_{\mathbf{x} \in \mathbb{C}^n, \|\mathbf{x}\|_p=1} \|\mathbf{A}\mathbf{x}\|_p.$$

**Example:**  $p$ -norm of a diagonal matrix  $\mathbf{D} = \text{diag}\{d_1, \dots, d_m\}$

$$\|\mathbf{D}\|_p = \max_{1 \leq i \leq m} |d_i|$$

**Example:**  $1, 2, \infty$ -norm

$$\|\mathbf{A}\|_1 = \max_j \sum_i |a_{ij}|, \quad \|\mathbf{A}\|_\infty = \max_i \sum_j |a_{ij}|$$

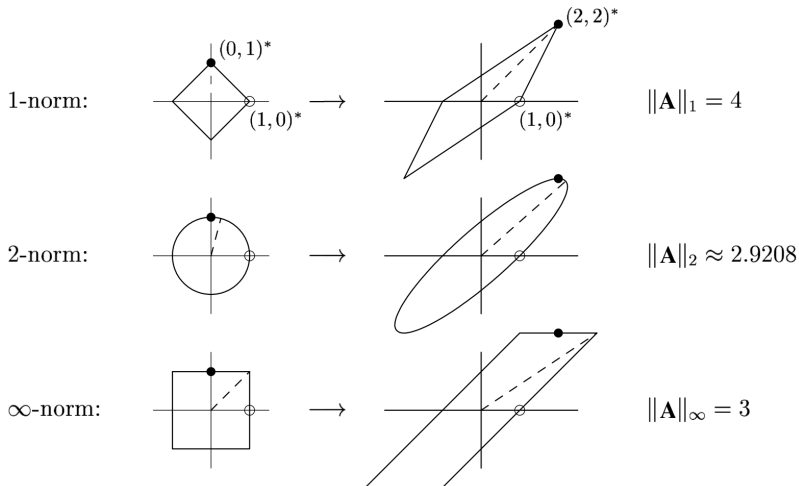
$$\|\mathbf{A}\|_2 = \sqrt{\lambda_{\max}(\mathbf{A}^* \mathbf{A})} = \sqrt{\lambda_{\max}(\mathbf{A} \mathbf{A}^*)} \leq \|\mathbf{A}\|_F$$

The norm  $\|\cdot\|_2$  on  $\mathbb{C}^{m \times n}$  is also called the spectral norm.

Inequalities:  $\|\mathbf{A}\|_\infty \leq \sqrt{n} \|\mathbf{A}\|_2, \quad \|\mathbf{A}\|_2 \leq \sqrt{m} \|\mathbf{A}\|_\infty.$

- Matlab: `norm` for  $1, 2, \infty$ -norm

Example:  $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix}$



### 3.3. Unitary invariance of $\|\cdot\|_2$ and $\|\cdot\|_F$ : $\forall \mathbf{A} \in \mathbb{C}^{m \times n}$

- If  $\mathbf{P}$  has orthonormal columns, i.e.,

$$\mathbf{P} \in \mathbb{C}^{p \times m}, \quad p \geq m, \quad \mathbf{P}^* \mathbf{P} = \mathbf{I}_m,$$

then

$$\|\mathbf{P}\mathbf{A}\|_2 = \|\mathbf{A}\|_2, \quad \|\mathbf{P}\mathbf{A}\|_F = \|\mathbf{A}\|_F.$$

- If  $\mathbf{Q}$  has orthonormal rows, i.e.,

$$\mathbf{Q} \in \mathbb{C}^{n \times q}, \quad n \leq q, \quad \mathbf{Q}\mathbf{Q}^* = \mathbf{I}_n,$$

then

$$\|\mathbf{A}\mathbf{Q}\|_2 = \|\mathbf{A}\|_2, \quad \|\mathbf{A}\mathbf{Q}\|_F = \|\mathbf{A}\|_F.$$

## 4. Unitary matrix

- For  $\mathbf{Q} \in \mathbb{C}^{m \times m}$ , if  $\mathbf{Q}^* = \mathbf{Q}^{-1}$ , i.e.,  $\mathbf{Q}^* \mathbf{Q} = \mathbf{I}$ ,  $\mathbf{Q}$  is called *unitary* (or *orthogonal* in the real case).

$$\begin{bmatrix} \mathbf{q}_1^* \\ \mathbf{q}_2^* \\ \vdots \\ \mathbf{q}_m^* \end{bmatrix} \begin{bmatrix} | & | & & | \\ \mathbf{q}_1 & \mathbf{q}_2 & \cdots & \mathbf{q}_m \\ | & | & & | \end{bmatrix} = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix}$$

**Exercise:** Let  $\mathbf{Q} \in \mathbb{C}^{m \times m}$  be a unitary matrix. Prove

$$\|\mathbf{Q}\|_2 = 1, \quad \|\mathbf{Q}\|_F = \sqrt{m}.$$

- A unitary matrix has both orthonormal rows and orthonormal columns.
- The columns of a unitary matrix form an orthonormal basis of  $\mathbb{C}^m$ .