高级工程数学
2021-2022 (1)

沈超敏
计算机科学与技术学院
cmshen@cs.ecnu.edu.cn

教书院 219

Ch. 2. Vector Spaces and Matrices

2.1. Vector and Matrix

A column n-vector: an array of n numbers

$$\boldsymbol{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$$

a_i: the ith component of the vector $\boldsymbol{a}$.

A row n-vector

$$[a_1, a_2, \cdots, a_n].$$

2.1. Vector and Matrix

Transpose: If

$$\boldsymbol{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix},$$

then

$$\boldsymbol{a}^\top = [a_1, \cdots, a_n]$$

2.1. Vector and Matrix

Transpose: If

$$\boldsymbol{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix},$$

then

$$\boldsymbol{a}^\top = [a_1, \cdots, a_n]$$

向量的相等: 两个向量的对应的元素均相等

2.1. Vector and Matrix

The operation of addition of vectors has the following properties:

1. The operation is commutative:

$$\boldsymbol{a} + \boldsymbol{b} = \boldsymbol{b} + \boldsymbol{a}.$$

2. The operation is associative:

$$(\boldsymbol{a} + \boldsymbol{b}) + \boldsymbol{c} = \boldsymbol{a} + (\boldsymbol{b} + \boldsymbol{c}).$$

3. There is a zero vector $\mathbf{0} = [0, 0, ..., 0]^T$ such that

$$\boldsymbol{a} + \mathbf{0} = \mathbf{0} + \boldsymbol{a} = \boldsymbol{a}.$$

2.1. Vector and Matrix

The vector

$$[a_1 - b_1, a_2 - b_2, \cdots, a_n - b_n]^\top$$

is called the difference between $\boldsymbol{a}$ and $\boldsymbol{b}$ and is denoted $\boldsymbol{a} - \boldsymbol{b}$.

The vector $\mathbf{0} - \boldsymbol{b}$ is denoted $-\boldsymbol{b}$.

$$\begin{gathered}
\boldsymbol{b} + (\boldsymbol{a} - \boldsymbol{b}) = \boldsymbol{a} \\
-(-\boldsymbol{b}) = \boldsymbol{b} \\
-(\boldsymbol{a} - \boldsymbol{b}) = \boldsymbol{b} - \boldsymbol{a}
\end{gathered}$$

2.1. Vector and Matrix

数乘: We define an operation of multiplication of a vector $\boldsymbol{a} \in \mathbb{R}^n$ by a real scalar $\alpha \in \mathbb{R}$ as

$$\alpha \boldsymbol{a} = [\alpha a_1, \cdots, \alpha a_n]^\top.$$

1. The operation is distributive: for any real scalars α and β,

$$\alpha(\boldsymbol{a} + \boldsymbol{b}) = \alpha \boldsymbol{a} + \alpha \boldsymbol{b},$$
$$(\alpha + \beta)\boldsymbol{a} = \alpha \boldsymbol{a} + \beta \boldsymbol{a}.$$

2. The operation is associative:

$$\alpha(\beta \boldsymbol{a}) = (\alpha\beta)\boldsymbol{a}.$$

3. The scalar 1 satisfies $1\boldsymbol{a} = \boldsymbol{a}$.

4. Any scalar α satisfies $\alpha \mathbf{0} = \mathbf{0}$.

5. The scalar 0 satisfies $0\boldsymbol{a} = \mathbf{0}$.

6. The scalar —1 satisfies $(-1)\boldsymbol{a} = -\boldsymbol{a}$.

杨振宁教授在清华的《普通物理》

https://www.bilibili.com/video/BV1Fx411T7sV?p=3

(from 第 15 分钟)

2.1. Vector and Matrix

$$\alpha \boldsymbol{a} = \mathbf{0} \quad \iff \quad \alpha = 0 \text{ or } \boldsymbol{a} = \mathbf{0}.$$

2.1. Vector and Matrix

$$\alpha \boldsymbol{a} = \mathbf{0} \quad \Longleftrightarrow \quad \alpha = 0 \text{ or } \boldsymbol{a} = \mathbf{0}.$$

$\alpha \boldsymbol{a} = \mathbf{0}$ is equivalent to $\alpha a_1 = \alpha a_2 = \cdots = \alpha a_n = 0$.

$$\alpha \boldsymbol{a} = \mathbf{0} \quad \Longleftrightarrow \quad \alpha = 0 \text{ or } \boldsymbol{a} = \mathbf{0}.$$

$\alpha \boldsymbol{a} = \mathbf{0}$ is equivalent to $\alpha a_1 = \alpha a_2 = \cdots = \alpha a_n = 0$.

Proof

$\Leftarrow$

$$\alpha \boldsymbol{a} = \mathbf{0} \quad \iff \quad \alpha = 0 \text{ or } \boldsymbol{a} = \mathbf{0}.$$

$\alpha \boldsymbol{a} = \mathbf{0}$ is equivalent to $\alpha a_1 = \alpha a_2 = \cdots = \alpha a_n = 0$.

Proof

$\Leftarrow$

If $\alpha = 0$ or $\boldsymbol{a} = \mathbf{0}$, then $\alpha \boldsymbol{a} = \mathbf{0}$. $\qquad \surd$

2.1. Vector and Matrix

$$\alpha \boldsymbol{a} = \mathbf{0} \quad \Longleftrightarrow \quad \alpha = 0 \text{ or } \boldsymbol{a} = \mathbf{0}.$$

$\alpha \boldsymbol{a} = \mathbf{0}$ is equivalent to $\alpha a_1 = \alpha a_2 = \cdots = \alpha a_n = 0$.

Proof

$\Leftarrow$

If $\alpha = 0$ or $\boldsymbol{a} = \mathbf{0}$, then $\alpha \boldsymbol{a} = \mathbf{0}$. $\qquad \surd$

$\Rightarrow$

$$\alpha \boldsymbol{a} = \mathbf{0} \quad \Longleftrightarrow \quad \alpha = 0 \text{ or } \boldsymbol{a} = \mathbf{0}.$$

$\alpha \boldsymbol{a} = \mathbf{0}$ is equivalent to $\alpha a_1 = \alpha a_2 = \cdots = \alpha a_n = 0$.

Proof

$\Leftarrow$

If $\alpha = 0$ or $\boldsymbol{a} = \mathbf{0}$, then $\alpha \boldsymbol{a} = \mathbf{0}$. $\surd$

$\Rightarrow$

(1) If $\boldsymbol{a} \neq \mathbf{0}$, then at least one of its components $a_k \neq 0$. For this component, $\alpha a_k = 0$, and hence we must have $\alpha = 0$.

2.1. Vector and Matrix

$$\alpha \boldsymbol{a} = \mathbf{0} \quad \Longleftrightarrow \quad \alpha = 0 \text{ or } \boldsymbol{a} = \mathbf{0}.$$

$\alpha \boldsymbol{a} = \mathbf{0}$ is equivalent to $\alpha a_1 = \alpha a_2 = \cdots = \alpha a_n = 0$.

Proof

$\Leftarrow$

If $\alpha = 0$ or $\boldsymbol{a} = \mathbf{0}$, then $\alpha \boldsymbol{a} = \mathbf{0}$. $\quad \surd$

$\Rightarrow$

(1) If $\boldsymbol{a} \neq \mathbf{0}$, then at least one of its components $a_k \neq 0$. For this component, $\alpha a_k = 0$, and hence we must have $\alpha = 0$.

(2) If $\alpha \neq 0$, **类似**

Video 3 结束

2.1. Vector and Matrix

A set of vectors $\{\boldsymbol{a}_1, \cdots, \boldsymbol{a}_k\}$ is linearly independent if

$$\alpha_1 \boldsymbol{a}_1 + \alpha_2 \boldsymbol{a}_2 + \cdots + \alpha_k \boldsymbol{a}_k = 0$$

implies that all coefficients α_i, $i = 1, \cdots, k$, are 0.

2.1. Vector and Matrix

A set of vectors $\{\boldsymbol{a}_1, \cdots, \boldsymbol{a}_k\}$ is linearly independent if

$$\alpha_1 \boldsymbol{a}_1 + \alpha_2 \boldsymbol{a}_2 + \cdots + \alpha_k \boldsymbol{a}_k = 0$$

implies that all coefficients α_i, $i = 1, \cdots, k$, are 0.

线性相关

2.1. Vector and Matrix

A set of vectors $\{\boldsymbol{a}_1, \cdots, \boldsymbol{a}_k\}$ is linearly independent if

$$\alpha_1 \boldsymbol{a}_1 + \alpha_2 \boldsymbol{a}_2 + \cdots + \alpha_k \boldsymbol{a}_k = 0$$

implies that all coefficients α_i, $i = 1, \cdots, k$, are 0.

线性相关

Note: 1) **如** $\{\boldsymbol{a}_1, \cdots, \boldsymbol{a}_k\}$ **中包含** $\mathbf{0}$, **则** $\{\boldsymbol{a}_1, \cdots, \boldsymbol{a}_k\}$ **线性相关**.

2) $\{\mathbf{0}\}$ **线性相关**

2.1. Vector and Matrix

A set of vectors $\{\boldsymbol{a}_1, \cdots, \boldsymbol{a}_k\}$ is linearly independent if

$$\alpha_1 \boldsymbol{a}_1 + \alpha_2 \boldsymbol{a}_2 + \cdots + \alpha_k \boldsymbol{a}_k = 0$$

implies that all coefficients α_i, $i = 1, \cdots, k$, are 0.

线性相关

Note: 1) **如** $\{\boldsymbol{a}_1, \cdots, \boldsymbol{a}_k\}$ **中包含** $\mathbf{0}$, **则** $\{\boldsymbol{a}_1, \cdots, \boldsymbol{a}_k\}$ **线性相关**.

2) $\{\mathbf{0}\}$ **线性相关**

A vector $\boldsymbol{a}$ is said to be a linear combination of vectors $\boldsymbol{a}_1, \boldsymbol{a}_2$, ..., $\boldsymbol{a}_k$ if there are scalars α_1, ..., αa_k such that

$$\boldsymbol{a} = \alpha_1 \boldsymbol{a}_1 + \cdots + \alpha_k \boldsymbol{a}_k.$$

2.1. Vector and Matrix

Proposition 2.1 A set of vectors $\{\boldsymbol{a}_1, ..., \boldsymbol{a}_k\}$ is linearly dependent $\Leftrightarrow$ one of the vectors from the set is a linear combination of the remaining vectors.

2.1. Vector and Matrix

Proposition 2.1 A set of vectors $\{\boldsymbol{a}_1, ..., \boldsymbol{a}_k\}$ is linearly dependent $\Leftrightarrow$ one of the vectors from the set is a linear combination of the remaining vectors.

Proof. $\Rightarrow$ If $\{\boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_k\}$ is linearly dependent, then

$$\alpha_1 \boldsymbol{a}_1 + \alpha_2 \boldsymbol{a}_2 + \cdots + \alpha_k \boldsymbol{a}_k = 0,$$

where at least one of the scalars $\alpha_i \neq 0$, whence

$$\boldsymbol{a}_i = -\frac{\alpha_1}{\alpha_i}\boldsymbol{a}_1 - \cdots - \frac{\alpha_k}{\alpha_i}\boldsymbol{a}_k$$

2.1. Vector and Matrix

Proposition 2.1 A set of vectors $\{\boldsymbol{a}_1, ..., \boldsymbol{a}_k\}$ is linearly dependent $\Leftrightarrow$ one of the vectors from the set is a linear combination of the remaining vectors.

Proof. $\Rightarrow$ If $\{\boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_k\}$ is linearly dependent, then

$$\alpha_1 \boldsymbol{a}_1 + \alpha_2 \boldsymbol{a}_2 + \cdots + \alpha_k \boldsymbol{a}_k = 0,$$

where at least one of the scalars $\alpha_i \neq 0$, whence

$$\boldsymbol{a}_i = -\frac{\alpha_1}{\alpha_i}\boldsymbol{a}_1 - \cdots - \frac{\alpha_k}{\alpha_i}\boldsymbol{a}_k$$

$\Leftarrow$ Suppose that

$$\boldsymbol{a}_1 = \alpha_2 \boldsymbol{a}_2 + \alpha_3 \boldsymbol{a}_3 + \cdots + \alpha_k \boldsymbol{a}_k,$$

i.e.,

$$(-1)\boldsymbol{a}_1 + \alpha_2 \boldsymbol{a}_2 + \alpha_3 \boldsymbol{a}_3 + \cdots + \alpha_k \boldsymbol{a}_k = \boldsymbol{0},$$

i.e., $\{\boldsymbol{a}_1, \cdots, \boldsymbol{a}_k\}$ is linearly independent.

2.1. Vector and Matrix

A subset V of $\mathbb{R}^n$ is called a subspace of $\mathbb{R}^n$ if V is closed under the operations of vector addition and scalar multiplication

2.1. Vector and Matrix

A subset V of $\mathbb{R}^n$ is called a subspace of $\mathbb{R}^n$ if V is closed under the operations of vector addition and scalar multiplication

That is, if $\boldsymbol{a}$ and $\boldsymbol{b}$ are vectors in V, then the vectors $\boldsymbol{a}+\boldsymbol{b}$ and $\alpha\boldsymbol{a}$ are also in V for every scalar α.

2.1. Vector and Matrix

A subset V of $\mathbb{R}^n$ is called a subspace of $\mathbb{R}^n$ if V is closed under the operations of vector addition and scalar multiplication

That is, if $\boldsymbol{a}$ and $\boldsymbol{b}$ are vectors in V, then the vectors $\boldsymbol{a}+\boldsymbol{b}$ and $\alpha\boldsymbol{a}$ are also in V for every scalar α.

Every subspace contains the zero vector $\mathbf{0}$, for if $\boldsymbol{a}$ is an element of the subspace, so is $(-1)\boldsymbol{a} = -\boldsymbol{a}$. Hence, $\boldsymbol{a} - \boldsymbol{a} = \mathbf{0}$ also belongs to the subspace.

2.1. Vector and Matrix

Let $\boldsymbol{a}_1$, ..., $\boldsymbol{a}_k$ be arbitrary vectors in $\mathbb{R}^n$. The set of all their linear combinations is called the span of $\boldsymbol{a}_1$, ..., $\boldsymbol{a}_k$ and is denoted

$$\mathrm{span}[\boldsymbol{a}_1, ..., \boldsymbol{a}_k] = \{\sum_{i=1}^{k} \alpha_i \boldsymbol{a}_i \mid \alpha_1, ..., \alpha_k \in \mathbb{R}\}.$$

2.1. Vector and Matrix

Let $\boldsymbol{a}_1$, ..., $\boldsymbol{a}_k$ be arbitrary vectors in $\mathbb{R}^n$. The set of all their linear combinations is called the span of $\boldsymbol{a}_1$, ..., $\boldsymbol{a}_k$ and is denoted

$$\mathrm{span}[\boldsymbol{a}_1, ..., \boldsymbol{a}_k] = \{\sum_{i=1}^{k} \alpha_i \boldsymbol{a}_i \mid \alpha_1, ..., \alpha_k \in \mathbb{R}\}.$$

Given a vector $\boldsymbol{a}$, the subspace $\mathrm{span}[\boldsymbol{a}]$ is composed of the vectors $\alpha \boldsymbol{a}$, where α is an arbitrary real number ($\alpha \in \mathbb{R}$).

2.1. Vector and Matrix

Let $\boldsymbol{a}_1$, ..., $\boldsymbol{a}_k$ be arbitrary vectors in $\mathbb{R}^n$. The set of all their linear combinations is called the span of $\boldsymbol{a}_1$, ..., $\boldsymbol{a}_k$ and is denoted

$$\mathrm{span}[\boldsymbol{a}_1, ..., \boldsymbol{a}_k] = \{\sum_{i=1}^{k} \alpha_i \boldsymbol{a}_i \mid \alpha_1, ..., \alpha_k \in \mathbb{R}\}.$$

Given a vector $\boldsymbol{a}$, the subspace $\mathrm{span}[\boldsymbol{a}]$ is composed of the vectors $\alpha \boldsymbol{a}$, where α is an arbitrary real number ($\alpha \in \mathbb{R}$).

If $\boldsymbol{a}$ is a linear combination of $\boldsymbol{a}_1$, ..., $\boldsymbol{a}_k$, then

$$\mathrm{span}[\boldsymbol{a}_1, ..., \boldsymbol{a}_k, \boldsymbol{a}] = \mathrm{span}[\boldsymbol{a}_1, ..., \boldsymbol{a}_k].$$

Basis

Given a subspace $\mathcal{V}$, any set of linearly independent vectors $\{\boldsymbol{a}_1, ..., \boldsymbol{a}_k\} \subset \mathcal{V}$ such that $\mathcal{V} = \mathrm{span}[\boldsymbol{a}_1, ..., \boldsymbol{a}_k]$ is referred to as a basis of the subspace $\mathcal{V}$.

Basis

Given a subspace $\mathcal{V}$, any set of linearly independent vectors $\{\boldsymbol{a}_1, ..., \boldsymbol{a}_k\} \subset \mathcal{V}$ such that $\mathcal{V} = \mathrm{span}[\boldsymbol{a}_1, ..., \boldsymbol{a}_k]$ is referred to as a basis of the subspace $\mathcal{V}$.

The dimension of $\mathcal{V}$

All bases of a subspace $\mathcal{V}$ contain the same number of vectors. This number is called the dimension of $\mathcal{V}$, denoted $\dim \mathcal{V}$.

2.1. Vector and Matrix

Proposition 2.2 If $\{\boldsymbol{a}_1, ..., \boldsymbol{a}_k\}$ is a basis of $\mathcal{V}$, then any vector $\boldsymbol{a}$ of $\mathcal{V}$ can be represented uniquely as

$$\boldsymbol{a} = \alpha_1 \boldsymbol{a}_1 + \cdots + \alpha_k \boldsymbol{a}_k,$$

where $\alpha_i \in \mathbb{R}$, $i = 1, 2, ..., k$.

Proposition 2.2 If $\{\boldsymbol{a}_1, ..., \boldsymbol{a}_k\}$ is a basis of $\mathcal{V}$, then any vector $\boldsymbol{a}$ of $\mathcal{V}$ can be represented uniquely as

$$\boldsymbol{a} = \alpha_1 \boldsymbol{a}_1 + \cdots + \alpha_k \boldsymbol{a}_k,$$

where $\alpha_i \in \mathbb{R}$, $i = 1, 2, ..., k$.

Proof. To prove the uniqueness of the representation of $\boldsymbol{a}$ in terms of the basis vectors, assume that

$$\boldsymbol{a} = \alpha_1 \boldsymbol{a}_1 + \cdots + \alpha_k \boldsymbol{a}_k,$$

and

$$\boldsymbol{a} = \beta_1 \boldsymbol{a}_1 + \cdots + \beta_k \boldsymbol{a}_k,$$

Then

$$(\alpha_1 - \beta_1)\boldsymbol{a}_1 + \cdots + (\alpha_k - \beta_k)\boldsymbol{a}_k = 0.$$

Because $\{\boldsymbol{a}_1, \cdots, \boldsymbol{a}_k\}$ is linearly independent, $\alpha_i = \beta_i$. □

2.1. Vector and Matrix

Suppose that we are given a basis $\{\boldsymbol{a}_1, ..., \boldsymbol{a}_k\}$ of $\mathcal{V}$ and a vector $\boldsymbol{a} \in \mathcal{V}$ such that

$$\boldsymbol{a} = \alpha_1 \boldsymbol{a}_1 + \cdots + \alpha_k \boldsymbol{a}_k,$$

The coefficients α_i, $i = 1, ..., k$, are called the coordinates of $\boldsymbol{a}$ with respect to the basis $\{\boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_k\}$.

2.1. Vector and Matrix

Suppose that we are given a basis $\{\boldsymbol{a}_1, ..., \boldsymbol{a}_k\}$ of $\mathcal{V}$ and a vector $\boldsymbol{a} \in \mathcal{V}$ such that

$$\boldsymbol{a} = \alpha_1 \boldsymbol{a}_1 + \cdots + \alpha_k \boldsymbol{a}_k,$$

The coefficients α_i, $i = 1, ..., k$, are called the coordinates of $\boldsymbol{a}$ with respect to the basis $\{\boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_k\}$.

The natural basis for $\mathbb{R}^n$ is the set of vectors

$$\boldsymbol{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \boldsymbol{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \cdots, \boldsymbol{e}_n = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}.$$

2.1. Vector and Matrix

Matrix

A matrix is a rectangular array of numbers, commonly denoted by uppercase bold letters (e.g., A). A matrix with m rows and n columns is called an $m \times n$ matrix, and we write

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

2.1. Vector and Matrix

Matrix

A matrix is a rectangular array of numbers, commonly denoted by uppercase bold letters (e.g., A). A matrix with m rows and n columns is called an $m \times n$ matrix, and we write

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

Transpose

$$A^\top = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{bmatrix}$$

Video 4 结束

2.2. Rank of a Matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

Let us denote the k-th column of A by $\boldsymbol{a}_k$, then $A = [\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n]$.

2.2. Rank of a Matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

Let us denote the k-th column of A by $\boldsymbol{a}_k$, then $A = [\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n]$.

The maximal number of linearly independent columns of A is called the rank of the matrix A, denoted $\operatorname{rank} A$. Note that $\operatorname{rank} A$ is the dimension of $\operatorname{span}[\boldsymbol{a}_1, ..., \boldsymbol{a}_n]$.

2.2. Rank of a Matrix

Proposition 2.3. The rank of a matrix A is invariant under the following operations:

1. Multiplication of the columns of A by nonzero scalars.

2. Interchange of the columns.

3. Addition to a given column a linear combination of other columns.

2.2. Rank of a Matrix

Proposition 2.3. The rank of a matrix A is invariant under the following operations:

1. Multiplication of the columns of A by nonzero scalars.

2. Interchange of the columns.

3. Addition to a given column a linear combination of other columns.

Proof.

1. Let $\boldsymbol{b}_k = \alpha_k \boldsymbol{a}_k$, where $\alpha_k \neq 0$, $k = 1, \cdots, n$, and let $B = [\boldsymbol{b}_1, \cdots, \boldsymbol{b}_n]$. Obviously,

$$\operatorname{span}[\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n] = \operatorname{span}[\boldsymbol{b}_1, \cdots, \boldsymbol{b}_n],$$

and thus

$$\operatorname{rank} A = \operatorname{rank} B.$$

2.2. Rank of a Matrix

Proposition 2.3. The rank of a matrix A is invariant under the following operations:

1. Multiplication of the columns of A by nonzero scalars.
2. Interchange of the columns.
3. Addition to a given column a linear combination of other columns.

2.2. Rank of a Matrix

Proposition 2.3. The rank of a matrix A is invariant under the following operations:

1. Multiplication of the columns of A by nonzero scalars.

2. Interchange of the columns.

3. Addition to a given column a linear combination of other columns.

Proof.
2. The number of linearly independent vectors does not depend on their order.

2.2. Rank of a Matrix

3. Addition to a given column a linear combination of other columns.

2.2. Rank of a Matrix

3. Addition to a given column a linear combination of other columns.

Proof of 3. Let

$$
\begin{aligned}
\boldsymbol{b}_1 &= \boldsymbol{a}_1 + c_2\boldsymbol{a}_2 + \cdots + c_n\boldsymbol{a}_n, \\
\boldsymbol{b}_2 &= \boldsymbol{a}_2 \\
&\vdots \\
\boldsymbol{b}_n &= \boldsymbol{a}_n
\end{aligned}
$$

So, $\alpha_1\boldsymbol{b}_1 + \alpha_2\boldsymbol{b}_2 + \cdots + \alpha_n\boldsymbol{b}_n = \alpha_1\boldsymbol{a}_1 + (\alpha_2 + \alpha_1 c_2)\boldsymbol{a}_2 + \cdots + (\alpha_n + \alpha_1 c_n)\boldsymbol{a}_n,.$
hence $\mathrm{span}[\boldsymbol{b}_1, \cdots, \boldsymbol{b}_n] \subset \mathrm{span}[\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n]$.

2.2. Rank of a Matrix

3. Addition to a given column a linear combination of other columns.

Proof of 3. Let

$$\begin{aligned} \boldsymbol{b}_1 &= \boldsymbol{a}_1 + c_2\boldsymbol{a}_2 + \cdots + c_n\boldsymbol{a}_n, \\ \boldsymbol{b}_2 &= \boldsymbol{a}_2 \\ &\vdots \\ \boldsymbol{b}_n &= \boldsymbol{a}_n \end{aligned}$$

So, $\alpha_1\boldsymbol{b}_1 + \alpha_2\boldsymbol{b}_2 + \cdots + \alpha_n\boldsymbol{b}_n = \alpha_1\boldsymbol{a}_1 + (\alpha_2 + \alpha_1 c_2)\boldsymbol{a}_2 + \cdots + (\alpha_n + \alpha_1 c_n)\boldsymbol{a}_n,$.
hence $\operatorname{span}[\boldsymbol{b}_1, \cdots, \boldsymbol{b}_n] \subset \operatorname{span}[\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n]$.
On the other hand

$$\begin{aligned} \boldsymbol{a}_1 &= \boldsymbol{b}_1 - c_2\boldsymbol{b}_2 - \cdots - c_n\boldsymbol{b}_n, \\ \boldsymbol{a}_2 &= \boldsymbol{b}_2 \\ &\vdots \\ \boldsymbol{a}_n &= \boldsymbol{b}_n \end{aligned}$$

hence $\operatorname{span}[\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n] \subset \operatorname{span}[\boldsymbol{b}_1, \cdots, \boldsymbol{b}_n]$.
Therefore, $\operatorname{rank} A = \operatorname{rank} B$.

Video 5 结束

2.2. Rank of a Matrix

对方阵 A, 可定义 determinant, 记为 $\det A$ or $|A|$.

2.2. Rank of a Matrix

对方阵 A, **可定义** determinant, **记为** $\det A$ or $|A|$.

Determinant **具有如下性质**

1. The determinant of the matrix $A = [\boldsymbol{a}_1, ..., \boldsymbol{a}_n]$ is a linear function of each column; that is,

$$
\begin{aligned}
&\det[\boldsymbol{a}_1, \cdots, \boldsymbol{a}_{k-1}, \alpha\boldsymbol{a}_k^{(1)} + \beta\boldsymbol{a}_k^{(2)}, \boldsymbol{a}_{k+1}, \cdots, \boldsymbol{a}_n] \\
&= \alpha \det[\boldsymbol{a}_1, \cdots, \boldsymbol{a}_k^{(1)}, \boldsymbol{a}_{k+1}, \cdots, \boldsymbol{a}_n] + \beta \det[\boldsymbol{a}_1, \cdots, \boldsymbol{a}_k^{(2)}, \boldsymbol{a}_{k+1}, \cdots, \boldsymbol{a}_n]
\end{aligned}
$$

2.2. Rank of a Matrix

对方阵 A, 可定义 determinant, 记为 $\det A$ or $|A|$.

Determinant 具有如下性质

1. The determinant of the matrix $A = [\boldsymbol{a}_1, ..., \boldsymbol{a}_n]$ is a linear function of each column; that is,

$$\begin{aligned}&\det[\boldsymbol{a}_1, \cdots, \boldsymbol{a}_{k-1}, \alpha\boldsymbol{a}_k^{(1)} + \beta\boldsymbol{a}_k^{(2)}, \boldsymbol{a}_{k+1}, \cdots, \boldsymbol{a}_n] \\ &= \alpha\det[\boldsymbol{a}_1, \cdots, \boldsymbol{a}_k^{(1)}, \boldsymbol{a}_{k+1}, \cdots, \boldsymbol{a}_n] + \beta\det[\boldsymbol{a}_1, \cdots, \boldsymbol{a}_k^{(2)}, \boldsymbol{a}_{k+1}, \cdots, \boldsymbol{a}_n]\end{aligned}$$

2. If for some k we have $\boldsymbol{a}_k = \boldsymbol{a}_{k+1}$, then

$$\det A = \det[\boldsymbol{a}_1, \cdots, \boldsymbol{a}_k, \boldsymbol{a}_{k+1}, \cdots, \boldsymbol{a}_n] = 0$$

2.2. Rank of a Matrix

对方阵 A, 可定义 determinant, 记为 $\det A$ or $|A|$.

Determinant 具有如下性质

1. The determinant of the matrix $A = [\boldsymbol{a}_1, ..., \boldsymbol{a}_n]$ is a linear function of each column; that is,

$$\begin{aligned}&\det[\boldsymbol{a}_1, \cdots, \boldsymbol{a}_{k-1}, \alpha\boldsymbol{a}_k^{(1)} + \beta\boldsymbol{a}_k^{(2)}, \boldsymbol{a}_{k+1}, \cdots, \boldsymbol{a}_n] \\ &= \alpha \det[\boldsymbol{a}_1, \cdots, \boldsymbol{a}_k^{(1)}, \boldsymbol{a}_{k+1}, \cdots, \boldsymbol{a}_n] + \beta \det[\boldsymbol{a}_1, \cdots, \boldsymbol{a}_k^{(2)}, \boldsymbol{a}_{k+1}, \cdots, \boldsymbol{a}_n]\end{aligned}$$

2. If for some k we have $\boldsymbol{a}_k = \boldsymbol{a}_{k+1}$, then

$$\det A = \det[\boldsymbol{a}_1, \cdots, \boldsymbol{a}_k, \boldsymbol{a}_{k+1}, \cdots, \boldsymbol{a}_n] = 0$$

3. Let

$$I_n = [\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n] = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

Then

$$\det I_n = 1$$

其它性质:

$$\begin{aligned}
&\det[\boldsymbol{a}_1, \cdots, \boldsymbol{a}_{k-1}, \boldsymbol{a}_k + \alpha\boldsymbol{a}_j, \boldsymbol{a}_{k+1}, \cdots, \boldsymbol{a}_j, \cdots, \boldsymbol{a}_n] \\
&= \det[\boldsymbol{a}_1, \cdots, \boldsymbol{a}_k, \boldsymbol{a}_{k+1}, \cdots, \boldsymbol{a}_n] + \alpha \det[\boldsymbol{a}_1, \cdots, \boldsymbol{a}_j, \cdots, \boldsymbol{a}_j, \cdots, \boldsymbol{a}_n] \\
&= \det[\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n]
\end{aligned}$$

Video 6 结束

2.2. Rank of a Matrix

A pth-order minor of an $m \times n$ matrix A, with $p < \min\{m, n\}$, is the determinant of a $p \times p$ matrix obtained from A by deleting $m - p$ rows and $n - p$ columns.

2.2. Rank of a Matrix

Proposition 2.4 If an $m \times n$ $(m > n)$ matrix A has a nonzero n-th-order minor, then the columns of A are linearly independent; that is, $\operatorname{rank} A = n$.

2.2. Rank of a Matrix

Proposition 2.4 If an $m \times n$ $(m > n)$ matrix A has a nonzero n-th-order minor, then the columns of A are linearly independent; that is, $\operatorname{rank} A = n$.

Proof. Suppose that A has a nonzero nth-order minor. Without loss of generality, we assume that the nth-order minor corresponding to the first n rows of A is nonzero. Let x_i, $i = 1, ..., n$, be scalars such that

$$x_1 \boldsymbol{a}_1 + \cdots + x_n \boldsymbol{a}_n = 0.$$

2.2. Rank of a Matrix

Proposition 2.4 If an $m \times n$ $(m > n)$ matrix A has a nonzero n-th-order minor, then the columns of A are linearly independent; that is, $\operatorname{rank} A = n$.

Proof. Suppose that A has a nonzero nth-order minor. Without loss of generality, we assume that the nth-order minor corresponding to the first n rows of A is nonzero. Let x_i, $i = 1, ..., n$, be scalars such that

$$x_1 \boldsymbol{a}_1 + \cdots + x_n \boldsymbol{a}_n = 0.$$

i.e.,

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0 \\ &\vdots \\ a_{n1}x_1 + a_{22}x_2 + \cdots + a_{nn}x_n &= 0 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= 0 \end{aligned}$$

2.2. Rank of a Matrix

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0 \\
&\vdots \\
a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= 0 \\
&\vdots \\
a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= 0
\end{aligned}
$$

记 $\tilde{\boldsymbol{a}}_i = \begin{bmatrix} a_{1i} \\ \vdots \\ a_{ni} \end{bmatrix}$, then $x_1\tilde{\boldsymbol{a}}_1 + \cdots + x_n\tilde{\boldsymbol{a}}_n = \mathbf{0}$.

The nth-order minor is $\det[\tilde{\boldsymbol{a}}_1, , ..., \tilde{\boldsymbol{a}}_n]$, assumed to be nonzero. From the properties of determinants it follows that the columns $\tilde{\boldsymbol{a}}_1, \cdots, \tilde{\boldsymbol{a}}_n$ are linearly independent. Therefore, all $x_i = 0$, $i = 1, ..., n$. Hence, the columns $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n$ an are linearly independent. $\square$

非异阵

对方阵 $A_{n\times n}$, A **非异** iff $\exists B_{n\times n}$, such that

$$AB = BA = I_n.$$

B **称为** A **的逆阵, 记为** $B = A^{-1}$.

Video 7 结束

2.3. Linear Equations

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\
&\vdots \\
a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m
\end{aligned}
$$

2.3. Linear Equations

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\
&\vdots \\
a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m
\end{aligned}
$$

i.e.,

$$x_1\boldsymbol{a}_1 + \cdots + x_n\boldsymbol{a}_n = \boldsymbol{b}.$$

i.e.,

$$A\boldsymbol{x} = \boldsymbol{b}$$

2.3. Linear Equations

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

i.e.,

$$x_1\boldsymbol{a}_1 + \cdots + x_n\boldsymbol{a}_n = \boldsymbol{b}.$$

i.e.,

$$A\boldsymbol{x} = \boldsymbol{b}$$

增广矩阵 $[A, \boldsymbol{b}] = [\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n, \boldsymbol{b}]$

2.3. Linear Equations

Theorem 2.1 The system of equations $A\boldsymbol{x} = \boldsymbol{b}$ has a solution if and only if

$$\operatorname{rank} A = \operatorname{rank}[A, \boldsymbol{b}].$$

2.3. Linear Equations

Theorem 2.1 The system of equations $A\boldsymbol{x} = \boldsymbol{b}$ has a solution if and only if

$$\operatorname{rank} A = \operatorname{rank}[A, \boldsymbol{b}].$$

Proof. $\Rightarrow$ Suppose that the system $A\boldsymbol{x} = \boldsymbol{b}$ has a solution. Therefore, $\boldsymbol{b}$ is a linear combination of the columns of A; that is, there exist $x_1, ..., x_n$ such that $x_1\boldsymbol{a}_1 + \cdots + x_n\boldsymbol{a}_n = \boldsymbol{b}$, 即 $\boldsymbol{b} \in \operatorname{span}[\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n]$. 因此

$$\begin{aligned}
\operatorname{rank} A &= \dim \operatorname{span}[\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n] \\
&= \dim \operatorname{span}[\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n, \boldsymbol{b}] \\
&= \operatorname{rank}[A, \boldsymbol{b}]
\end{aligned}$$

$\Leftarrow$ Suppose that $\operatorname{rank} A = \operatorname{rank}[A, \boldsymbol{b}] = r$. Thus, we have r linearly independent columns of A. Without loss of generality, let $\boldsymbol{a}_1, ..., \boldsymbol{a}_r$ be these columns. Therefore, $\boldsymbol{a}_1, ..., \boldsymbol{a}_r$ are also linearly independent columns of the matrix $[A, \boldsymbol{b}]$. Because $\operatorname{rank}[A, \boldsymbol{b}] = r$, the remaining columns of $[A, \boldsymbol{b}]$ can be expressed as linear combinations of $\boldsymbol{a}_1, ..., \boldsymbol{a}_r$. In particular, $\boldsymbol{b}$ can be expressed as a linear combination of these columns. Hence, there exist $x_1, ..., x_n$ such that $x_1\boldsymbol{a}_1 + \cdots + x_n\boldsymbol{a}_n = \boldsymbol{b}$. □

2.3. Linear Equations

Theorem 2.2 Consider the equation $A\boldsymbol{x} = \boldsymbol{b}$, where $A \in \mathbb{R}^{m \times n}$ and $\operatorname{rank} A = m$. A solution to $A\boldsymbol{x} = \boldsymbol{b}$ can be obtained by assigning arbitrary values for $n - m$ variables and solving for the remaining ones.

2.3. Linear Equations

Theorem 2.2 Consider the equation $A\boldsymbol{x} = \boldsymbol{b}$, where $A \in \mathbb{R}^{m\times n}$ and $\operatorname{rank} A = m$. A solution to $A\boldsymbol{x} = \boldsymbol{b}$ can be obtained by assigning arbitrary values for $n - m$ variables and solving for the remaining ones.

Proof. We have $\operatorname{rank} A = m$, and therefore we can find m linearly independent columns of A. Without loss of generality, let $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m$ be such columns. Rewrite the equation $A\boldsymbol{x} = \boldsymbol{b}$ as

$$x_1\boldsymbol{a}_1 + x_m\boldsymbol{a}_m = \boldsymbol{b} - x_{m+1}\boldsymbol{a}_{m+1} - \cdots - x_n\boldsymbol{a}_n.$$

Assign to x_{m+1}, x_{m+2}, ..., x_n arbitrary values, say

$$x_{m+1} = d_{m+1}, x_{m+2} = d_{m+2}, \cdots, x_n = d_m,$$

and let

$$B = [\boldsymbol{a}_1, \cdots, \boldsymbol{a}_m] \in \mathbb{R}^{m\times m}.$$

Note that $\det B \neq 0$. We can represent the system of equations as

$$B \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} = [\boldsymbol{b} - d_{m+1}\boldsymbol{a}_{m+1} - \cdots - d_n\boldsymbol{a}_n].$$

因为 B 可逆, 所以

$$\begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} = B^{-1}[\boldsymbol{b} - d_{m+1}\boldsymbol{a}_{m+1} - \cdots - d_n\boldsymbol{a}_n].$$

$\square$

Video 8 结束

解释下列现象

例如二维空间中的 (1,2) 和 (2,1)，在换了一组基 (1,1), (0,1) 的情况下，它们的坐标分别变成了 (1,1) 和 (2,-1)，此时内积从 4 变成了 1。

这个过程哪里出问题了？

问题 2

$$\begin{cases} x + y = 1 \\ x - y = 1 \\ x + 2y = 3 \end{cases}$$

2.4 Inner Products and Norms

对于绝对值, 有以下的公式:

1. $|a| = |-a|$.

2. $-|a| \leq a < |a|$.

3. $|a+b| < |a| + |b|$.

4. $||a| - |b|| < |a-b| < |a| + |b|$.

5. $|ab| = |a|\,|b|$.

6. $|a| < c$ and $|b| < d$ imply that $|a+b| < c+d$.

7. The inequality $|a| < b$ is equivalent to $-b < a < b$ (i.e., $a < b$ and $-a < b$). The same holds if we replace every occurrence of "$<$" by "$\leq$"

8. The inequality $|a| > b$ is equivalent to $a > b$ or $-a > b$. The same holds if we replace every occurrence of "$>$" by "$>$"

2.4 Inner Products and Norms

For $\boldsymbol{x}$, $\boldsymbol{y} \in \mathbb{R}^n$, we define the Euclidean inner product by

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \sum_{i=1}^{n} x_i y_i = \boldsymbol{x}^\top \boldsymbol{y}.$$

内积具有以下性质:

1. Positivity: $\langle \boldsymbol{x}, \boldsymbol{x} \rangle > 0$, $\langle \boldsymbol{x}, \boldsymbol{x} \rangle = 0$ iff $\boldsymbol{x} = 0$.
2. Symmetry: $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{y}, \boldsymbol{x} \rangle$.
3. Additivity: $\langle \boldsymbol{x} + \boldsymbol{y}, \boldsymbol{z} \rangle = \langle \boldsymbol{x}, \boldsymbol{z} \rangle + \langle \boldsymbol{y}, \boldsymbol{z} \rangle$.
4. Homogeneity (**齐性**): $\langle r\boldsymbol{x}, \boldsymbol{y} \rangle = r\langle \boldsymbol{x}, \boldsymbol{y} \rangle$ for every $r \in \mathbb{R}$.

The properties of additivity and homogeneity in the second vector also hold;

$$\langle \boldsymbol{x}, \boldsymbol{y} + \boldsymbol{z} \rangle = \langle \boldsymbol{x}, \boldsymbol{y} \rangle + \langle \boldsymbol{x}, \boldsymbol{z} \rangle,$$

$$\langle \boldsymbol{x}, r\boldsymbol{y} \rangle = r\langle \boldsymbol{x}, \boldsymbol{y} \rangle \text{ for every } r \in \mathbb{R}.$$

2.4 Inner Products and Norms

The vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ are said to be orthogonal if $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = 0$.

The Euclidean norm of a vector $\boldsymbol{x}$ is defined as

$$\|\boldsymbol{x}\| = \sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle} = \sqrt{\boldsymbol{x}^\top \boldsymbol{x}}$$

2.4 Inner Products and Norms

The vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ are said to be orthogonal if $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = 0$.

The Euclidean norm of a vector $\boldsymbol{x}$ is defined as

$$\|\boldsymbol{x}\| = \sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle} = \sqrt{\boldsymbol{x}^\top \boldsymbol{x}}$$

Theorem 2.3 Cauchy-Schwarz Inequality. For any two vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ in $\mathbb{R}^n$, the Cauchy-Schwarz inequality

$$|\langle \boldsymbol{x}, \boldsymbol{y} \rangle| \leq \|\boldsymbol{x}\| \cdot \|\boldsymbol{y}\|$$

holds. Furthermore, equality holds iff $\boldsymbol{x} = \alpha \boldsymbol{y}$ for some $\alpha \in \mathbb{R}$.

Proof. First assume that $\boldsymbol{x}$ and $\boldsymbol{y}$ are unit vectors; that is, $\|\boldsymbol{x}\| = \|\boldsymbol{y}\| = 1$. Then,

$$\begin{aligned} 0 \leq \|\boldsymbol{x} - \boldsymbol{y}\|^2 &= \langle \boldsymbol{x} - \boldsymbol{y}, \boldsymbol{x} - \boldsymbol{y} \rangle \\ &= \|\boldsymbol{x}\|^2 - 2\langle \boldsymbol{x}, \boldsymbol{y} \rangle + \|\boldsymbol{y}\|^2 \\ &= 2 - 2\langle \boldsymbol{x}, \boldsymbol{y} \rangle \end{aligned}$$

or

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle \leq 1,$$

等号成立当且仅当 $\boldsymbol{x} = \boldsymbol{y}$.

Next, assuming that neither $\boldsymbol{x}$ nor $\boldsymbol{y}$ is zero (for the inequality obviously holds if one of them is zero), we replace $\boldsymbol{x}$ and $\boldsymbol{y}$ by the unit vectors $\boldsymbol{x}/\|\boldsymbol{x}\|$ and $\boldsymbol{y}/\|\boldsymbol{y}\|$. Then

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

Replacing $\boldsymbol{x}$ by $-\boldsymbol{x}$, we have

$$-\langle \boldsymbol{x}, \boldsymbol{y} \rangle \leq \|\boldsymbol{x}\| \cdot \|\boldsymbol{y}\|.$$

所以,

$$|\langle \boldsymbol{x}, \boldsymbol{y} \rangle| \leq \|\boldsymbol{x}\| \cdot \|\boldsymbol{y}\|$$

等号成立当且仅当 $\boldsymbol{x}/\|\boldsymbol{x}\| = \pm \boldsymbol{y}/\|\boldsymbol{y}\|$; that is, $\boldsymbol{x} = \alpha \boldsymbol{y}$ for some $\alpha \in \mathbb{R}$.

2.4 Inner Products and Norms

The Euclidean norm of a vector $\|\boldsymbol{x}\|$ has the following properties:

1. Positivity: $\|\boldsymbol{x}\| > 0$, $\|\boldsymbol{x}\| = \mathbf{0}$ iff $\boldsymbol{x} = \mathbf{0}$.
2. Homogeneity: $\|r\boldsymbol{x}\| = |r| \cdot \|\boldsymbol{x}\|$, $r \in \mathbb{R}$.
3. Triangle inequality: $\|\boldsymbol{x} + \boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$.

2.4 Inner Products and Norms

The Euclidean norm of a vector $\|\boldsymbol{x}\|$ has the following properties:

1. Positivity: $\|\boldsymbol{x}\| > 0$, $\|\boldsymbol{x}\| = \mathbf{0}$ iff $\boldsymbol{x} = \mathbf{0}$.
2. Homogeneity: $\|r\boldsymbol{x}\| = |r| \cdot \|\boldsymbol{x}\|$, $r \in \mathbb{R}$.
3. Triangle inequality: $\|\boldsymbol{x} + \boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$.

三角不等式可用 Cauchy-Schwarz 不等式证明:

$$
\begin{aligned}
\|\boldsymbol{x} + \boldsymbol{y}\|^2 \quad & \overset{def}{=\!=} \quad \|\boldsymbol{x}\|^2 + 2\langle \boldsymbol{x}, \boldsymbol{y} \rangle + \|\boldsymbol{y}\|^2 \\
& \overset{C-S}{\leq} \quad \|\boldsymbol{x}\|^2 + 2\|\boldsymbol{x}\|\|\boldsymbol{y}\| + \|\boldsymbol{y}\|^2 \\
& = \quad (\|\boldsymbol{x}\| + \|\boldsymbol{y}\|)^2
\end{aligned}
$$

2.4 Inner Products and Norms

The Euclidean norm of a vector $\|\boldsymbol{x}\|$ has the following properties:

1. Positivity: $\|\boldsymbol{x}\| > 0$, $\|\boldsymbol{x}\| = \mathbf{0}$ iff $\boldsymbol{x} = \mathbf{0}$.

2. Homogeneity: $\|r\boldsymbol{x}\| = |r| \cdot \|\boldsymbol{x}\|$, $r \in \mathbb{R}$.

3. Triangle inequality: $\|\boldsymbol{x} + \boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$.

三角不等式可用 Cauchy-Schwarz 不等式证明:

$$
\begin{aligned}
\|\boldsymbol{x} + \boldsymbol{y}\|^2 &\overset{def}{=\!=} \|\boldsymbol{x}\|^2 + 2\langle \boldsymbol{x}, \boldsymbol{y} \rangle + \|\boldsymbol{y}\|^2 \\
&\overset{C-S}{\leq} \|\boldsymbol{x}\|^2 + 2\|\boldsymbol{x}\|\|\boldsymbol{y}\| + \|\boldsymbol{y}\|^2 \\
&= (\|\boldsymbol{x}\| + \|\boldsymbol{y}\|)^2
\end{aligned}
$$

If $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = 0$, then $\|\boldsymbol{x} + \boldsymbol{y}\|^2 = \|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2$.

2.4 Inner Products and Norms

p-norm

刚才的 Euclidean norm 是以下形式的 norm 取 $p = 2$ 的情形

$$\|\boldsymbol{x}\|_p = \begin{cases} (|x_1|^p + \cdots + |x_n|^p)^{1/p} & \text{if } 1 \leq p < \infty \\ \max\{|x_1|, ..., |x_n|\} & \text{if } p = \infty \end{cases}$$

2.4 Inner Products and Norms

p-norm

刚才的 Euclidean norm 是以下形式的 norm 取 $p = 2$ 的情形

$$\|\boldsymbol{x}\|_p = \begin{cases} (|x_1|^p + \cdots + |x_n|^p)^{1/p} & \text{if } 1 \leq p < \infty \\ \max\{|x_1|, ..., |x_n|\} & \text{if } p = \infty \end{cases}$$

Norm 可以用来描述连续函数.

A function $\boldsymbol{f} : \mathbb{R}^n \to \mathbb{R}^m$ is continuous at $\boldsymbol{x}$ if for all $\varepsilon > 0$, there exists $\delta > 0$ such that $\|\boldsymbol{y} - \boldsymbol{x}\| < \delta \Rightarrow \|\boldsymbol{f}(\boldsymbol{y}) - \boldsymbol{f}(\boldsymbol{x})\| < \varepsilon$.

If the function $\boldsymbol{f}$ is continuous at every point in $\mathbb{R}^n$, we say that it is continuous on $\mathbb{R}^n$. Note that $\boldsymbol{f} = [f_1, ..., f_m]^\top$ is continuous iff each component f_i, $i = 1, ..., m$, is continuous

2.4 Inner Products and Norms

For the complex vector space $\mathbb{C}^n$, we define an inner product $\langle \boldsymbol{x}, \boldsymbol{y} \rangle$ to be $\sum_{i=1}^{n} x_i \bar{y}_i$.

The inner product on $\mathbb{C}^n$ is a complex-valued function having the following properties:

1. Positivity: $\langle \boldsymbol{x}, \boldsymbol{x} \rangle > 0$, $\langle \boldsymbol{x}, \boldsymbol{x} \rangle = 0$ iff $\boldsymbol{x} = 0$.
2. Symmetry: $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \overline{\langle \boldsymbol{y}, \boldsymbol{x} \rangle}$.
3. Additivity: $\langle \boldsymbol{x} + \boldsymbol{y}, \boldsymbol{z} \rangle = \langle \boldsymbol{x}, \boldsymbol{z} \rangle + \langle \boldsymbol{y}, \boldsymbol{z} \rangle$.
4. Homogeneity (齐性): $\langle r\boldsymbol{x}, \boldsymbol{y} \rangle = r\langle \boldsymbol{x}, \boldsymbol{y} \rangle$ for every $r \in \mathbb{C}$.

Also,

$$\langle \boldsymbol{x}, r_1\boldsymbol{y} + r_2\boldsymbol{z} \rangle = \bar{r}_1\langle \boldsymbol{x}, \boldsymbol{y} \rangle + \bar{r}_2\langle \boldsymbol{x}, \boldsymbol{z} \rangle.$$

Video 9 结束

Transformations

3.1 Linear Transformations

A function $\mathcal{L} : \mathbb{R}^n \to \mathbb{R}^m$ is called a linear transformation if:

1. $\mathcal{L}(a\boldsymbol{x}) = a\mathcal{L}(\boldsymbol{x})$ for every $\boldsymbol{x} \in \mathbb{R}^n$ and $a \in \mathbb{R}$.
2. $\mathcal{L}(\boldsymbol{x}_1 + \boldsymbol{x}_2) = \mathcal{L}(\boldsymbol{x}_1) + \mathcal{L}(\boldsymbol{x}_2)$ for $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{R}^n$.

3.1 Linear Transformations

A function $\mathcal{L} : \mathbb{R}^n \to \mathbb{R}^m$ is called a linear transformation if:

1. $\mathcal{L}(a\boldsymbol{x}) = a\mathcal{L}(\boldsymbol{x})$ for every $\boldsymbol{x} \in \mathbb{R}^n$ and $a \in \mathbb{R}$.
2. $\mathcal{L}(\boldsymbol{x}_1 + \boldsymbol{x}_2) = \mathcal{L}(\boldsymbol{x}_1) + \mathcal{L}(\boldsymbol{x}_2)$ for $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{R}^n$.

If we fix the bases for $\mathbb{R}^n$ and $\mathbb{R}^m$, then the linear transformation $\mathcal{L}$ can be represented by a matrix.

Suppose that $\boldsymbol{x} \in \mathbb{R}^n$ is a given vector, and $\boldsymbol{x}'$ is the representation of $\boldsymbol{x}$ with respect to the given basis for $\mathbb{R}^n$. If $\boldsymbol{y} = \mathcal{L}(\boldsymbol{x})$, and $\boldsymbol{y}'$ is the representation of $\boldsymbol{y}$ with respect to the given basis for $\mathbb{R}^m$. If there exists $A \in \mathbb{R}^{m \times n}$ such that the following representation holds, then

$$\boldsymbol{y}' = A\boldsymbol{x}'.$$

We call A the matrix representation of $\mathcal{L}$ with respect to the given bases for $\mathbb{R}^n$ and $\mathbb{R}^m$. In the special case where we assume the natural bases for $\mathbb{R}^n$ and $\mathbb{R}^m$, the matrix representation A satisfies

$$\mathcal{L}(\boldsymbol{x}) = A\boldsymbol{x}.$$

3.1 Linear Transformations

We call A the matrix representation of $\mathcal{L}$ with respect to the given bases for $\mathbb{R}^n$ and $\mathbb{R}^m$.

In the special case where we assume the natural bases for $\mathbb{R}^n$ and $\mathbb{R}^m$, the matrix representation A satisfies

$$\mathcal{L}(\boldsymbol{x}) = A\boldsymbol{x}.$$

3.1 Linear Transformation

这本书上用以下记号 (同华师大大一线性代数教科书不同)

Let $\{\boldsymbol{e}_1, \boldsymbol{e}_2, ..., \boldsymbol{e}_n\}$ and $\{\boldsymbol{e}'_1, \boldsymbol{e}'_2, ..., \boldsymbol{e}'_n\}$ be two bases for $\mathbb{R}^n$. Define the matrix

$$T = [\boldsymbol{e}'_1, \boldsymbol{e}'_2, ..., \boldsymbol{e}'_n]^{-1}[\boldsymbol{e}_1, \boldsymbol{e}_2, ..., \boldsymbol{e}_n].$$

We call T the transformation matrix from $\{\boldsymbol{e}_1, \boldsymbol{e}_2, ..., \boldsymbol{e}_n\}$ to $\{\boldsymbol{e}'_1, \boldsymbol{e}'_2, ..., \boldsymbol{e}'_n\}$. It is clear that

$$[\boldsymbol{e}'_1, \boldsymbol{e}'_2, ..., \boldsymbol{e}'_n]T = [\boldsymbol{e}_1, \boldsymbol{e}_2, ..., \boldsymbol{e}_n];$$

i.e., T 的第 i 列是 $\boldsymbol{e}_i$ 在基 $\{\boldsymbol{e}'_1, \cdots, \boldsymbol{e}'_n\}$ 下的坐标.

3.1 Linear Transformation

这本书上用以下记号 (同华师大大一线性代数教科书不同)

Let $\{\boldsymbol{e}_1, \boldsymbol{e}_2, ..., \boldsymbol{e}_n\}$ and $\{\boldsymbol{e}_1', \boldsymbol{e}_2', ..., \boldsymbol{e}_n'\}$ be two bases for $\mathbb{R}^n$. Define the matrix

$$T = [\boldsymbol{e}_1', \boldsymbol{e}_2', ..., \boldsymbol{e}_n']^{-1}[\boldsymbol{e}_1, \boldsymbol{e}_2, ..., \boldsymbol{e}_n].$$

We call T the transformation matrix from $\{\boldsymbol{e}_1, \boldsymbol{e}_2, ..., \boldsymbol{e}_n\}$ to $\{\boldsymbol{e}_1', \boldsymbol{e}_2', ..., \boldsymbol{e}_n'\}$. It is clear that

$$[\boldsymbol{e}_1', \boldsymbol{e}_2', ..., \boldsymbol{e}_n']T = [\boldsymbol{e}_1, \boldsymbol{e}_2, ..., \boldsymbol{e}_n];$$

i.e., T 的第 i 列是 $\boldsymbol{e}_i$ 在基 $\{\boldsymbol{e}_1', \cdots, \boldsymbol{e}_n'\}$ 下的坐标.

我们大一教科书

$$(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \cdots, \boldsymbol{\eta}_n) = (\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \cdots, \boldsymbol{\varepsilon}_n)C,$$

我们称矩阵 C 为由基 $\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \cdots, \boldsymbol{\varepsilon}_n$ 到基 $\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \cdots, \boldsymbol{\eta}_n$ 的过渡矩阵.

Video 10 结束