



Inner Product and Orthogonality

CE282: Linear Algebra

Computer Engineering Department

Sharif University of Technology

Hamid R. Rabiee

Maryam Ramezani



Definition

Suppose V and W are vector spaces over the same field \mathbb{F} . Then a function $f: V \times W \rightarrow \mathbb{F}$ is called a **bilinear form** if it satisfies the following properties:

- a) It is linear in its first argument:
 - i. $f(\mathbf{v}_1 + \mathbf{v}_2, \mathbf{w}) = f(\mathbf{v}_1, \mathbf{w}) + f(\mathbf{v}_2, \mathbf{w})$ and
 - ii. $f(c\mathbf{v}_1, \mathbf{w}) = cf(\mathbf{v}_1, \mathbf{w})$ for all $c \in \mathbb{F}$, $\mathbf{v}_1, \mathbf{v}_2 \in V$, and $\mathbf{w} \in W$.
- b) It is linear in its second argument:
 - i. $f(\mathbf{v}, \mathbf{w}_1 + \mathbf{w}_2) = f(\mathbf{v}, \mathbf{w}_1) + f(\mathbf{v}, \mathbf{w}_2)$ and
 - ii. $f(\mathbf{v}, c\mathbf{w}_1) = cf(\mathbf{v}, \mathbf{w}_1)$ for all $c \in \mathbb{F}$, $\mathbf{v} \in V$, and $\mathbf{w}_1, \mathbf{w}_2 \in W$.



Note

Let V be a vector space over a field \mathbb{F} . Then the **dual** of V , denoted by V^* , is the vector space consisting of all linear forms on V .

Example

Let V be a vector space over a field \mathbb{F} . Show that the function $g: V^* \times V \rightarrow \mathbb{F}$ defined by

$$g(f, \mathbf{v}) = f(\mathbf{v}) \text{ for all } f \in V^*, \mathbf{v} \in V$$

is a bilinear form.



□ An inner product on V is a function $\langle, \rangle: V \times V \rightarrow \mathbb{R}$ such that

1. $\langle v, v \rangle \geq 0$ for all $v \in V$.
2. $\langle v, v \rangle = 0$ if and only if $v = 0$.
3. $\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$ for all $u, v, w \in V$.
4. $\langle cw, u \rangle = c\langle w, u \rangle$ for all $u, w \in V$ and $c \in \mathbb{R}$.
5. $\langle w, v \rangle = \langle v, w \rangle$.



Definition

Suppose that $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$ and that V is a vector space over \mathbb{F} . Then an **inner product** on V is a function $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{F}$ such that the following three properties hold for all $c \in \mathbb{F}$ and all $\mathbf{v}, \mathbf{w}, \mathbf{x} \in V$:

a) $\langle \mathbf{v}, \mathbf{w} \rangle = \overline{\langle \mathbf{w}, \mathbf{v} \rangle}$ (conjugate symmetry)

b) $\langle \mathbf{v} + c\mathbf{x}, \mathbf{w} \rangle = \langle \mathbf{v}, \mathbf{w} \rangle + c\langle \mathbf{x}, \mathbf{w} \rangle$ (linearity)

c) $\langle \mathbf{v}, \mathbf{v} \rangle \geq 0$, with equality if and only if $\mathbf{v} = \mathbf{0}$.
(pos. definiteness)



Note

□ $F = \mathbb{R}$ bilinear forms

□ $F = \mathbb{C}$ sesquilinear forms—they are linear in their first argument, but only conjugate linear in their second argument

$$\langle v, w + cx \rangle = \overline{\langle w + cx, v \rangle} = \overline{\langle w, v \rangle} + \overline{c \langle x, v \rangle} = \langle v, w \rangle + \bar{c} \langle v, x \rangle$$



Example

Show that the function $\langle \cdot, \cdot \rangle : \mathbb{C}^n \times \mathbb{C}^n \rightarrow \mathbb{C}$ defined by

$$\langle \mathbf{v}, \mathbf{w} \rangle = \mathbf{v} \mathbf{w}^* = \sum_{i=1}^n v_i \overline{w_i} \quad \text{for all } \mathbf{v}, \mathbf{w} \in \mathbb{C}^n$$

is an inner product on \mathbb{C}^n .



Example

Let $a < b$ be real numbers and let $C[a, b]$ be the vector space of continuous functions on the real interval $[a, b]$.

Show that the function $\langle \cdot, \cdot \rangle : C[a, b] \times C[a, b] \rightarrow \mathbb{R}$ defined by

$$\langle f, g \rangle = \int_a^b f(x)g(x) dx \quad \text{for all } f, g \in C[a, b]$$

is an inner product on $C[a, b]$.



Example

Find $\langle p, q \rangle$, $\|p\|$, $\|p - q\|$ which $p(x) = 3 - x + 2x^2$ and $q(x) = 4x + x^2$.



Theorem

Take any inner product $\langle \cdot, \cdot \rangle$ and define $f(x) = \sqrt{\langle x, x \rangle}$. Then f is a norm.

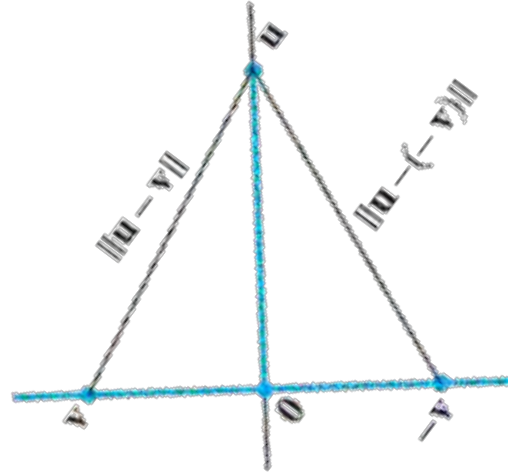
Proof

Note

Every inner product gives rise to a norm, but not every norm comes from an inner product. (Think about norm 2 and norm max)



- Geometry
- Algebra



Two vectors \mathbf{u} and \mathbf{v} in \mathbb{R}^n are **orthogonal** (to each other) if $\mathbf{u} \cdot \mathbf{v} = 0$.

Suppose V is an inner product space.

Two vectors $\mathbf{v}, \mathbf{w} \in V$ are called **orthogonal** if $\langle \mathbf{v}, \mathbf{w} \rangle = 0$.

The Pythagorean Theorem

Two vectors \mathbf{u} and \mathbf{v} are orthogonal if and only if $\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$

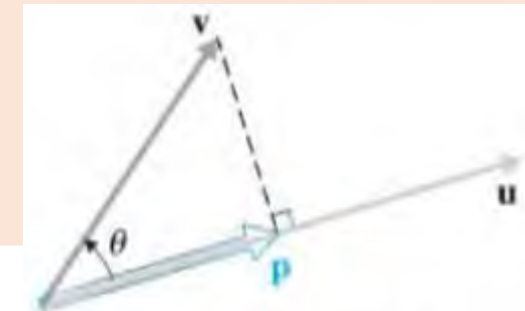
- Finding the distance from a point B to line l = Finding the length of line segment BP
- AP : projection of AB onto the line l



Definition

If \mathbf{u} and \mathbf{v} are vectors in \mathbb{R}^n and $\mathbf{u} \neq \mathbf{0}$, then the **projection of \mathbf{v} onto \mathbf{u}** is the vector $proj_{\mathbf{u}}(\mathbf{v})$ defined by

$$proj_{\mathbf{u}}(\mathbf{v}) = \left(\frac{\mathbf{u} \cdot \mathbf{v}}{\mathbf{u} \cdot \mathbf{u}} \right) \mathbf{u}$$



The projection of \mathbf{v} onto \mathbf{u}



- A set of vectors $\{a_1, \dots, a_k\}$ in R^n is **orthogonal** set if each pair of distinct vectors is orthogonal (**mutually orthogonal vectors**).

Theorem

If $S = \{a_1, \dots, a_k\}$ is an orthogonal set of nonzero vectors in R^n , then S is linearly independent and is a basis for the subspace spanned by S .

Proof

If $k = n$, then prove that S is a basis for R^n



Definition

A basis B of an inner product space V is called an **orthonormal basis** of V if

a) $\langle \mathbf{v}, \mathbf{w} \rangle = 0$ for all $\mathbf{v} \neq \mathbf{w} \in B$, and

(mutual orthogonality)

b) $\|\mathbf{v}\| = 1$ for all $\mathbf{v} \in B$.

(normalization)

□ set of n -vectors a_1, \dots, a_k are (*mutually*) *orthogonal* if $a_i \perp a_j$ for $i \neq j$

□ they are *normalized* if $\|a_i\| = 1$ for $i = 1, \dots, k$

□ they are *orthonormal* if both hold

□ can be expressed using inner products as

$$a_i^T a_j = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$



Independence-dimension inequality

If the n -vectors a_1, \dots, a_k are linearly independent, then $k \leq n$.

- orthonormal sets of vectors are linearly independent
- by independence-dimension inequality, must have $k \leq n$
- when $k = n$, a_1, \dots, a_n are an *orthonormal basis*

Example

- Standard unit n-vectors e_1, \dots, e_n

- The 3-vectors

$$\begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix}, \quad \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$$

- The 2-vectors shown below



- The standard basis in $P^p[a, b]$ (be the set of real-valued polynomials of degree at most p.)



- A simple way to check if an n -vector y is a linear combination of the orthonormal vectors a_1, \dots, a_k , if and only if:

$$y = (a_1^T y) a_1 + \dots + (a_k^T y) a_k$$

- For orthogonal vectors a_1, \dots, a_k :

$$y = c_1 a_1 + \dots + c_k a_k$$

$$c_j = \frac{y \cdot a_j}{a_j \cdot a_j}$$



Example

Write x as a linear combination of a_1, a_2, a_3 ?

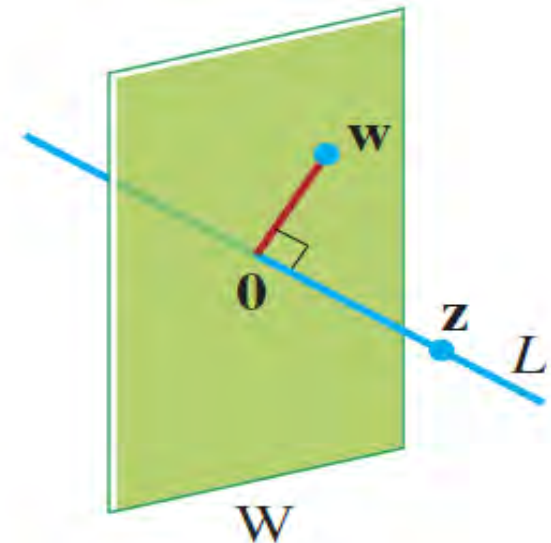
$$x = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \quad a_1 = \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix}, \quad a_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad a_3 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$$

Definition

- ❑ If a vector z is orthogonal to every vector in a subspace W of \mathbb{R}^n , then z is said to be orthogonal to W .
- ❑ **The set of all vectors z that are orthogonal to W** is called the orthogonal complement of W and is denoted by W^\perp

Example

W be a plane through the origin in \mathbb{R}^3 .
 $L = W^\perp$ and $W = L^\perp$





Note

- 1) A vector x is in W^\perp if and only if x is orthogonal to every vector in a set that spans W .
- 2) W^\perp is a subspace of \mathbb{R}^n .

Important

We emphasize that W_1 and W_2 can be orthogonal without being complements.

$W_1 = \text{span}((1, 0, 0))$ and $W_2 = \text{span}((0, 1, 0))$.

Orthogonal Projection of y onto W



The Orthogonal Decomposition Theorem

Let W be a subspace of \mathbb{R}^n . Then each y in \mathbb{R}^n can be written **uniquely** in the form:

$$y = \hat{y} + z \quad \text{where } \hat{y} = \text{proj}_W y \quad (1)$$

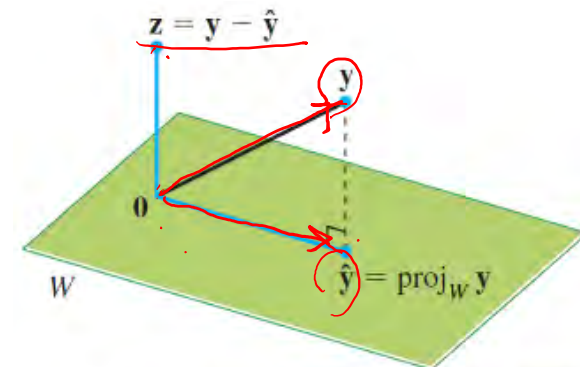
where \hat{y} is in W and z is in W^\perp . In fact, if $\{u_1, \dots, u_p\}$ is any orthogonal basis of W , then

$$\hat{y} = \frac{y \cdot u_1}{u_1 \cdot u_1} u_1 + \dots + \frac{y \cdot u_p}{u_p \cdot u_p} u_p \quad (2)$$

and $z = y - \hat{y}$

Important

The uniqueness of the decomposition (1) shows that the orthogonal projection \hat{y} depends only on W and not on the particular basis used in (2).



The orthogonal projection of y onto W .

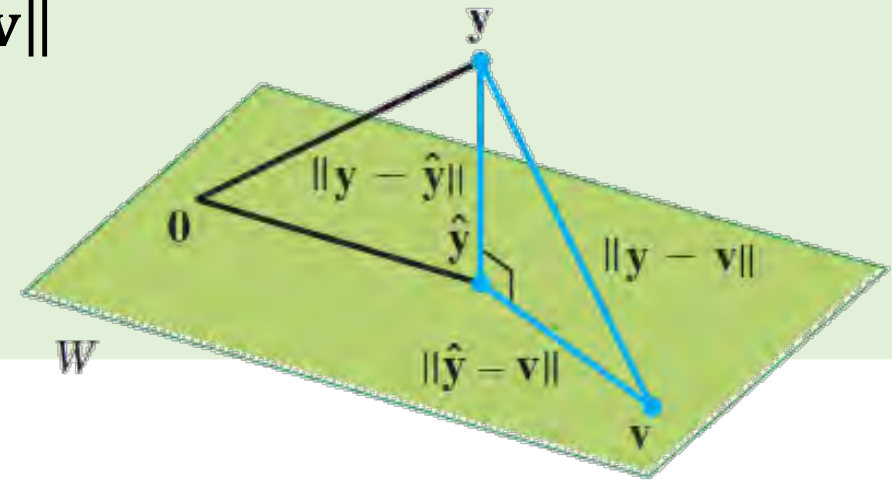
The Best Approximation Theorem

Let W be a subspace of \mathbb{R}^n . let \mathbf{y} be any vector in \mathbb{R}^n , and let $\hat{\mathbf{y}}$ be the orthogonal projection of \mathbf{y} onto W . Then $\hat{\mathbf{y}}$ is the closest point in W to \mathbf{y} , in the sense that

$$\|\mathbf{y} - \hat{\mathbf{y}}\| < \|\mathbf{y} - \mathbf{v}\|$$

for all \mathbf{v} in W distinct from $\hat{\mathbf{y}}$.

Proof

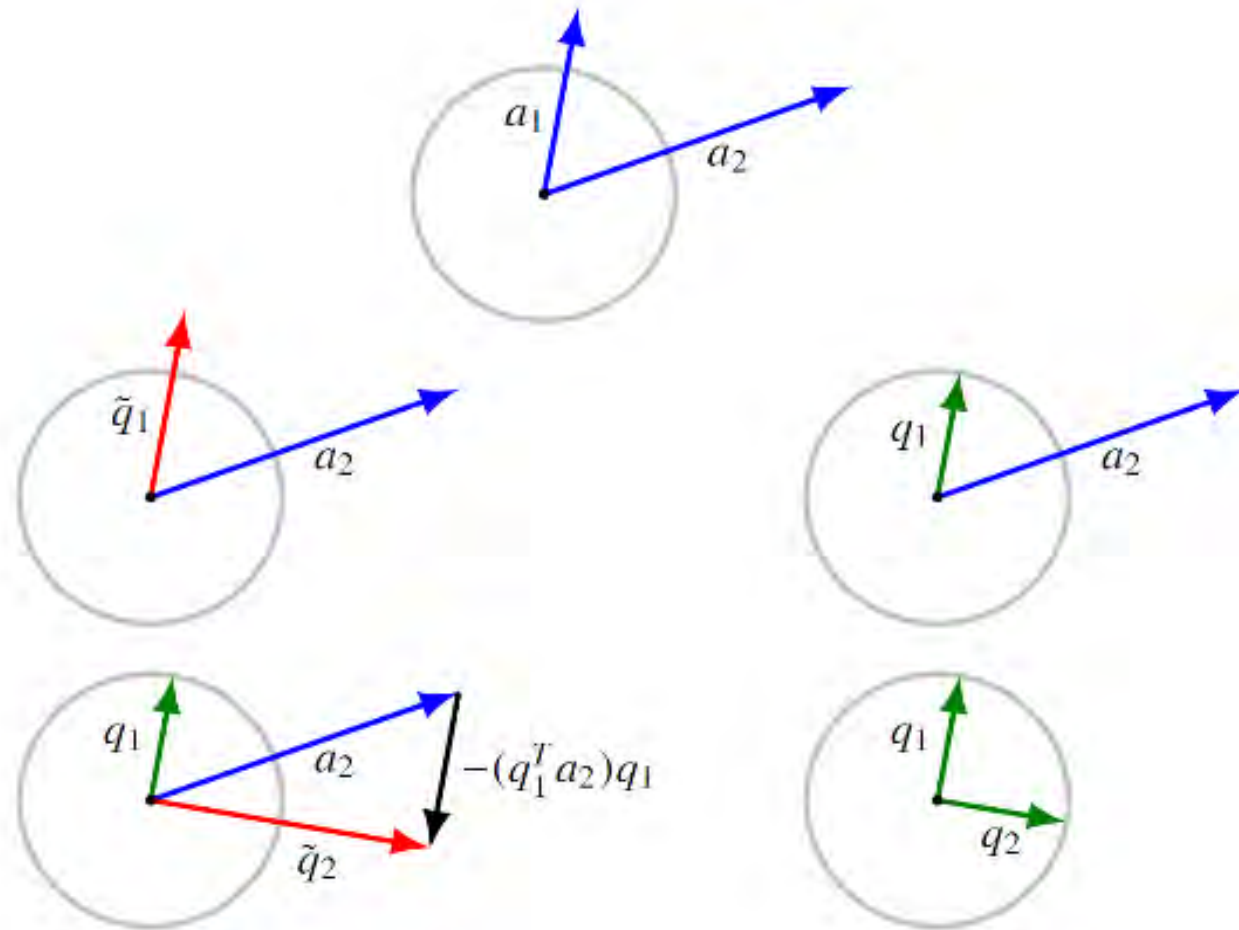


The orthogonal projection of \mathbf{y} onto W is the closest point in W to \mathbf{y} .

Gram–Schmidt (orthogonalization) algorithm



- Find orthonormal basis for $\text{span}\{a_1, a_2, \dots, a_k\}$
- Geometry:





- Find orthonormal basis for $\text{span}\{a_1, a_2, \dots, a_k\}$
- Algebra:

$$1) q_1 = \frac{a_1}{\|a_1\|}$$

$$2) \widetilde{q}_2 = a_2 - (q_1^T a_2)q_1 \rightarrow q_2 = \frac{\widetilde{q}_2}{\|\widetilde{q}_2\|}$$

$$3) \widetilde{q}_3 = a_3 - (q_1^T a_3)q_1 - (q_2^T a_3)q_2 \rightarrow q_3 = \frac{\widetilde{q}_3}{\|\widetilde{q}_3\|}$$

•

•

•

$$k) \widetilde{q}_k = a_k - (q_1^T a_k)q_1 - \dots - (q_{k-1}^T a_k)q_{k-1} \rightarrow q_k = \frac{\widetilde{q}_k}{\|\widetilde{q}_k\|}$$



□ Why $\{q_1, q_2, \dots, q_k\}$ is a orthonormal basis for $\text{span}\{a_1, a_2, \dots, a_k\}$?

- $\{q_1, q_2, \dots, q_k\}$ are normalized.
- $\{q_1, q_2, \dots, q_k\}$ is a orthogonal set
- a_i is a linear combination of $\{q_1, q_2, \dots, q_i\}$



$$\text{span}\{q_1, q_2, \dots, q_k\} = \text{span}\{a_1, a_2, \dots, a_k\}$$

□ q_i is a linear combination of $\{a_1, a_2, \dots, a_i\}$



□ Given n -vectors a_1, \dots, a_k

for $i = 1, \dots, k$

1. Orthogonalization: $\tilde{q}_i = a_i - (q_1^T a_i)q_1 - \dots - (q_{i-1}^T a_i)q_{i-1}$
2. Test for linear dependence: if $\tilde{q}_i = 0$, quit
3. Normalization: $q_i = \frac{\tilde{q}_i}{\|\tilde{q}_i\|}$

Note

- If G-S does not stop early (in step 2), a_1, \dots, a_k are linearly independent.
- If G-S stops early in iteration $i = j$, then a_j is a linear combination of a_1, \dots, a_{j-1} (so a_1, \dots, a_k are linearly dependent)

$$a_j = (q_1^T a_j)q_1 + \dots + (q_{j-1}^T a_j)q_{j-1}$$



□ Gram-Schmidt algorithm gives us an explicit method for determining if a list of vectors is linearly dependent or independent.

□ Given n -vectors a_1, \dots, a_k

for $i = 1, \dots, k$

1. Orthogonalization: $\tilde{q}_i = a_i - (q_1^T a_i)q_1 - \dots - (q_{i-1}^T a_i)q_{i-1}$
2. Test for linear dependence: if $\tilde{q}_i = 0$, quit
3. Normalization: $q_i = \frac{\tilde{q}_i}{\|\tilde{q}_i\|}$

- https://docs.google.com/forms/d/e/1FAIpQLSdy5qeQrADjDvt5ca2Iefi3on-oMtUABeZmxBHzIe34DipiQ/viewform?usp=sf_link
- <https://forms.gle/GBTGgCwsKtN7SFiC6>



□ Given n -vectors a_1, \dots, a_k
for $i = 1, \dots, k$

1. Orthogonalization: $\tilde{q}_i = a_i - (q_1^T a_i)q_1 - \dots - (q_{i-1}^T a_i)q_{i-1}$
2. Test for linear dependence: if $\tilde{q}_i = 0$, quit
3. Normalization: $q_i = \frac{\tilde{q}_i}{\|\tilde{q}_i\|}$



- Determining if a vector is a linear combination of linearly independent vectors.
- Checking if a collection of vectors is a basis.



Theorem

Suppose $B = \{a_1, a_2, \dots, a_n\}$ is a basis of an inner product space A . Then $C = \{q_1, q_2, \dots, q_n\}$ is an orthonormal basis of $\text{span}\{a_1, a_2, \dots, a_n\}$.

$$q_1 = \frac{a_1}{\|a_1\|} \quad q_k = \frac{a_k - \sum_{i=1}^{k-1} \langle q_i, a_k \rangle q_i}{\left\| a_k - \sum_{i=1}^{k-1} \langle q_i, a_k \rangle q_i \right\|} \quad \text{for } 2 \leq k \leq n$$

Proof

We prove this result by induction on k .



Example

Find an orthonormal basis for $P^2[-1, 1]$ with respect to the inner product

$$\langle f, g \rangle = \int_{-1}^1 f(x)g(x)dx$$



Existence of Orthonormal Bases

- Every finite-dimensional inner product space has an orthonormal basis.
- Since finite-dimensional inner product spaces (by definition) have a basis consisting of finitely many vectors, and the Gram-Schmidt process tells us how to convert that basis into an orthonormal basis, we now know that every finite-dimensional inner product space has an orthonormal basis.



- ❑ Chapter 1: Advanced Linear and Matrix Algebra, Nathaniel Johnston
- ❑ Chapter 6: Linear Algebra David Cherney
- ❑ Linear Algebra and Optimization for Machine Learning
- ❑ Introduction to Applied Linear Algebra Vectors, Matrices, and Least Squares