1. 线性代数概念

线性：只含加法和数乘运算.

线性变换的几何直觉：保持网格线平行且等距分布的变换.

高斯消元 $\longrightarrow$ 最简行阶梯形矩阵 (Reduced row echelon matrix, RREF)

行阶梯形矩阵 (row echelon matrix)

解的存在性：利用RREF的最后一行不为0无解，否则有解 ($r=n$ 唯一解，$r<n$ 无穷解)

解的个数：无解，唯一解，无穷解.

2. 矩阵乘法

Matrix-Vector Multiplication

Definition I

Let $A=(a_{ij})_{m\times n}=\begin{bmatrix}\vec{\mathbf{a}}_1\\ \vec{\mathbf{a}}_2\\ \vdots\\ \vec{\mathbf{a}}_m\end{bmatrix}\in\mathbb{R}^{m\times n}$, where $\vec{\mathbf{a}}_1,\vec{\mathbf{a}}_2,\cdots,\vec{\mathbf{a}}_m$ are row vectors and $\mathbf{u}=(u_i)_{n\times 1}$ is a column vector, then

$$A\mathbf{u}=\begin{bmatrix}\vec{\mathbf{a}}_1\mathbf{u}\\ \vec{\mathbf{a}}_2\mathbf{u}\\ \vdots\\ \vec{\mathbf{a}}_m\mathbf{u}\end{bmatrix}$$

where $$\vec{\mathbf{a}}_i\mathbf{u}=a_{i1}u_1+a_{i2}u_2+\cdots+a_{in}u_n$$

is the scalar product of $\vec{\mathbf{a}}_i$ and $\mathbf{u}$.

Matrix-Vector Multiplication: Second Definition

Definition II

Let $A=(a_{ij})_{m\times n}=[\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_n]\in\mathbb{R}^{m\times n}$, where $\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_n$ are column vectors and $\mathbf{u}=(u_i)_{n\times 1}$ is also a column vector, then the matrix-vector product $A\mathbf{u}$ is

$$u_1\mathbf{a}_1+u_2\mathbf{a}_2+\cdots+u_n\mathbf{a}_n$$

which is a **linear Combination** of column vectors $\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_n$ with weights $u_1,\cdots,u_n$.

Remark: Two definitions produce the same results (They are equivalent).

★ properties：

$AB\neq BA$

$A(B+C)=AB+AC$

$(A+B)C=AC+BC$

✓ $(AB)x=A(Bx)$

✓ $(AB)C=A(BC)$

$(AB)^T=B^TA^T$

proof：

$$\begin{aligned}AB+AC&=[A\mathbf{b}_1,\ A\mathbf{b}_2,\ \cdots,\ A\mathbf{b}_l]+[A\mathbf{c}_1,\ A\mathbf{c}_2,\ \cdots,\ A\mathbf{c}_l]\\&=[A(\mathbf{b}_1+\mathbf{c}_1),\ A(\mathbf{b}_2+\mathbf{c}_2),\ \ldots,\ A(\mathbf{b}_l+\mathbf{c}_l)]\\&=A(B+C)\end{aligned}$$

$$BA+CA=\begin{bmatrix}\vec{b}_1\\ \vec{b}_2\\ \vdots\\ \vec{b}_n\end{bmatrix}A+\begin{bmatrix}\vec{c}_1\\ \vec{c}_2\\ \vdots\\ \vec{c}_n\end{bmatrix}A$$

(简单

$$=\begin{bmatrix}(\vec{b}_1+\vec{c}_1)A\\ (\vec{b}_2+\vec{c}_2)A\\ \vdots\\ (\vec{b}_n+\vec{c}_n)A\end{bmatrix}=(B+C)A$$

常量

$$\begin{aligned}A(Bx)&=A(b_1x_1+b_2x_2+\cdots+b_mx_m)\\&=x_1Ab_1+x_2Ab_2+\cdots+x_mAb_m\\&=[Ab_1,Ab_2,\cdots,Ab_m]\,x\\&=ABx\end{aligned}$$

(把x拆开再还原)

矩阵乘法证明题通常使用两种def.

• 矩阵转置

properties: $(A+B)^T = A^T + B^T$

$(\alpha A)^T = \alpha A^T$

$(A^T)^T = A$

☆ $(AB)^T = B^T A^T$

Proof. The (i,j)-entry of AB is $\sum_{k=1}^{n} a_{ik}b_{kj}$, and the (i,j)-entry of $(AB)^T$ is $\sum_{k=1}^{n} a_{jk}b_{ki}$.

(i,k)-entry of B^T is b_{ki}, and (k,j)-entry of A^T is a_{jk}, and (i,j)-entry of B^TA^T is $\sum_{k=1}^{n} b_{ki}a_{jk} = \sum_{k=1}^{n} a_{jk}b_{ki}$.

转置矩阵证明题变换下标

Thus, (i,j)-entry of $(AB)^T$ and B^TA^T are the same.

E.g. Show that AA^T and A^TA are symmetric matrix.

proof: (i,j)-entry of AA^T is $\sum_k a_{ik} a^T_{kj} = \sum_k a_{ik}a_{jk}$

(j,i)-entry of AA^T is $\sum_k a_{jk}a_{ik}$. $=$

A^TA is similar.

Property 5.14 Any square matrix can be written as a sum of a symmetric matrix and a skew-symmetric matrix.

$$A = \underbrace{\frac{A+A^T}{2}}_{\text{symmetric}} + \frac{A-A^T}{2} \rightarrow \text{skew-symmetric}$$

• 矩阵分块乘法

Matrix with blocks which are zero or identity matrix

Example

$$A = \left[\begin{array}{cc|cc} 1 & 1 & 0 & 0 \\ \hline 0 & 0 & 1 & 1 \\ 0 & 0 & 2 & 2 \end{array}\right] = \begin{bmatrix} A_{11} & O \\ O & A_{22} \end{bmatrix}$$

$$B = \left[\begin{array}{cc|cc} 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ \hline 3 & 1 & 1 & 0 \\ 3 & 2 & 0 & 1 \end{array}\right] = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & I_2 \end{bmatrix}$$

$$AB = \begin{bmatrix} A_{11}B_{11} & A_{11}B_{12} \\ A_{22}B_{21} & A_{22}B_{22} \end{bmatrix} = \left[\begin{array}{cc|cc} 2 & 3 & 2 & 2 \\ \hline 6 & 3 & 1 & 1 \\ 12 & 6 & 2 & 2 \end{array}\right]$$

• 可逆矩阵

invertible matrix (nonsingular / nondegenerate) 唯一性

$\begin{bmatrix} 2 & 1 \\ 0 & 0 \end{bmatrix}$ 不存在. ☆

$AB = I_n$, $BA = I_n \Leftrightarrow ABA = AI_n \Leftrightarrow I_nA = AI_n$ ✓

properties: $(A^T)^{-1} = (A^{-1})^T$

$(\alpha A)^{-1} = \frac{1}{\alpha}A^{-1}$

$(A^{-1})^{-1} = A$

$(AB)^{-1} = B^{-1}A^{-1}$

$(A+B)^{-1} \neq A^{-1} + B^{-1}$

proof: $A^T(A^{-1})^T = (A^{-1}A)^T = I$

$A(BB^{-1})A^{-1} = AIA^{-1} = I$

求逆矩阵

$\underbrace{E_k \cdots E_2E_1}_{A^{-1}}[A \mid I] \rightarrow [I \mid P]$

$A^{-1}I = P \Rightarrow A^{-1} = PI^{-1} = P$

初等矩阵：单位矩阵只进行一次行变换

(3) The elementary matrix corresponding to elementary row operation 3 ($R_j \to \beta R_i + R_j$) is (elementary matrix type III)

$$
E_{\beta R_i+R_j} = \begin{array}{c}
\left[\begin{array}{ccc|c|ccc|c|ccc}
1 & & & & & & & & & & \\
 & \ddots & & & & & & & & & \\
 & & 1 & & & & & & & & \\
\hline
 & & & 1 & & & & & & & \\
\hline
 & & & & 1 & & & & & & \\
 & & & & & \ddots & & & & & \\
 & & & & & & 1 & & & & \\
\hline
 & & & \beta & & & & 1 & & & \\
\hline
 & & & & & & & & 1 & & \\
 & & & & & & & & & \ddots & \\
 & & & & & & & & & & 1
\end{array}\right]
\end{array}
\begin{array}{l}
\\ \\ \\ i\text{th row} \\ \\ \\ \\ j\text{th row} \\ \\ \\ \\
\end{array}
$$

ith column $\quad$ jth column

（用于描述 row operation）

- **行变换矩阵 E**

properties： $\boxed{E^{-1}_{R_iR_j} = E_{R_iR_j} = E^{T}_{R_iR_j}}$

$E^{-1}_{\alpha R_i} = E_{\frac{1}{\alpha}R_i}$

$E^{-1}_{\beta R_i+R_j} = E_{-\beta R_i+R_j}$

（E^{-1} 即为逆行变换）

Property 6.10 For a permutation matrix P, $P^{-1} = P^T$, since

$$
\begin{aligned}
P^{-1} &= (E_{R_{i_k}R_{j_k}} \cdots E_{R_{i_2}R_{j_2}} E_{R_{i_1}R_{j_1}})^{-1} \\
&= E^{-1}_{R_{i_1}R_{j_1}} E^{-1}_{R_{i_2}R_{j_2}} \cdots E^{-1}_{R_{i_k}R_{j_k}} \\
&= E_{R_{i_1}R_{j_1}} E_{R_{i_2}R_{j_2}} \cdots E_{R_{i_k}R_{j_k}} \\
&= E^{T}_{R_{i_1}R_{j_1}} E^{T}_{R_{i_2}R_{j_2}} \cdots E^{T}_{R_{i_k}R_{j_k}} \\
&= P^T
\end{aligned}
$$

upper / lower triangular matrix 同类型相乘不改变性质.

用途：

- **LU decomposition** ☆ （减小空间）

First solve the linear system $L\underline{\mathbf{y}} = \mathbf{b}$ by using **forward substitution**:

解 y 即 Ux 的值.

$$
\begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ 2 & -3 & 1 \end{bmatrix}
\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}
= \begin{bmatrix} -4 \\ -7 \\ 15 \end{bmatrix}
\Rightarrow
\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}
= \begin{bmatrix} -4 \\ -5 \\ 8 \end{bmatrix}
$$

$-7 - \frac{1}{2}y_1 = -5$

$15 - 2y_1 + 3y_2 = 8$

方便求解.

Then solve the linear system $U\underline{\mathbf{x}} = \mathbf{y}$ by using **backward substitution**:

$$
\begin{bmatrix} 2 & 4 & 2 \\ 0 & 3 & 1 \\ 0 & 0 & 8 \end{bmatrix}
\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}
= \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}
\Rightarrow
\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}
= \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}
$$

$x_1 \to \frac{1}{2}(-4-2-4(-2)) = 1$

$x_2 \to \frac{1}{3}(-5-1) = -2$

$x_3 \to 1$

triangular matrix 方便求解.

In software, Matlab's command $x = A \setminus b$ use LU-decomposition, along with forward and backward substitution to solve the linear system when the matrix A is nonsingular.

Note that here A is nonsingular, the diagonal entries of U are nonzero, U can be further decomposed into

$$U = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 2 & 3 \\ 0 & 0 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & \frac{3}{2} \\ 0 & 0 & 1 \end{bmatrix} = D\hat{U}$$

where D is a diagonal matrix with nonzero diagonal entries, $\hat{U}$ is a unit upper triangular matrix (an upper triangular matrix whose diagonal entries are all 1's).
Thus

coefficient matrix

$$PA = LU = \underline{LD}\hat{U}$$ → unit triangular matrix

This is the LDU decomposition.

可逆矩阵的等价判定条件☆

Theorem 8.2 (**Equivalent conditions for invertible matrix**)
$A \in \mathbb{R}^{n\times n}$, the following are equivalent:
(1) A is invertible,
(2) the linear system $\boxed{A\mathbf{x} = \mathbf{0}}$ has only a trivial solution,
(3) matrix A is row equivalent to I_n,
(4) A is a product of elementary matrices,
(5) there exists an invertible matrix $E \in \mathbb{R}^{n\times n}$ such that $EA = I_n$.
(6) $\boxed{A\mathbf{x} = \mathbf{b}}$ has a unique solution for any $\mathbf{b}$.

Proof.
(1)⇒(2)
Since A is invertible, then $A\mathbf{x} = \mathbf{0}$ has a unique solution $\mathbf{x} = A^{-1}\mathbf{0} = \mathbf{0}$.

两边同乘 A^{-1}(等价变换)

(2)⇒(3)
Suppose

$$[A|\mathbf{0}] \xrightarrow{\text{elementary row operations}} [B|\mathbf{0}](\textit{reduced row} - \textit{echelon form})$$

Since the linear system $A\mathbf{x} = \mathbf{0}$ has only a trivial solution, thus each row of B must has a leading 1.
Thus, $B = I_n$.

(3)⇒ (4)
By theorem 8.2, there are elementary matrices $E_1, \cdots, E_k$, such that $E_k \cdots E_1 A = I_n$. Thus $A = E_1^{-1} \cdots E_k^{-1}$ is a product of elementary matrices since $E_1^{-1}, \cdots, E_k^{-1}$ are also elementary matrices.

(4)⇒ (5)
Suppose $A = E_1 \cdots E_k$ ($E_1, \cdots, E_k$ are elementary matrices), then $E_k^{-1} \cdots E_1^{-1} A = I_n$. Let $E = E_k^{-1} \cdots E_1^{-1}$, then E is invertible and $EA = I_n$.

(5)⇒ (1)
Since $EA = I_n$ and E is invertible, then $A = E^{-1}$ is also invertible, and $A^{-1} = E$.

(1)⇒ (6)
$A\mathbf{x} = \mathbf{b} \Rightarrow \mathbf{x} = A^{-1}\mathbf{b}$ is the unique solution for any $\mathbf{b}$.
(6)⇒ (1) 反证法☆
If A is singular, then $A\mathbf{x} = \mathbf{0}$ has infinity many solutions from (2) by the contrapositive statement. Now taking $\mathbf{z} \neq \mathbf{0}$ is the solution of $A\mathbf{x} = \mathbf{0}$, and suppose that $\mathbf{y}$ is the unique solution of $A\mathbf{x} = \mathbf{b}$, then $A(\mathbf{y}+\mathbf{z}) = A\mathbf{y} + A\mathbf{z} = \mathbf{b} + \mathbf{0} = \mathbf{b}$, thus $\mathbf{y}+\mathbf{z}$ is also the solution of $A\mathbf{x} = \mathbf{b}$. But $\mathbf{y} \neq \mathbf{y}+\mathbf{z}$. This is a contradiction.

Remark.
If A is invertible, then A is row equivalent to I, i.e., $A \xrightarrow{\text{row op}_1,\cdots,\text{op}_k} I$, suppose the corresponding elementary matrices for the row operations A^{-1} $\text{op}_1, \text{op}_2, \cdots, \text{op}_k$ are $E_1, E_2, \cdots, E_k$, then $I = \boxed{E_k \cdots E_1}A$. Thus, $A^{-1} = E_k \cdots E_1 I$. Thus, for the same series of elementary row operations, it will transforms a nonsingular matrix A to I and transform I to A^{-1}.
This suggests a method to find A^{-1} by performing row operations for augmented matrix $[A|I]$. → $[I|A^{-1}]$

$[A|B] \to [I|A^{-1}B]$

$AX = B \to X = A^{-1}B$

Method to find A^{-1} (A is invertible)

$$[A|I] \xrightarrow{\text{Gauss Jordan elimination}} [I|P], \text{ then } P = A^{-1}.$$

Method to find X such that $AX = B$ (A is invertible)

$$[A|B] \xrightarrow{\text{Gauss Jordan elimination}} [I|X], \text{ then } X = A^{-1}B.$$

Theorem 8.5 (**Nonsingular Product has Nonsingular Terms**)
Suppose that A and B are square matrices with the same size. The product AB is nonsingular if and only if A and B are both nonsingular.

Proof.
$\Leftarrow$ Since $(AB)^{-1} = B^{-1}A^{-1}$, $(AB)^{-1}$ exists.

$\Rightarrow$ For this portion of the proof we will form the logically-equivalent contrapositive 对应的 statement: "AB is nonsingular implies A and B are both nonsingular" is equivalent to "A or B is singular implies AB is singular". 逆否命题.

方法二：
AB is invertible
$\Rightarrow \exists C.\ ABC = I$
$A(BC) = I$, $A^{-1} = BC$
$\Rightarrow \exists C.\ CAB = I$
$(CA)B = I$, $B^{-1} = CA$.

Case (1). Suppose B is singular, then there is z nonzero vector $\mathbf{z}$ s.t. $B\mathbf{z} = \mathbf{0}$, thus

$$AB\mathbf{z} = A(B\mathbf{z}) = A\mathbf{0} = \mathbf{0}$$

Therefore $AB\mathbf{x} = \mathbf{0}$ has a nonzero solution, this means that AB is singular.

Case (2). Suppose B is invertible and A is singular, then if AB is invertible, then there is a invertible matrix C, s.t. $ABC = I$, thus $A = C^{-1}B^{-1}$ which is invertible. This is a contradiction with the assumption A is singular. 反证法.

Theorem 8.6 (**One-Sided Inverse Verification is Sufficient**) Suppose $A, B \in \mathbb{R}^{n\times n}$. If $BA = I_n$, then $AB = I_n$.

Proof. By theorem 8.5, both A and B are invertible. By theorem 8.2, there exists $E \in \mathbb{R}^{n\times n}$, s.t. $EB = I$. Thus

$$AB = IAB = (EB)AB = E(BA)B = EIB = EB = I.$$

3. 向量空间

Definition 9.1 (**Euclidean Vector Space**) The set defined as
欧式向量空间 (线性空间)

$$\mathbb{R}^m = \left\{ \begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix} \middle| u_i \in \mathbb{R} \right\}.$$

is called the Euclidean Vector Space.

向量空间要求对加法，数乘封闭 (保证结果仍为合法向量)

Definition 11.1 (**Vector Space**) Let V be a set of **elements** associated with two operations:

(i) Addition "+": $\mathbf{u} + \mathbf{v} \in V$, $\forall\, \mathbf{u},\ \mathbf{v} \in V$. (**Additive Closure**)

(ii) Scalar multiplication: $\alpha\mathbf{u} \in V$, $\forall\, \alpha \in \mathbb{R},\ \mathbf{u} \in V$. (**Scalar Closure**)

向量空间和可逆矩阵的联系.

Theorem 10.4 (**More Equivalent Conditions for Invertible Matrix**)
Suppose that A is a square matrix. The following are equivalent.

(7). The column vectors of A are linearly independent.

☆ 线性方程组解的参数向量形式

Example 9.3

$$\begin{aligned} 2x_1 + x_2 + 7x_3 - 7x_4 &= 8 \\ -3x_1 + 4x_2 - 5x_3 - 6x_4 &= -12 \\ x_1 + x_2 + 4x_3 - 5x_4 &= 4 \end{aligned}$$

Its augmented matrix can be reduced into reduced row-echelon form as

$$\begin{array}{c} \begin{array}{cccc} x_1 & x_2 & x_3 & x_4 \end{array} \\ \left[\begin{array}{cccc|c} \boxed{1} & 0 & 3 & -2 & 4 \\ 0 & \boxed{1} & 1 & -3 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] \\ \text{dependent} \quad \text{independent} \end{array}$$

自由项作为参数.

Null(A)
||

$$\mathbf{x} = \underbrace{\begin{bmatrix} 4 \\ 0 \\ 0 \\ 0 \end{bmatrix}}_{w_0} + x_3 \begin{bmatrix} -3 \\ -1 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 2 \\ 3 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \operatorname{span}\left\{ \begin{bmatrix} -3 \\ -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \\ 0 \\ 1 \end{bmatrix} \right\}$$

w_0 为特解，S_h 为 $Ax=0$ 的齐次方程解，则 $Ax=b$ 的解可表示为 $S = \{ s \mid s = w_0 + s_h,\ s_h \in S_h \}$

齐次方程解.

线性无关与线性相关

Definition 10.1 (**Linearly independent**) Given a set of vectors $\mathcal{U} = \{\mathbf{u}_1, \cdots, \mathbf{u}_n\} \subseteq \mathbb{R}^m$, we say that $\mathcal{U}$ is linearly independent if $c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_n\mathbf{u}_n = \mathbf{0}$ $(c_1, \cdots, c_n \in \mathbb{R})$ is valid only when $c_1 = c_2 = \cdots = c_n = 0$.

Definition 10.2 (**Linearly dependent**) Given a set of vectors $\mathcal{U} = \{\mathbf{u}_1, \cdots, \mathbf{u}_n\} \subseteq \mathbb{R}^m$, then $\mathcal{U}$ is linearly dependent if there exists a set of real numbers $c_1, \cdots, c_n$ which are not all zeros $((c_1, \cdots, c_n) \neq (0, 0, \cdots, 0))$, such that $c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_n\mathbf{u}_n = \mathbf{0}$.

- 生成空间（span） 所有线性组合形成的空间
- 多项式空间 （polynomial vector space）
- 连续函数空间 （?）

· 向量空间的性质

Then V is called a **vector space over** $\mathbb{R}$ if the following eight axioms are satisfied:

(A1) $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}, \forall\ \mathbf{u},\ \mathbf{v} \in V$.

(A2) $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + \mathbf{v} + \mathbf{w}, \forall\ \mathbf{u},\ \mathbf{v},\ \mathbf{w}\ \in V$.

(A3) There exists a element $\mathbf{0}$ s.t. $\mathbf{u} + \mathbf{0} = \mathbf{u}$, $\forall\ \mathbf{u} \in V$. ✓ 0的性质

(A4) If $\mathbf{u} \in V$, then there exists $-\mathbf{u} = (-1)\mathbf{u}$, s.t. $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$. ✓

(A5) $\alpha(\mathbf{u} + \mathbf{v}) = \alpha\mathbf{u} + \alpha\mathbf{v}, \forall\ \alpha \in \mathbb{R},\ \mathbf{u},\ \mathbf{v} \in V$.

(A6) $(\alpha + \beta)\mathbf{u} = \alpha\mathbf{u} + \beta\mathbf{u}, \forall\ \alpha,\ \beta \in \mathbb{R},\ \mathbf{u} \in V$.

(A7) $\alpha(\beta\mathbf{u}) = (\alpha\beta)\mathbf{u}, \forall\ \alpha,\ \beta \in \mathbb{R},\ \mathbf{u} \in V$.

(A8) $1\mathbf{u} = \mathbf{u}$. 1的定义

And we call the elements in the set V are **"vectors"**.

Theorem 11.6 (**More Properties of Vector Space**) Let V be a vector space over $\mathbb{R}$, then

(1) Zero vector is unique.

(2) $0\mathbf{x} = \mathbf{0}, \forall\ \mathbf{x} \in V$.

(3) $c\mathbf{0} = \mathbf{0}, \forall\ c\ \in \mathbb{R}$.

(4) (The additive inverse if unique) For each $\mathbf{u} \in V$ there is a unique $\mathbf{y} \in V$ so that $\mathbf{x} + \mathbf{y} = \mathbf{0}$. (This element $\mathbf{y}$ is denoted by $-\mathbf{x}$.)

Proof.

❑ The zero vector $\mathbf{0} \in V$ is unique.

Proof: Suppose $\mathbf{0}$ and $\mathbf{0}'$ are both zero vectors. We have $\mathbf{0} \overset{A3}{=} \mathbf{0} + \mathbf{0}' \overset{A1}{=} \mathbf{0}' + \mathbf{0} \overset{A3}{=} \mathbf{0}'$

❑ $0\mathbf{x} = \mathbf{0}$ for each $\mathbf{x} \in V$.

Proof: $\mathbf{x} \overset{A8}{=} 1\mathbf{x} = (1+0)\mathbf{x} \overset{A6}{=} \boxed{\mathbf{x} + 0\mathbf{x}}$

$\mathbf{0} \overset{A4}{=} -\mathbf{x} + \mathbf{x} = -\mathbf{x} + (\boxed{\mathbf{x} + 0\mathbf{x}}) \overset{A2}{=} (-\mathbf{x} + \mathbf{x}) + 0\mathbf{x} \overset{A4}{=} \mathbf{0} + 0\mathbf{x} \overset{A1,A3}{=} 0\mathbf{x}$

❑ $c\mathbf{0} = \mathbf{0}$ for any scalar c.

Proof: $c\mathbf{0} = c\mathbf{0} + \mathbf{0} = c\mathbf{0} + (c\mathbf{0} + (-c\mathbf{0})) = (c\mathbf{0} + c\mathbf{0}) + (-c\mathbf{0})$
$= c(\mathbf{0} + \mathbf{0}) + (-c\mathbf{0}) = c\mathbf{0} + (-c\mathbf{0}) = \mathbf{0}$

❑ For any $\mathbf{x}, \mathbf{y} \in V$, if $\mathbf{y} + \mathbf{x} = \mathbf{0}$ then $\mathbf{y} = -\mathbf{x}$. $\mathbf{y} \overset{A3}{=} \mathbf{y} + \mathbf{0} \overset{A4}{=} \mathbf{y} + (\mathbf{x} + (-\mathbf{x}))$

Definition 11.8 (**Subspace [alternative, equivalent definition]**) Let V be a vector space over $\mathbb{R}$. A subset $W \subseteq V$ is a subspace of V if the following are satisfied:

(1). $\mathbf{0} \in W$.

运算封闭

(2). W is closed under vector addition: $\forall\ \mathbf{u}, \mathbf{v} \in W$, we have $\mathbf{u} + \mathbf{v} \in W$.

(3). W is closed under scalar multiplication: $\forall\ \alpha \in \mathbb{R}, \mathbf{u} \in W$, we have $\alpha\mathbf{u} \in W$.

基底和维

Definition 12.1 (**Basis**) A subset $\mathcal{U} = \{\mathbf{u}_1, \cdots, \mathbf{u}_m\}$ from the vector space V is a **basis** of V if
(1) $\mathcal{U}$ is linearly independent.
(2) **Span**$(\mathcal{U}) = V$, that is $\mathcal{U}$ spans V.

Example 12.3 (**Standard basis for $V = \mathbb{R}^n$**) Let

标准基

$$\mathbf{e}_i = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad i\text{th row},$$

then $\mathcal{E} = \{\mathbf{e}_1, \cdots, \mathbf{e}_n\}$ is a basis for $\mathbb{R}^n$. In particular, $\mathcal{E}$ is called the standard basis.

Lemma 12.6 (**A vector set is linearly dependent if # of vectors in the set is larger than # of vectors in the basis**) Let $\mathcal{U} = \{\mathbf{u}_1, \cdots, \mathbf{u}_n\}$ be the basis of vector space V, then $T = \{\mathbf{v}_1, \cdots, \mathbf{v}_m\} \subseteq V$ is linearly dependent if $m > n$.

Definition 12.8 (**Dimension**) Let V be a vector space and let $\mathcal{U}$ be a basis of V, then the number of vectors in $\mathcal{U}$ is called the dimension of V. Denoted by $\dim(V)$.

零向量自身是线性相关的.

Example 12.9 (0) $\dim(\{\mathbf{0}\}) = 0$, since $\{\mathbf{0}\}$ has no basis vector.

Example 12.9 (1) $\dim(\mathbb{R}^n) = n$, since $\{\mathbf{e}_i, i = 1, \cdots, n\}$ is the standard basis.

基底转换

Coordinate transformation in general vector spaces

Theorem 12.15 (**Transition Matrix between two bases**) Let $\mathcal{U} = \{\mathbf{u}_1, \cdots, \mathbf{u}_n\}$ and $\mathcal{V} = \{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$ be two bases of vector space V. Then

逐项分解 $\overbrace{}^{n\times n\text{次}}$

$$[\mathbf{x}]_{\mathcal{V}} = A[\mathbf{x}]_{\mathcal{U}}$$

$[x]=[u_1|u_2|\cdots|u_n]\begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{bmatrix}$

$=[v_1|v_2|\cdots|v_n]\begin{bmatrix} d_1' \\ d_2' \\ \vdots \\ d_n' \end{bmatrix}$

where the jth column of A is $[\mathbf{u}_j]_{\mathcal{V}}$.

Proof.

Let $[\mathbf{x}]_{\mathcal{U}} = \underbrace{(d_1, d_2, \cdots, d_n)}_{\text{coefficient}}{}^T$, then $\mathbf{x} = d_1\mathbf{u}_1 + d_2\mathbf{u}_2 + \cdots + d_n\mathbf{u}_n$. Suppose $[\mathbf{u}_j]_{\mathcal{V}} = \mathbf{a}_j (j = \cdots, n)$. According to the above lemma 12.17,

线性可加

$x=c_1u_1+c_2u_2+\cdots+c_nu_n$

$=[u_1|u_2|\cdots|u_n]\begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$

$(=U[x]_u)$

$x=d_1v_1+d_2v_2+\cdots+d_nv_n$

$=[v_1|v_2|\cdots|v_n]\begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{bmatrix}$

$(=V[x]_v)$

$$\begin{aligned}
[\mathbf{x}]_{\mathcal{V}} &= d_1[\mathbf{u}_1]_{\mathcal{V}} + d_2[\mathbf{u}_2]_{\mathcal{V}} + \cdots + d_n[\mathbf{u}_n]_{\mathcal{V}} \\
&= d_1\mathbf{a}_1 + d_2\mathbf{a}_2 + \cdots + d_n\mathbf{a}_n \\
&= [\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_n](d_1, d_2, \cdots, d_n)^T \quad \text{(乘法原理)} \\
&= A[\mathbf{x}]_{\mathcal{U}}
\end{aligned}$$

☆ a_i 第 j 行对应最终 v_j 的贡献. (做题有感)

A is the transition matrix corresponding to the change of basis from $\mathcal{U}$ to $\mathcal{V}$.

$[x]_v = V^{-1}U[x]_u$ → transition matrix

Example 12.16

Find the transition matrix corresponding to the change of basis from

$$\mathcal{U} = \left\{ \begin{bmatrix} 5 \\ 2 \end{bmatrix}, \begin{bmatrix} 7 \\ 3 \end{bmatrix} \right\} \text{ to } \mathcal{V} = \left\{ \begin{bmatrix} 3 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}.$$

$[x]_v = A[x]_u$

Since

$$\mathbf{u}_1 = 3\begin{bmatrix} 3 \\ 2 \end{bmatrix} - 4\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\mathbf{u}_2 = 4\begin{bmatrix} 3 \\ 2 \end{bmatrix} - 5\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Let $A = [\mathbf{a}_1, \mathbf{a}_2]$, then

$$\mathbf{a}_1 = [\mathbf{u}_1]_{\mathcal{V}} = \begin{bmatrix} 3 \\ -4 \end{bmatrix}, \quad \mathbf{a}_2 = [\mathbf{u}_2]_{\mathcal{V}} = \begin{bmatrix} 4 \\ -5 \end{bmatrix}$$

Thus

$$A = \begin{bmatrix} 3 & 4 \\ -4 & -5 \end{bmatrix}$$

零空间 null space

Definition 13.1 (**Null Space**) Let $A \in \mathbb{R}^{m\times n}$, then the null space of A is the set Null(A)=$\{\mathbf{x} \in \mathbb{R}^n | A\mathbf{x} = \mathbf{0}\}$. The Null Space of A is also the solution set for $A\mathbf{x} = \mathbf{0}$.

Theorem 13.2 (**Null(A) is subspace**)
Let $A \in \mathbb{R}^{m\times n}$, then Null$(A)$ is a subspace of $\mathbb{R}^n$, hence is a vector space.

slide 13

列空间 column space

Definition 13.4 Let $A = [\mathbf{a}_1, \cdots, \mathbf{a}_n] \in \mathbb{R}^{m\times n}$, the column space of A is defined as the vector space

$$\mathrm{Col}\,(A) = \mathbf{Span}\,(\mathbf{a}_1, \cdots, \mathbf{a}_n) \subseteq \mathbb{R}^m$$

Some book uses C(A) to denote the column space of A.

Together with dim(Col (A))=the number of pivot columns of A. We have the following corollary:

Corollary: Let A be an $m \times n$ matix.

$$\dim(\mathrm{Col}\,(A)) + \dim(\mathrm{Null}\,(A)) = n.$$

行空间 row space

Definition 14.1 Let $A = (a_{ij})_{m\times n} = \begin{bmatrix} \vec{\mathbf{a}}_1 \\ \vec{\mathbf{a}}_2 \\ \vdots \\ \vec{\mathbf{a}}_m \end{bmatrix}$, then

$A^T = [(\vec{\mathbf{a}}_1)^T, \cdots, (\vec{\mathbf{a}}_m)^T]$, the row space of A is defined as

$$\mathrm{Row}\,(A) = \mathbf{Span}\left((\vec{\mathbf{a}}_1)^T, \cdots, (\vec{\mathbf{a}}_m)^T\right) \subseteq \mathbb{R}^n$$

By definition,

$$\mathrm{Row}\,(A) = \mathrm{Col}\,(A^T)$$

秩 Rank.

Theorem 14.10 (**Definition for Rank**) $r = \dim(\mathrm{Row}\,(A)) = \dim(\mathrm{Col}\,(A))$ is called the rank of A, denoted by $\mathrm{rank}(A)$. By above theorem 14.9, one can readily see that $\mathrm{rank}(A) = \mathrm{rank}(A^T)$

满秩矩阵.

Definition 14.11 (**Full Rank Matrix**) Let $A \in \mathbb{R}^{m\times n}$, if $\mathrm{rank}(A) = \min(m, n)$, then A is called the full rank matrix.

Definition 14.13 (**Nullity**) Let $A \in \mathbb{R}^{m\times n}$, then the nullity of A is $n(A) = \dim(\mathrm{Null(A)})$.

rank (A) + n(A) = n.
rank-nullity theorem 秩-零度化定理.

Theorem 14.15 ☆ (**Rank-Nullity Theorem**) Let $A \in \mathbb{R}^{m\times n}$, then

$$\mathrm{rank}(A) + n(A) = n$$

Theorem[Invertible Matrix Theorem] Given a $n \times n$ matrix $A \in \mathbb{R}^{n \times n}$ (square!). The following are equivalent: (Only some important equivalent conditions are listed here)

1. A is invertible
2. the only solution to $A\mathbf{x} = \mathbf{0}$ is the trivial solution
3. $A\mathbf{x} = \mathbf{b}$ has a solution for every $\mathbf{b}$
4. A is row equivalent to I
5. A is a product of elementary matrices
6. the columns/rows of A are linearly independent
7. the columns/rows of A span $\mathbb{R}^n$, i.e., $\text{Row}\,(A) = \text{Col}\,(A) = \mathbb{R}^n$.
8. $\text{Null}\,(A) = \{\mathbf{0}\}$
9. $\text{rank}(A) = n$
10. $\det(A) \neq 0$.

行列式

(Definition of the determinant by using its property)

- For two $n \times n$ matrices A and B

$$\det(AB) = \det(A)\det(B)$$

- For an $n \times n$ triangular matrix A with diagonal entries $a_{11}, \cdots, a_{nn}$:

$$\det(A) = a_{11}a_{22}\cdots a_{nn}$$

$$\det(A^{-1}) = \frac{1}{\det(A)}$$

Elementary matrix :

I : $R_i \leftrightarrow R_j$ $\qquad \det(\mathrm{I}) = -1$

II : $R_i \to \alpha R_i$ $\qquad \det(\mathrm{II}) = \alpha$

III : $R_i \to R_i + \alpha R_j$ $\qquad \det(\mathrm{III}) = 1$

$E_1 = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \qquad E_2 = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \qquad E_3 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \qquad E_4 = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$

$$\underset{A}{\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}} \xrightarrow{\mathrm{III}} \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \xrightarrow{\mathrm{II}} \begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix} \xrightarrow{\mathrm{III}} \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \xrightarrow{\mathrm{III}} \underset{B}{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}}$$

$\det(\mathrm{III}) = \det(E_1)\det(E_2)\det(E_3)\det(E_4) = 1 \times (-1) \times 1 \times 1 = -1$.

Determinant of a Matrix with LU Decomposition

- In general, permutation of the rows of A is required to obtain an LU decomposition, i.e., $PA = LU$, where P is a permutation matrix. $A = P^{-1}LU = P^{T}LU$
- $\det(A) = \det(P^{T})\det(L)\det(U)$.

↳ 只需消元时记录交换次数. $(-1)^{k}$

✓ $\det(L) = 1$ ✓三角矩阵行列式相等.

✓ $\det(A^{T}) = \det(U^{T})\det(L^{T})\det(P)$

$= \det(U)\det(L)\det(P^{T})$

$= \det(P^{T})\det(L)\det(U)$

$= \det(A)$

Property 6

Consider $n \times n$ matrices A, B and C defined as follows:

$$A = \begin{bmatrix} \vec{\mathbf{a}} \\ \vec{\mathbf{c}}_2 \\ \vdots \\ \vec{\mathbf{c}}_n \end{bmatrix}, \quad B = \begin{bmatrix} \vec{\mathbf{b}} \\ \vec{\mathbf{c}}_2 \\ \vdots \\ \vec{\mathbf{c}}_n \end{bmatrix}, \quad C = \begin{bmatrix} \vec{\mathbf{a}} + \vec{\mathbf{b}} \\ \vec{\mathbf{c}}_2 \\ \vdots \\ \vec{\mathbf{c}}_n \end{bmatrix}.$$

We have $\det(C) = \det(A) + \det(B)$.

① 若 $\vec{c}_2, \ldots, \vec{c}_n$ 是 linearly dependent.

$\det(A) = \det(B) = \det(C) = 0$.

② 若 $\vec{c}_2, \ldots, \vec{c}_n$ 是 linearly independent. ($\vec{a}, \vec{b}, \vec{c}_2, \ldots, \vec{c}_n$ 一定线性相关)

(1) $\vec{a}, \vec{c}_2, \ldots, \vec{c}_n$ 线性相关.

$\vec{a} = k_2\vec{c}_2 + \ldots + k_n\vec{c}_n$.

$\det(C) = \det(B) + 0 = \det(B) + \det(A)$

(2) $\vec{a}$ 与 $\vec{c}_2, \ldots, \vec{c}_n$ 线性无关. $\vec{a}, \vec{c}_2, \cdots, \vec{c}_n$ 构成 basis.

$$B = \begin{bmatrix} \vec{b} \\ \vec{c}_2 \\ \vdots \\ \vec{c}_n \end{bmatrix} = \begin{bmatrix} \alpha\vec{a} + \sum kc \\ \vec{c}_2 \\ \vdots \\ \vec{c}_n \end{bmatrix}$$

$\det(C) = (1+\alpha)\det(A) = \det(A) + \det(B)$.

Definition 16.2

For matrix (a_{ij}), we call $\det(M_{ij})$ the minor 次矩阵 of a_{ij}, and $C_{ij} = (-1)^{i+j}\det(M_{ij})$ the cofactor 代数余子式 of a_{ij}.

Property 2

If A is an $n \times n$ matrix with $n \geq 2$, then

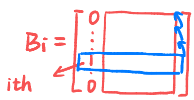
$$\begin{aligned}\det(A) &= a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in} \\ &= \underline{a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj}}\end{aligned}$$

for $i, j = 1, 2, \ldots, n$. (i, j 为定值)

lemma: $\det\left(\begin{bmatrix} a & \cdots \\ 0 & M \end{bmatrix}\right) = a\det(M)$, 上面的性质※.

构造 $B_i = [e_i \ a_2 \cdots a_n]$, $A = \sum_{i=1}^{n} a_{1i} B_i$.

由性质※ $\det(A) = \sum_{i=1}^{n} a_{1i} \det(B_i)$



$\det(B_i) = (-1)^{i-1}\det(M_{1i})$

$= (-1)^{i+1}\det(M_{1i})$

$\therefore \det(A) = a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj}$.

Adjoint Matrix 邻接矩阵.

- Define the adjoint of matrix A by 矩阵乘法法则

$$\mathrm{adj}(A) = \begin{bmatrix} C_{11} & C_{12} & \cdots & C_{1n} \\ C_{21} & C_{22} & \cdots & C_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ C_{n1} & C_{n2} & \cdots & C_{nn} \end{bmatrix}^T = \begin{bmatrix} C_{11} & C_{21} & \cdots & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n2} \\ \vdots & & & \\ C_{1n} & C_{2n} & \cdots & C_{nn} \end{bmatrix}.$$

$$\underline{A\ \mathrm{adj}(A)} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & & & \vdots \\ a_{n1} & & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} C_{11} & \cdots & C_{j1} & \cdots & C_{n1} \\ C_{12} & & C_{j2} & & \vdots \\ \vdots & & \vdots & & \vdots \\ C_{1n} & \cdots & C_{jn} & \cdots & C_{nn} \end{bmatrix}$$

$S_{ij} = \begin{cases} \det(A), & i = j \\ 0, & i \neq j \end{cases}$

$= \det(A)\ \underline{I}$

$$A^{-1} = \frac{\mathrm{adj}(A)}{\det(A)} \qquad A^{-1}_{ij} = \frac{C_{ji}}{\det(A)}$$

$$\begin{array}{c} \\ \text{ith} \\ \text{jth} \\ \\ \end{array} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ a_{n1} & & \cdots & a_{nn} \end{bmatrix}$$

Cramer's Rule

Theorem Let $A = [\mathbf{a}_1, \cdots, \mathbf{a}_n]$ ($\mathbf{a}_1, \cdots, \mathbf{a}_n$ are column vectors of A) be a nonsingular $n \times n$ matrix and let $\mathbf{b} \in \mathbb{R}^n$ and $A_i = [\mathbf{a}_1, \cdots, \mathbf{a}_{i-1}, \mathbf{b}, \mathbf{a}_{i+1}, \cdots, \mathbf{a}_n]$, then the unique solution of $A\mathbf{x} = \mathbf{b}$ is given by

$$x_i = \frac{\det(A_i)}{\det(A)}, \quad \text{for } i = 1, 2, \cdots, n$$

Proof.

$$\mathbf{x} = A^{-1}\mathbf{b} = \frac{1}{\det(A)}\mathrm{adj}(A)\mathbf{b}.$$

$$x_i = \frac{b_1 C_{1i} + b_2 C_{2i} + \cdots + b_n C_{ni}}{\det(A)} = \frac{\det(A_i)}{\det(A)}$$

where $\det(A_i) = b_1 C_{1i} + b_2 C_{2i} + \cdots + b_n C_{ni}$ (expand along ith column) and $A_i = [\mathbf{a}_1, \cdots, \mathbf{a}_{i-1}, \mathbf{b}, \mathbf{a}_{i+1}, \cdots, \mathbf{a}_n]$ $\square$

Leibniz Formula

- Observe that the determinant of an $n \times n$ ($n = 2$ or 3) matrix can be written as the summation of $n!$ terms, each of the multiplication of n entries in A, which are not on the same row nor column.
- We can derive that

$$\det(A) = \sum_{\tau \in S_n} \mathrm{sgn}(\tau) a_{1\tau(1)} a_{2\tau(2)} \cdots a_{n\tau(n)}, \tag{1}$$

  where S_n is the set of all permutations of n elements $(1, 2, \cdots, n)$, and $\mathrm{sgn}(\tau) \in \{-1, 1\}$ is the sign function of the permutation τ.

If a permutation τ is odd, $\mathrm{sgn}(\tau) = -1$, otherwise 1.

To derive the Leibniz formula, follow the proof of Property 2,

$$\det(A) = \sum_{i=1}^{n} a_{i1} \begin{vmatrix} \mathbf{e}_i & \mathbf{a}_2 & \mathbf{a}_3 & \cdots & \mathbf{a}_n \end{vmatrix}.$$

We can continue this expansion for the second column, third column, ..., and obtain

$$\det(A) = \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \cdots \sum_{i_n=1}^{n} a_{i_1 1} a_{i_2 2} \cdots a_{i_n n} \begin{vmatrix} \mathbf{e}_{i_1} & \mathbf{e}_{i_2} & \cdots & \mathbf{e}_{i_n} \end{vmatrix}.$$

Noting that $\begin{bmatrix} \mathbf{e}_{i_1} & \mathbf{e}_{i_2} & \cdots & \mathbf{e}_{i_n} \end{bmatrix}$ is a permutation matrix if $i_1 \neq i_2 \neq \cdots \neq i_n$, and has determinant 0 if two columns are the same, we obtain the Leibniz formula.

5. 线性变换

Definition 17.1 **(Linear transformation)** Let V, W be two vector spaces, and the mapping L from V to W is said to be a linear transformation if the following condition is satisfied:

$$L(\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2) = \alpha_1 L(\mathbf{v}_1) + \alpha_2 L(\mathbf{v}_2), \quad \forall \alpha_1,\ \alpha_2 \in \mathbb{R}, \quad \forall \mathbf{v}_1,\ \mathbf{v}_2 \in V. (*)$$

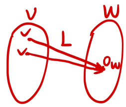
V is called the **domain** of the linear transformation, and W is called the **codomain** 陪域 of the linear transformation. In particular, if $V = W$, then L is a linear operator on V.

Remark
It can be easily checked that the above condition (*) is equivalent to the following two conditions:
(1) $L(\mathbf{u}+\mathbf{v}) = L(\mathbf{u}) + L(\mathbf{v}), \forall\ \mathbf{u}, \mathbf{v} \in V$
(2) $L(\alpha\mathbf{u}) = \alpha L(\mathbf{u}), \forall\ \mathbf{u} \in V, \alpha \in \mathbb{R}$

线性变换的基本性质.

Definition 17.8 **(Kernel 核 of the Linear transformation)** Let L be a linear transformation from V to W, then the kernel of L, denoted by $\ker(L)$ is defined as

$$\ker(L) = \{\mathbf{v} \in V | L(\mathbf{v}) = \mathbf{0}_W\}$$



Definition 17.9 **(Image and Range)** Let L be a linear transformation from V to W and let S be a subspace of V, the **image** of S, denoted by $L(S)$, is defined by

像的集合 $imf = \{L(v) \mid v \in S\}$.

$$L(S) = \{\mathbf{w} \in W | \exists\ \mathbf{v} \in S,\ s.t.\ L(\mathbf{v}) = \mathbf{w}\}$$

(unique)

The image of the entire vector space V, i.e., $L(V)$ is called the **range** of L.

Matrix representation of Linear Transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$

Theorem 17.16 **(Matrix Representation for linear transformation between Eulerian vector spaces w.r.t. standard bases)** If L is a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$, there is a $m \times n$ matrix A such that

$L(x) = L(e_1x_1 + e_2x_2 + \cdots + e_nx_n)$
$= L(e_1x_1) + L(e_2x_2) + \cdots + L(e_nx_n)$
$= L(e_1)x_1 + L(e_2)x_2 + \cdots + L(e_n)x_n$

$$L(\mathbf{x}) = A\mathbf{x} = a_1x_1 + a_2x_2 + \cdots + a_nx_n.$$

for each $\mathbf{x} \in \mathbb{R}^n$. In fact, the jth column vector of $A = [\mathbf{a}_1, \cdots, \mathbf{a}_n]$ is given by

$$\mathbf{a}_j = L(\mathbf{e}_j), \quad j = 1, 2, \cdots, n$$

where $\mathcal{E}_n = \{\mathbf{e}_1, \cdots, \mathbf{e}_n\}$ is the standard basis of $\mathbb{R}^n$.

基向量变换决定了整个变换

Remark 1. Linear transformation and its matrix representation is completely characterized by its action on a basis of its domain.
Remark 2. The matrix A in this theorem is the matrix representation of the linear transformation L w.r.t. standard bases of $\mathbb{R}^n$ and $\mathbb{R}^m$.

Theorem 18.1 (**Matrix Representation for General Vector Spaces**)
If $\mathcal{V} = \{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n\}$ is a basis for vector space V and $\mathcal{W} = \{\mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_m\}$ is a basis for vector space W, and L is a linear transformation mapping from vector space V to vector space W, then there is a $m \times n$ matrix A such that

$$[L(\mathbf{u})]_{\mathcal{W}} = A[\mathbf{u}]_{\mathcal{V}}, \quad \forall \mathbf{u} \in V$$

$[L(u)]_W = A[u]_V.$

And in fact, the jth column of A is given by

$$\mathbf{a}_j = [L(\mathbf{v}_j)]_{\mathcal{W}}$$

$a_j = [L(v_j)]_W.$

and $A = [\mathbf{a}_1, \cdots, \mathbf{a}_n]$.
Remark. Linear transformation is completely characterized by its action on a basis of its domain.

S

Theorem 18.4 (**Similarity Result in general vector space**) Let $E = \{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n\}$ and $F = \{\mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_n\}$ are two ordered bases for a vector space V, and L be a linear transformation from V to V. Let S be the transition matrix corresponding to the coordinate change from F to E. If A is the matrix representation of L w.r.t. E (taking E as the basis for both domain and co-domain) and B is the matrix representation of L w.r.t. F (taking F as the basis for both domain and co-domain), then

$[L]_E \triangleq A$ $\quad [L(x)]_E = A[x]_E$

$$B = S^{-1}AS$$

$[L]_F \triangleq B$ $\quad [L(x)]_F = B[x_F]$

$[x]_E = S[x_F]$

1、以F为底的L变换
2、转换基底 F→E.

$SB[x_F] = A[x_E] = AS[x_F]$

(以E为底的L变换)

Definition 18.5 (**Similar**) Let A and B are two $n \times n$ matrices, B is said to be similar to A if there exists a nonsingular matrix S such that $B = S^{-1}AS$.

$\Rightarrow SB = AS$
$B = S^{-1}AS.$

slide 18 p.12-16.

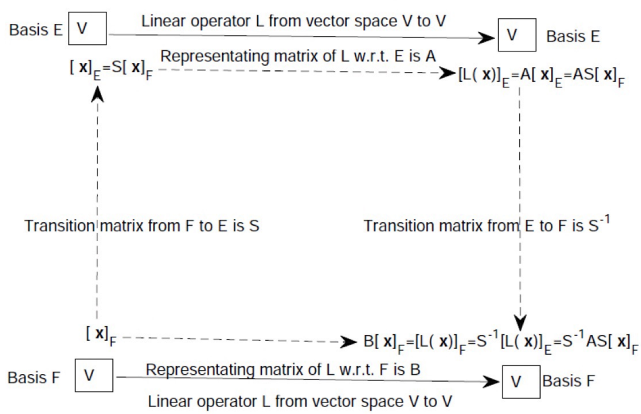


Figure: Illustration for the coordinates relation using two different bases and matrices representation of the linear operator L from V to V.

正交

Theorem 19.5 (**Scalar Product in terms of Vector Length**) Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, suppose θ is the angle between two nonzero vectors, then

$$\mathbf{x}^T\mathbf{y} = \| \mathbf{x} \| \| \mathbf{y} \| \cos\theta, \quad 0 \leq \theta \leq \pi.$$

Corollary 19.6 (**Cauchy-Schwartz Inequality**)

$$|\mathbf{x}^T\mathbf{y}| \leq \| \mathbf{x} \| \| \mathbf{y} \|$$

Definition 19.7 (**Orthogonal Vectors in $\mathbb{R}^n$**) Two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ are said to be **orthogonal** if $\mathbf{x}^T\mathbf{y} = 0$. Denote $\mathbf{x} \perp \mathbf{y}$.

Definition 19.11 (**Orthogonal Subspaces in $\mathbb{R}^n$**) Two subspaces X and Y of $\mathbb{R}^n$ are said to be orthogonal if

$$\mathbf{x}^T\mathbf{y} = 0, \ \forall \, \mathbf{x} \in X, \ \mathbf{y} \in Y.$$

Denoted by $X \perp Y$.

Theorem 20.2 (**Direct sum of $\mathbb{R}^n$**) If S is a subspace of $\mathbb{R}^n$, then

$$\mathbb{R}^n = S \oplus S^{\perp}$$

suppose $\alpha_1 u_1 + \cdots + \alpha_n u_n = 0$

$\underbrace{\alpha_1 u_1 + \cdots + \alpha_r u_r}_{\in S} = \underbrace{-\alpha_{r+1} u_{r+1} - \cdots - \alpha_n u_n}_{\in S^{\perp}}$

Theorem 19.19 If S is a subspace of $\mathbb{R}^n$, then

可能同时含有某一维

$$\dim S + \dim S^{\perp} = n.$$

e.g. $S = \mathrm{span}\left(\begin{bmatrix}1\\1\\0\end{bmatrix}\right)$

$S^{\perp} = \mathrm{span}\left(\begin{bmatrix}1\\-1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\end{bmatrix}\right)$

$\because S \cap S^{\perp} = \{0\}$ $\therefore$ LHS = RHS = 0

$\therefore \alpha_1 = \cdots = \alpha_n = 0$.

Furthermore, if $\{\mathbf{u}_1, \cdots, \mathbf{u}_r\}$ is a basis for S and $\{\mathbf{u}_{r+1}, \cdots, \mathbf{u}_n\}$ is a basis for $S^{\perp}$, then $\{\mathbf{u}_1, \cdots, \mathbf{u}_r, \mathbf{u}_{r+1}, \cdots, \mathbf{u}_n\}$ is a basis for $\mathbb{R}^n$. $\Leftrightarrow$ $\{u\}$ are linearly independent.

Definition 20.10 (**Idempotent**) Let A be a square matrix that satisfies $A = A^2$, then A is called an idempotent matrix.

Theorem 19.17 (**Fundamental Subspaces Theorem**)

Let $A \in \mathbb{R}^{m \times n} = [\mathbf{a}_1, \cdots, \mathbf{a}_n] = \begin{bmatrix} \vec{\mathbf{a}}_1 \\ \vec{\mathbf{a}}_2 \\ \vdots \\ \vec{\mathbf{a}}_m \end{bmatrix}$, $Col(A) = \mathrm{Span}\{\mathbf{a}_1, \cdots, \mathbf{a}_n\}$ is the column space of A, $Row(A) = Col(A^T) = \mathrm{Span}\{\vec{\mathbf{a}}_1^T, \vec{\mathbf{a}}_2^T, \cdots, \vec{\mathbf{a}}_m^T\}$ is the row space of A, then

(1) $\mathrm{Null}(A) = Col(A^T)^{\perp} = Row(A)^{\perp}$

(2) $\mathrm{Null}(A^T) = Col(A)^{\perp} = Row(A^T)^{\perp}$

proof: $\mathrm{Null}(A) = \{x \mid \vec{a}_i x = 0, \ \forall \, i = 1, 2, \cdots, n\}$

$= \{x \mid \alpha_i \vec{a}_i x = 0, \ \alpha_i \in R, \ i = 1, 2, \cdots, n\}$

$= \{x \mid y^T x = 0, \ y \in Row(A)\}$

$= Row(A)^{\perp}$

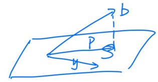
$= Col(A^T)^{\perp}$

最小二乘解 (least square solution)

Definition 20.6 (**Least square solution**) Given linear system $A\mathbf{x} = \mathbf{b}(A \in \mathbb{R}^{m\times n}, \mathbf{b} \in \mathbb{R}^m)$, a vector $\hat{\mathbf{x}}(\mathbf{x} \in \mathbb{R}^n)$ that satisfies the minimum residual condition

$$\| r(\hat{\mathbf{x}}) \| = \min_{\mathbf{x}} \| r(\mathbf{x}) \|$$

is called the least square solution for $A\mathbf{x} = \mathbf{b}$.



For any b, inside the subspace S.
$\exists\, p$. s.t. ~~① $(b-p)\cdot p=0$~~ $(b-p \in S^{\perp})$.
② $\|b-p\| \geq \|b-y\|$.

proof ①: $\mathbb{R}^n = S \oplus S^{\perp}$
$b = p + z$, $\because p \in S$ $\therefore z \in S^{\perp}$
p is the optimal sol.

if $b-p \in S^{\perp}$.
$\|b-y\|^2 = \|b-p+p-y\|^2$
$= \|b-p\|^2 + \|p-y\|^2 \geq \|b-p\|^2$.

$Ax = b$.

求 $Ax = p \in Col(A)$ $\Leftrightarrow$ $b-p = b-Ax$.
$b - Ax = b - p \in Col(A)^{\perp} = Null(A^T)$
$\therefore A^T\cdot(b-p) = 0$
$A^T\cdot(b - A\hat{x}) = 0$
$\therefore A^TA\hat{x} = A^Tb$.

核心思想.

Theorem 20.8 (**Normal equations for the linear system**) Given the linear system $A\mathbf{x} = \mathbf{b}$ $(A \in \mathbb{R}^{m\times n}, \mathbf{b} \in \mathbb{R}^m)$, let the projection of $\mathbf{b}$ onto the subspace $Col(A)$ is $\mathbf{p}$, then there exists a vector $\hat{\mathbf{x}} \in \mathbb{R}^n$, s.t.
$\mathbf{p} = A\hat{\mathbf{x}} \in Col(A), \mathbf{b} - A\hat{\mathbf{x}} \in Col(A)^{\perp} = Null(A^T)$ and
$\| \mathbf{b} - A\mathbf{x} \| \geq \| \mathbf{b} - A\hat{\mathbf{x}} \|$ for any $\mathbf{x} \in \mathbb{R}^n$. 由正交性判断最优解.
$\mathbf{b} - A\hat{\mathbf{x}} \in Col(A)^{\perp} = Null(A^T)$ gives the condition

$$A^T(\mathbf{b} - A\hat{\mathbf{x}}) = 0$$

i.e.,

$$A^TA\hat{\mathbf{x}} = A^T\mathbf{b}$$

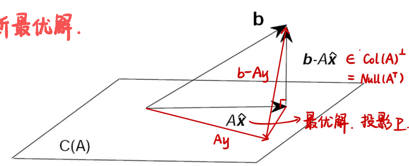


Figure: Projection of $\mathbf{b} \in V$ onto column space $Col(A)$.

which is called the **normal equation**, and it is a $n \times n$ linear system.

Remark
The normal equations may not have a unique solution, but the **projection vector p** of $\mathbf{b}$ onto $Col(A)$ is unique, i.e., there are possible two vectors $\hat{\mathbf{x}}, \hat{\mathbf{y}}$ satisfies $A\hat{\mathbf{x}} = A\hat{\mathbf{y}} = \mathbf{p}$.

Theorem 20.9 (**Unique Solution Condition for the Normal Equations**) If A is a $m \times n$ matrix of rank n, the normal equations

$$A^TA\hat{\mathbf{x}} = A^T\mathbf{b}$$

$\Leftrightarrow$ A^{-1} exists.

have a unique solution

$$\hat{\mathbf{x}} = (A^TA)^{-1}A^T\mathbf{b}$$

and $\hat{\mathbf{x}}$ is the unique least square solution for the linear system $A\hat{\mathbf{x}} = \mathbf{b}$.

本质: A is nonsingular $\Rightarrow$ A^TA is nonsigular.
Assume $A^TAz = 0$,
$z^TA^TAz = 0$
$\Rightarrow \|Az\| = 0$
$\because A$ is nonsigular
$\therefore z = 0$.
$\therefore A^TA$ is nonsigular.

内积空间

Definition 21.1 (**Inner Product Space over Real Number Field**)
Let V be a vector space, an **inner product** is an operation on V which assigns a real number $\langle \mathbf{x}, \mathbf{y} \rangle$ for each pair of vectors $\mathbf{x}, \mathbf{y} \in V$. The operation $\langle \cdot, \cdot \rangle$ satisfies:

(1) $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$ with equality if and only if $\mathbf{x} = \mathbf{0}$.

内积的基本性质.

(2) $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle, \ \forall \ \mathbf{x}, \ \mathbf{y} \in V$.

(3) $\langle \alpha\mathbf{x} + \beta\mathbf{y}, \mathbf{z} \rangle = \alpha\langle \mathbf{x}, \mathbf{z} \rangle + \beta\langle \mathbf{y}, \mathbf{z} \rangle, \ \forall \ \mathbf{x}, \mathbf{y}, \mathbf{z} \in V$, and $\alpha, \ \beta \in \mathbb{R}$.

If the vector space V has an inner product operation on V, then V is called the inner product space.

定义 (内积)： V上的内积就是一个函数，它把V中元素的每一个有序对 (u, v) 都映成一个数 $< u, v > \in F$ ，并且满足：正定性（1和2），线性（3和4），共轭对称性（5，在 $F = R$ 的时候就是交换律）

内积空间完备性质.

1. 正性：对所有 $v \in V$ 均有 $< v, v > \geq 0$
2. 定性： $< v, v > = 0$ 当且仅当 $v = 0$
3. 加性：对所有 $u, v, w \in V$ 均有 $< u + v, w > = < u, w > + < v, w >$
4. 齐性：对所有 $\lambda \in F$ 和所有 $u, v \in V$ 均有 $< \lambda u, v > = \lambda < u, v >$
5. 共轭对称性：对所有 $u, v \in V$ 均有 $< u, v > = < \bar{v, u} >$

定义（正交）： 两个向量 u, v 正交若 $< u, v > = 0$

正交用非常简洁的数学语言描述了 **"不相干"** 这种状态，是几何中**垂直**这个概念在更高维度上的推广。

赋范向量空间

Definition 21.6 (**Normed Vector Space**) A vector space V is said to be a normed linear space if, each vector $\mathbf{v} \in V$ is associated with a real number $\| \mathbf{v} \| \in \mathbb{R}$, called the **norm** of $\mathbf{v}$, satisfying:

范数的基本性质.

(I) $\| \mathbf{v} \| \geq 0$ with equality if and only if $\mathbf{v} = 0$.
(II) $\| \alpha\mathbf{v} \| = |\alpha| \ \| \mathbf{v} \|$ for any scalar α.
(III) $\| \mathbf{u} + \mathbf{v} \| \leq \| \mathbf{u} \| + \| \mathbf{v} \|$ for any $\mathbf{u}, \mathbf{v} \in V$ (triangle inequality).

默认将范数定义为 $\|v\|_2 = \sqrt{<v, v>}$

Inner product defined on $\mathbb{R}^{m \times n}$ (**Frobenius inner product**)
Given $A, B \in \mathbb{R}^{m \times n}$, we can define an inner product as

$$\langle A, B \rangle = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} b_{ij}$$

弗罗贝尼乌斯内积.

The vector space $C[a, b]$. For $f, g \in C[a, b]$, the inner product on $C[a, b]$ is defined by

$$\langle f, g \rangle = \int_a^b f(x) g(x) dx$$

傅立叶内积

正交矩阵.

Definition 21.8 (**Orthogonal Set**) Let $\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n\}$ be nonzero vectors in an inner product space V. If $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0$ when $i \neq j$, then $\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n\}$ is said to be an **orthogonal set** of vectors.

orthogonal matrix.

Definition 21.9 (**Orthonormal Set**) An **orthonormal set** of vectors is an orthogonal set of **unit** vectors, where the **unit** vector means the norm of the vector is 1.

Theorem 21.10 (**Orthogonal set are linearly independent**) Let $\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n\}$ be the set of orthogonal vectors in an inner product space V, then $\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n\}$ is linearly independent.

Assume $V_n = C_1V_1 + \dots + C_rV_r$

$V_n V_n = (C_1V_1 + \dots + C_rV_r) V_n = 0$

矛盾！

Remark: If $B = \{\mathbf{u}_1, \cdots, \mathbf{u}_m\}$ is an orthogonal set and is the basis for the inner product vector space V, then for any $\mathbf{v} \in V$, one has:

$$\mathbf{v} = \sum_{i=1}^{m} \frac{< \mathbf{v}, \mathbf{u}_i >}{\| \mathbf{u}_i \|^2} \mathbf{u}_i$$

向量的正交基分解.

Definition 21.15 (**Orthogonal Matrix**)
Let $Q \in \mathbb{R}^{n \times n}$, Q is said to be the orthogonal matrix if the column vectors of Q is an orthonormal set in $\mathbb{R}^n$.

Theorem 21.16 (**Equivalent Condition for Orthogonal Matrix**) An $n \times n$ matrix Q is orthogonal matrix if and only if $Q^{-1} = Q^T$.

矩阵正交的等价条件

Recall: For square matrices $A, B \in \mathbb{R}^{n \times n}$, $AB = I_n$ implies $BA = I_n$.
Proof. Since Q is a square matrix, $Q^{-1} = Q^T$ is equivalent to $Q^TQ = I_n$. Thus, only need to show that Q is orthogonal matrix if and only if $Q^TQ = I_n$. Let $Q = [\mathbf{q}_1, \cdots, \mathbf{q}_n]$, then $Q^T = \begin{bmatrix} \mathbf{q}_1^T \\ \vdots \\ \mathbf{q}_n^T \end{bmatrix}$.

$Q = [\mathbf{q}_1, \cdots, \mathbf{q}_n]$ is an orthogonal matrix
$\Leftrightarrow \{\mathbf{q}_1, \cdots, \mathbf{q}_n\}$ is an orthonormal set in $\mathbb{R}^n$
$\Leftrightarrow \{\mathbf{q}_1, \cdots, \mathbf{q}_n\}$ is an orthonormal basis for $\mathbb{R}^n$
$\Leftrightarrow (\mathbf{q}_i^T\mathbf{q}_j)_{n \times n} = Q^TQ = (\delta_{ij})_{n \times n} = I_n$.

运用 inner product 的正交性.

$$\delta_{ij} = \begin{cases} 1, & i=j \\ 0, & i \neq j. \end{cases}$$

Property 21.18 (**Properties for orthogonal matrix**) If Q is an $n \times n$ orthogonal matrix, then

(a) the column vectors of Q form an orthonormal basis for $\mathbb{R}^n$.

(b) $Q^{-1} = Q^T$

(c) $Q^TQ = I_n$

(d) $\langle Q\mathbf{x}, Q\mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle$,
$\langle Q\mathbf{x}, Q\mathbf{y} \rangle = (Q\mathbf{y})^T Q\mathbf{x} = \mathbf{y}^T Q^T Q\mathbf{x} = \mathbf{y}^T\mathbf{x} = \langle \mathbf{x}, \mathbf{y} \rangle$

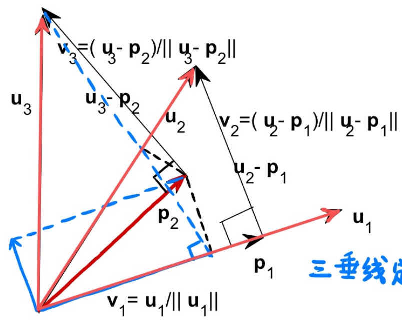
(e) $\| Q\mathbf{x} \| = \| \mathbf{x} \|$ $(Q\mathbf{x})^T Q\mathbf{x} = \mathbf{x}^T Q^T Q\mathbf{x} = \mathbf{x}^T\mathbf{x}$

☆ 格拉姆-施密特正交化 Gram-Schmidt process

将向量空间的一组基转化为标准正交基.

Lemma 22.1 (**Projection onto a subspace**) Let S be a subspace of the inner product space V and $\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_m\}$ is the orthonormal basis for S, for any $\mathbf{x} \in V$. And let $\mathbf{p}$ be the projection vector of $\mathbf{x}$ onto S $(\mathbf{x} - \mathbf{p} \perp S)$, then $\mathbf{p}$ is uniquely determined by

$$\mathbf{p} = \sum_{i=1}^{m} < \mathbf{x}, \mathbf{u}_i > \mathbf{u}_i$$



三垂线定理.

step 1: $\boxed{\mathbf{v}_1} = \mathbf{u}_1 / \|\mathbf{u}_1\|$

step 2: $\mathbf{p}_1 = < \mathbf{u}_2, \mathbf{v}_1 > \mathbf{v}_1$ S_1平面投影

$\boxed{\mathbf{v}_2} = (\mathbf{u}_2 - \mathbf{p}_1) / \|\mathbf{u}_2 - \mathbf{p}_1\|$

step 3: $\mathbf{p}_2 = < \mathbf{u}_3, \mathbf{v}_1 > \mathbf{v}_1 + < \mathbf{u}_3, \mathbf{v}_2 > \mathbf{v}_2$ S_2平面投影.

$\boxed{\mathbf{v}_3} = (\mathbf{u}_3 - \mathbf{p}_2) / \|\mathbf{u}_3 - \mathbf{p}_2\|$

求投影 — 得正交向量 — 扩充平面 (subspace)

Theorem 22.5 (**QR decomposition**) Let $A \in \mathbb{R}^{m \times n}$ ($m > n$) be an $m \times n$ real matrix and rank$(A) = n$ (column vectors are linearly independent), then A can be factorized as $A = QR$, where Q is an $m \times n$ matrix with orthonormal column vectors and R is an upper triangular $n \times n$ matrix with all positive diagonal elements.

QR decomposition

记录施密特正交化的过程

A　　　Q　　　R

$$A = [\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_n] = [\mathbf{q}_1, \mathbf{q}_2, \cdots, \mathbf{q}_n] \begin{bmatrix} \|\mathbf{a}_1\| & <\mathbf{q}_1, \mathbf{a}_2> & \ldots & <\mathbf{q}_1, \mathbf{a}_n> \\ 0 & \|\mathbf{r}_1\| & \ldots & <\mathbf{q}_2, \mathbf{a}_n> \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \|\mathbf{r}_{n-1}\| \end{bmatrix}$$

linearly independent set　　orthogonormal set

$$\triangleq QR$$

Here $Q = [\mathbf{q}_1, \mathbf{q}_2, \cdots, \mathbf{q}_n]$, and

$$R = \begin{bmatrix} \mathbf{q}_1^T\mathbf{a}_1 & \mathbf{q}_1^T\mathbf{a}_2 & \ldots & \mathbf{q}_1^T\mathbf{a}_n \\ 0 & \mathbf{q}_2^T\mathbf{a}_2 & \ldots & \mathbf{q}_2^T\mathbf{a}_n \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \mathbf{q}_n^T\mathbf{a}_n \end{bmatrix}$$

since $\mathbf{q}_1^T\mathbf{a}_1 =< \mathbf{a}_1, \mathbf{q}_1 >= \|\mathbf{a}_1\|$, $\quad \mathbf{q}_j^T\mathbf{a}_j =< \mathbf{a}_j, \mathbf{q}_j >=\| \mathbf{r}_{j-1} \|> 0 \quad (j = 2, \cdots, n)$.

In fact, from $A = QR$, one has

$$R = Q^TA = \begin{bmatrix} \mathbf{q}_1^T\mathbf{a}_1 & \mathbf{q}_1^T\mathbf{a}_2 & \ldots & \mathbf{q}_1^T\mathbf{a}_n \\ \mathbf{q}_2^T\mathbf{a}_1 & \mathbf{q}_2^T\mathbf{a}_2 & \ldots & \mathbf{q}_2^T\mathbf{a}_n \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{q}_n^T\mathbf{a}_1 & \mathbf{q}_n^T\mathbf{a}_2 & \ldots & \mathbf{q}_n^T\mathbf{a}_n \end{bmatrix}$$

Use $A = QR$ to solve linear system.

$Ax = b$, $\quad A^T = R^TQ^T$

$A^TA = R^TQ^TQR = R^TR$

$$Q^TQ = \begin{bmatrix} \vec{q}_1 \\ \vec{q}_2 \\ \vdots \\ \vec{q}_n \end{bmatrix} [q_1\ q_2 \cdots q_n] = I$$

$(|\vec{q}_i|^2 = 1,\ \vec{q}_i \cdot \vec{q}_j = 0,\ i \neq j)$

$$= \begin{bmatrix} \mathbf{q}_1^T\mathbf{a}_1 & \mathbf{q}_1^T\mathbf{a}_2 & \ldots & \mathbf{q}_1^T\mathbf{a}_n \\ 0 & \mathbf{q}_2^T\mathbf{a}_2 & \ldots & \mathbf{q}_2^T\mathbf{a}_n \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \mathbf{q}_n^T\mathbf{a}_n \end{bmatrix}$$

where

$$\mathbf{q}_j \perp \mathrm{Span}(\mathbf{q}_1, \cdots, \mathbf{q}_{j-1}) = \mathrm{Span}(\mathbf{a}_1, \cdots, \mathbf{a}_{j-1}), (j = 2, \cdots, n)$$

are used.

Question 1

Matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ is non-singular. For any vector $\boldsymbol{x} \in \mathbb{R}^n$, define

$$\|\boldsymbol{x}\|_{\boldsymbol{A}} = \|\boldsymbol{A}\boldsymbol{x}\|_2$$

show that $\|\boldsymbol{x}\|_{\boldsymbol{A}}$ defines a norm in $\mathbb{R}^n$.

范数定义的三个条件：

① $\|x\|_A \geqslant 0 \iff \|Ax\|_2 \geqslant 0$ 当且仅当 $x=0$ 时，$\|Ax\|=0$.

$\iff A$ is non-singular.

② $\|\alpha x\|_A = \|\alpha Ax\|_2 = \alpha\|Ax\|_2 = \alpha\|x\|_A$.

③ $\|u\|_A + \|v\|_A \geqslant \|u+v\|_A$

$\iff \|Au\|_2 + \|Av\|_2 \geqslant \|A(u+v)\|_2$ 成立.

Question 3

Find the equation of the circle that gives the best least squares circle fit to the points $(-1,-2), (0,2.4), (1.1,-4)$ and $(2.4,-1.6)$.

$$\begin{cases} (x+1)^2 + (y+2)^2 = r^2 \\ x^2 + (y-2.4)^2 = r^2 \\ (x-1.1)^2 + (y+4)^2 = r^2 \\ (x-2.4)^2 + (y+1.6)^2 = r^2 \end{cases}$$

$(x-a)^2 + (y-b)^2 = r^2$

$2ax + 2by + (r^2 - x^2 - y^2) = a^2 + b^2$.

$$\begin{bmatrix} -1 & -2 & 1 \\ 0 & 2.4 & 1 \\ 1.1 & -4 & 1 \\ 2.4 & -1.6 & 1 \end{bmatrix} \begin{bmatrix} 2x \\ 2y \\ r^2 - x^2 - y^2 \end{bmatrix} = \begin{bmatrix} a_1^2 + b_1^2 \\ a_2^2 + b_2^2 \\ a_3^2 + b_3^2 \\ a_4^2 + b_4^2 \end{bmatrix}$$

$A\underset{\smile}{x} = b$

所有未知数集中在一个 vector.

Solve $\hat{\boldsymbol{x}} = \left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T\mathbf{b}$, we can get

$$(x - 0.5755)^2 + (y + 0.6426)^2 = 7.4265$$

Question 4

Let $\boldsymbol{x}$ and $\boldsymbol{y}$ be linearly independent vectors in $\mathbb{R}^n$ and let $\mathcal{S} = \text{Span}\{\boldsymbol{x}, \boldsymbol{y}\}$. Construct a matrix by $\boldsymbol{A} = \boldsymbol{x}\boldsymbol{y}^T + \boldsymbol{y}\boldsymbol{x}^T$.

(a) Show that $\boldsymbol{A}$ is symmetric.

(b) Show that $Null(\boldsymbol{A}) = \mathcal{S}^{\perp}$.

(a) $A^T = (xy^T + yx^T)^T = yx^T + xy^T = A$

(b) $Az = 0$

$(xy^T + yx^T)z = 0$

$x(y^Tz) + y(x^Tz) = 0$ $\quad$ y^Tz, x^Tz is scalar.

$\because$ x, y are linearly independent.

$\therefore$ $y^Tz = x^Tz = 0$. $\quad$ $z \in \text{span}\{x, y\}^{\perp}$

Question 3

Compute the Gram-Schimdt QR factorization of the matrix.

$$A = \begin{array}{c} \begin{array}{ccc} a_1 & a_2 & a_3 \end{array} \\ \begin{bmatrix} 1 & -2 & -1 \\ 2 & 0 & 1 \\ 2 & -4 & 2 \\ 4 & 0 & 0 \end{bmatrix} \end{array}$$

$$q_1 = \begin{bmatrix} \frac{1}{5} \\ \frac{2}{5} \\ \frac{2}{5} \\ \frac{4}{5} \end{bmatrix}$$

$$r_1 = a_2 - \underbrace{\langle a_2, q_1 \rangle}_{=-2} q_1 = a_2 + 2q_1 = \begin{bmatrix} -\frac{8}{5} \\ \frac{4}{5} \\ -\frac{16}{5} \\ \frac{8}{5} \end{bmatrix}$$

$$q_2 = \frac{r_1}{\|r_1\|} = \begin{bmatrix} -\frac{2}{5} \\ \frac{1}{5} \\ -\frac{4}{5} \\ \frac{2}{5} \end{bmatrix}$$

(算比例,凑1)

$$r_2 = a_3 - \underbrace{\langle a_3, q_1 \rangle}_{=1} q_1 - \underbrace{\langle a_3, q_2 \rangle}_{=-1} q_2 = \begin{bmatrix} -\frac{8}{5} \\ \frac{4}{5} \\ \frac{4}{5} \\ -\frac{2}{5} \end{bmatrix}$$

$$q_3 = \begin{bmatrix} -\frac{4}{5} \\ \frac{2}{5} \\ \frac{2}{5} \\ -\frac{1}{5} \end{bmatrix}$$

$$A = QR = \begin{bmatrix} \frac{1}{5} & -\frac{2}{5} & -\frac{4}{5} \\ \frac{2}{5} & \frac{1}{5} & \frac{2}{5} \\ \frac{2}{5} & -\frac{4}{5} & \frac{2}{5} \\ \frac{4}{5} & \frac{2}{5} & -\frac{1}{5} \end{bmatrix} \begin{bmatrix} 5 & -2 & 1 \\ 0 & 4 & -1 \\ 0 & 0 & 2 \end{bmatrix}$$

Question 4

Find the ~~least squares~~ solution of $A\mathbf{x} = \mathbf{b}$ using QR factorization, where

$$A = \begin{bmatrix} 1 & -2 & -1 \\ 2 & 0 & 1 \\ 2 & -4 & 2 \\ 4 & 0 & 0 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} -1 \\ 1 \\ 1 \\ -2 \end{bmatrix}$$

$$Ax = b, \quad QRx = b$$

$$Rx = Q^T b = \begin{bmatrix} \frac{1}{5} & \frac{2}{5} & \frac{2}{5} & \frac{4}{5} \\ -\frac{2}{5} & \frac{1}{5} & -\frac{4}{5} & \frac{2}{5} \\ -\frac{4}{5} & \frac{2}{5} & \frac{2}{5} & -\frac{1}{5} \end{bmatrix} \begin{bmatrix} -1 \\ 1 \\ 1 \\ -2 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}$$

$$\left[\begin{array}{ccc|c} 5 & -2 & 1 & -1 \\ 0 & 4 & -1 & -1 \\ 0 & 0 & 2 & 2 \end{array}\right] \Rightarrow x = \begin{bmatrix} -\frac{2}{5} \\ 0 \\ 1 \end{bmatrix}$$

特征值和特征向量.

Definition 23.1 (**Eigenvalue and eigenvectors**) Let A be a square matrix with size $n \times n$ ($A \in \mathbb{R}^{n\times n}$ or $A \in \mathbb{C}^{n\times n}$), if there exists a scalar λ ($\lambda \in \mathbb{R}$ or $\lambda \in \mathbb{C}$) and nonzero vector $\mathbf{x}$ such that $A\mathbf{x} = \lambda\mathbf{x}$, then λ is called the **eigenvalue** (or **characteristic value**) and $\mathbf{x}$ is called the **eigenvector** (or **characteristic vector**) w.r.t λ. (λ, x) is a eigen pair

Theorem 23.3 Let Let A be a square matrix with size $n \times n$ ($A \in \mathbb{R}^{n\times n}$ or $A \in \mathbb{C}^{n\times n}$) and λ ($\lambda \in \mathbb{R}$ or $\lambda \in \mathbb{C}$), then the following statements are equivalent:

(a) λ is an eigenvalue of A.

(b) $(A - \lambda I)\mathbf{x} = \mathbf{0}$ has nontrivial solutions. $Ax = \lambda x$

(c) $\mathrm{Null}(A - \lambda I) \neq \{\mathbf{0}\}$, where $\mathrm{Null}(A - \lambda I)$ is a subspace of $\mathbb{R}^n$ when $\lambda \in \mathbb{R}$ and $\mathrm{Null}(A - \lambda I)$ is a subspace of $\mathbb{C}^n$ when $\lambda \in \mathbb{C}$. $\mathrm{Null}(A - \lambda I)$ is called the **eigenspace** corresponding to λ.

(d) $A - \lambda I$ is singular.

(e) $\det(A - \lambda I) = 0$. 此性质可以用来计算特征值 λ.

Proof. Using the definition of Null space, determinant, matrix singular. It is easy to show that: λ is an eigenvalue of $A \Leftrightarrow (A - \lambda I)\mathbf{x} = \mathbf{0}$ has a nontrivial solution$\Leftrightarrow \mathrm{Null}(A - \lambda I) \neq \{\mathbf{0}\} \Leftrightarrow A - \lambda I$ is singular$\Leftrightarrow$ $\det(A - \lambda I) = 0$.

特征多项式

Definition 23.4 (**Characteristic Polynomial**) Let A is a $n \times n$ matrix ($A \in \mathbb{R}^{n\times n}$ or $A \in \mathbb{C}^{n\times n}$) and λ is a variable, then

$$p_A(\lambda) = \det(A - \lambda I) = \begin{vmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{vmatrix}$$

is called the characteristic polynomial of A, and

$$p_A(\lambda) = 0$$

is called the characteristic equation of A.
特征方程

代数基本定理

Theorem 23. 9 (**Fundamental theorem in Algebra**) Every degree n polynomial with complex coefficients has exactly n complex roots. (Counting with multiplicity).

Theorem 23.10 (**Product and Sum of Eigenvalues**) Let $A = (a_{ij})_{n\times n}$ be a square matrix ($A \in \mathbb{R}^{n\times n}$ or $A \in \mathbb{C}^{n\times n}$), $\lambda_i (i = 1, 2, \cdots, n)$ are the eigenvalues, then

(1) $\det(A) = \prod_{i=1}^{n} \lambda_i$

(2) $\sum_{i=1}^{n} a_{ii} = \sum_{i=1}^{n} \lambda_i$, where $\sum_{i=1}^{n} a_{ii} = a_{11} + a_{22} + \cdots + a_{nn}$ is called the trace of A. Denoted as $\mathrm{Trace}(A) = \sum_{i=1}^{n} a_{ii}$.

$$\det(A-\lambda I) = \begin{vmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & & \vdots \\ \vdots & & \ddots & \\ a_{n1} & \cdots & & a_{nn}-\lambda \end{vmatrix}$$

展开

产生λ^n和λ^{n-1}.

$$P_A(\lambda) = (a_{11}-\lambda)\begin{vmatrix} a_{22}-\lambda & & \\ & \ddots & \\ & & a_{nn}-\lambda \end{vmatrix} - a_{21}\begin{vmatrix} \cdots \\ \cdots \end{vmatrix}$$

行列式展开 观察 λ^n 和 λ^{n-1} 前的系数

由观察知，λ^n 和 λ^{n-1} 由 $\prod_{i=1}^{n}(a_{ii}-\lambda)$ 产生.
(行列式展开一般式)

由代数基本定理，$P_A(\lambda) = (\underbrace{\lambda_1}_{可求}-\lambda)(\lambda_2-\lambda)\cdots(\lambda_n-\lambda)$ (根式)

$\det(A) = P_A(0) = \prod_{i=1}^{n}\lambda_i$

比较λ^{n-1}前的系数：$(-1)^{n-1}\sum_{i=1}^{n} a_{ii} = (-1)^{n-1}\sum_{i=1}^{n}\lambda_i$

$\Rightarrow \sum_{i=1}^{n} a_{ii} = \sum_{i=1}^{n}\lambda_i$.

Remark

1. A is nonsingular $\Leftrightarrow \det(A) \neq 0 \Leftrightarrow$ all eigenvalues $\lambda_i \neq 0$

2. If A is nonsingular and λ is the eigenvalue of A, then λ^{-1} is also the eigenvalue of A^{-1} since $A\mathbf{x} = \lambda\mathbf{x}$ implies that $A^{-1}\mathbf{x} = \frac{1}{\lambda}\mathbf{x}$ if A is nonsingular.
逆变换

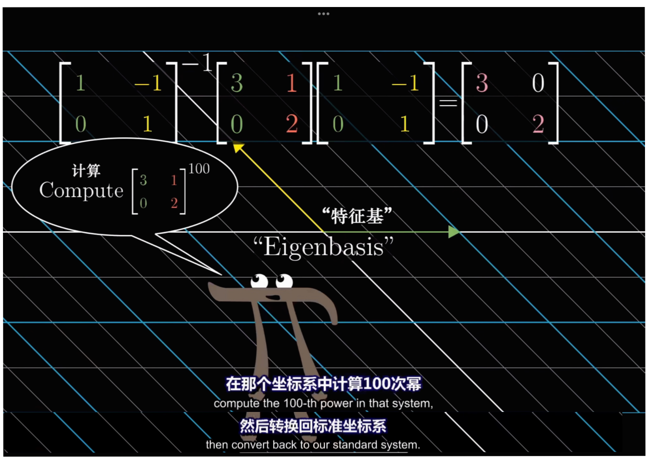
Theorem 23.12 (**Similar Matrices Have the Same Eigenvalues**) Let A, B are both $n \times n$ real matrices, if A B are similar, then two matrices have the same characteristic polynomial, and hence have the same eigenvalues.

☆ 对角化

Definition 24.1 (**Diagonalizable**) A $n \times n$ matrix A is said to be diagonalizable if there exists a nonsingular matrix X and a diagonal matrix D such that

$$X^{-1}AX = D$$

Thus, $A = XDX^{-1}$, and X diagonalizes A.



Theorem 24.3 (**Sufficient and Necessary Condition for Diagonalization**) A $n \times n$ matrix A is diagonalizable if and only if A has n linearly independent eigenvectors. (可能对应相同λ)

Proof. $\Leftarrow$ Suppose that A has n linearly independent eigenvectors $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n$, then we have 对角化的充要条件.

为了保证 X^{-1} 存在.

$$A\mathbf{x}_i = \lambda_i \mathbf{x}_i$$

where the associated eigenvalue for the eigenvector $\mathbf{x}_i$ is λ_i $(i = 1, \cdots, n)$. Thus

$$\begin{aligned} A[\underbrace{\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n}_{\text{eigenvectors}}] &\triangleq X \\ =&[\lambda_1 \mathbf{x}_1, \lambda_2 \mathbf{x}_2, \cdots, \lambda_n \mathbf{x}_n] \\ =&[\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n]\mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n) \end{aligned}$$

Let $X = [\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n]$ (= eigenvectors), then X is nonsingular since $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n$ are linearly independent. Thus $X^{-1}AX = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$. Thus A is diagonalizable.

$\Rightarrow$ Suppose that A is diagonalizable, there exists a nonsingular matrix $X = [\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n]$, s.t. $X^{-1}AX = D = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$, then

$$A[\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n] = [\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n]\mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$$

Thus

$$A\mathbf{x}_i = \lambda_i \mathbf{x}_i, \quad i = 1, \cdots, n.$$

Thus, $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n$ are eigenvectors. Moreover, since X is nonsingular, thus $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n$ is linearly independent.

Remark. Square matrices that is not diagonalizable is called **defective** matrices.

Theorem 24.7 Let $A \in \mathbb{R}^{n \times n}$ be the real symmetric matrix with eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_n$, then
(1) $\lambda_i \in \mathbb{R}$, $\forall\, i = 1, \cdots, n$. (**The eigenvalues of real symmetric matrices are real numbers.**)
(2) If $\lambda_i \neq \lambda_j$, $\mathbf{x}_i$ is the eigenvectors w.r.t λ_i, $\mathbf{x}_j$ is the is the eigenvectors w.r.t λ_j, then $\mathbf{x}_i$, $\mathbf{x}_j$ are orthogonal.(**For real symmetric matrices, the eigenvectors belonging to different eigenvalues are orthogonal.**)

$A = Q\Lambda Q^T$

Suppose the symmetric matric $A \in \mathbb{R}^{n \times n}$ and $Q^{-1}AQ = \Lambda$ (Λ is a diagonal matrix), where Q is an orthogonal matrix and Λ is the diagonal matrix($\Lambda = diag(\lambda_1, \cdots, \lambda_n)$, $\lambda_1, \cdots, \lambda_n$ are real numbers), then

$$A = Q\Lambda Q^T = [\mathbf{q}_1, \cdots, \mathbf{q}_n] diag(\lambda_1, \cdots, \lambda_n) \begin{bmatrix} \mathbf{q}_1^T \\ \vdots \\ \mathbf{q}_n^T \end{bmatrix}$$

Thus cross product.

$$A = \lambda_1 \mathbf{q}_1 \mathbf{q}_1^T + \cdots + \lambda_n \mathbf{q}_n \mathbf{q}_n^T$$

This is called the **eigendecomposition** or **eigenvalue decomposition**. 分解. 必须用 orthgonal matrix

Algebraic multiplicity and geometric multiplicity

Definition 24.10 (**Geometric multiplicity**) 几何重数
Let λ_0 be an eigenvalue of the $n \times n$ square matrix A. The **geometric multiplicity** w.r.t. the eigenvalue λ_0 is the Nullity of $A - \lambda_0 I$, denoted by $\gamma_A(\lambda_0)$, i.e., $\gamma_A(\lambda_0) = \dim(Null(A - \lambda_0 I))$.

Definition 24.11 (**Algebraic multiplicity**) 代数重数.
Let λ_0 be an eigenvalue of the $n \times n$ square matrix A and the characteristic polynomial is $p_A(\lambda)$.

The **algebraic multiplicity** w.r.t. the eigenvalue λ_0 is the integer $\mu_A(\lambda_0)$ such that $(\lambda_0 - \lambda)^{\mu_A(\lambda_0)}$ is a factor of $p_A(\lambda)$ while $(\lambda_0 - \lambda)^{\mu_A(\lambda_0)+1}$ is not a factor of $p_A(\lambda)$. $\mu_A(\lambda_0)$ is the highest power such that $(\lambda_0 - \lambda)^{\mu_A(\lambda_0)}$ is a factor of $p_A(\lambda)$.

Theorem 24.12 Let λ be an eigenvalue of $n \times n$ matrix A, then

algebra multiplicity.

$$1 \leq \gamma_A(\lambda) \leq \mu_A(\lambda) \leq n$$

Proof. Skipped.

$= \dim(Null(A-\lambda I))$

geometric multiplicity

Theorem 24.13 (**Sufficient and Necessary Condition for Diagonalization**) Let A be a square matrix with size $n \times n$, then A is diagonalizable $\Leftrightarrow \gamma_A(\lambda) = \mu_A(\lambda)$ for any eigenvalue λ.
Proof. Skipped.

矩阵对角化的充要条件.

Complex Inner Product

Definition 24.15 (**Complex Inner Product**) Let V be a vector space over complex numbers. An inner product on V is an operation that assigns to each pair of vectors $\mathbf{z}$ and $\mathbf{w}$ in V a complex number $\langle \mathbf{z}, \mathbf{w} \rangle$ satisfying the following conditions:
(I) $\langle \mathbf{z}, \mathbf{z} \rangle \geq 0$ with equality valid if only if $\mathbf{z} = \mathbf{0}$.
(II) $\langle \mathbf{z}, \mathbf{w} \rangle = \overline{\langle \mathbf{w}, \mathbf{z} \rangle}$ for $\mathbf{z}$ and $\mathbf{w}$ in V. (This is main difference from the inner product defined over real number field)
(III) $\langle \alpha \mathbf{u} + \beta \mathbf{v}, \mathbf{w} \rangle = \alpha \langle \mathbf{u}, \mathbf{w} \rangle + \beta \langle \mathbf{v}, \mathbf{w} \rangle$ 抽象.

Remark
$\langle \mathbf{u}, \alpha \mathbf{v} + \beta \mathbf{w} \rangle = \overline{\langle \alpha \mathbf{v} + \beta \mathbf{w}, \mathbf{u} \rangle} = \overline{\alpha \langle \mathbf{v}, \mathbf{u} \rangle + \beta \langle \mathbf{w}, \mathbf{u} \rangle} = \overline{\alpha} \overline{\langle \mathbf{v}, \mathbf{u} \rangle} + \overline{\beta} \overline{\langle \mathbf{w}, \mathbf{u} \rangle} = \overline{\alpha} \langle \mathbf{u}, \mathbf{v} \rangle + \overline{\beta} \langle \mathbf{u}, \mathbf{w} \rangle$

$\mathbb{R}^n$	$\mathbb{C}^n$
$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T \mathbf{y}$	$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{y}^H \mathbf{x}$
$\mathbf{y}^T \mathbf{x} = \mathbf{x}^T \mathbf{y}$	$\mathbf{y}^H \mathbf{x} = \overline{\mathbf{x}^H \mathbf{y}}$
$\lVert \mathbf{x} \rVert^2 = \mathbf{x}^T \mathbf{x}$	$\lVert \mathbf{x} \rVert^2 = \mathbf{x}^H \mathbf{x}$
$\mathbf{x} \perp \mathbf{y} \Leftrightarrow \mathbf{x}^T \mathbf{y} = 0$	$\mathbf{x} \perp \mathbf{y} \Leftrightarrow \mathbf{y}^H \mathbf{x} = 0$

$\alpha, \beta \in \mathbb{C}$

Let $A, B \in \mathbb{C}^{m\times n}$, $C \in \mathbb{C}^{n\times l}$, then

(1) $(A^H)^H = A$.

(2) $(\alpha A + \beta B)^H = \bar{\alpha}A^H + \bar{\beta}B^H$. $\Leftarrow$ $\overline{z_1 z_2} = \bar{z}_1 \cdot \bar{z}_2$.

(3) $(AC)^H = C^H A^H$.

厄米矩阵.

Definition 24.17 (**Hermitian Matrix**) Let $M = A + iB = (m_{ij})_{m\times n}$, where $A, B \in \mathbb{R}^{n\times n}$, then conjugate of M is the matrix $\bar{M} = A - iB = (\bar{m}_{ij})_{m\times n}$. Define

$$M^H = (\bar{M})^T$$

$$M = \begin{bmatrix} 3 & 2-i \\ 2+i & 4 \end{bmatrix}$$

共轭转置 = 原矩阵.

$$M^H = \begin{bmatrix} 3 & \overline{2+i} \\ \overline{2-i} & 4 \end{bmatrix} = \begin{bmatrix} 3 & 2-i \\ 2+i & 4 \end{bmatrix} = M$$

酉矩阵.

Definition 24.18 (**Unitary Matrix**) A matrix $U \in \mathbb{C}^{n\times n}$ is said to be **unitary matrix** if its column vectors form an orthonormal set in $\mathbb{C}^n$. In particular, if $U \in \mathbb{R}^{n\times n}$ and its column vectors form an orthonormal set in $\mathbb{R}^n$, then U is called the **orthogonal matrix**. $U^T = U^{-1}$

Remark. U is unitary$\Leftrightarrow$ $U^H U = I \Leftrightarrow U^{-1} = U^H$.
A real unitary matrix is the orthogonal matrix.

$$A = \begin{bmatrix} \frac{i}{\sqrt{2}} & \frac{-i}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \qquad A^H A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$$

$$A^H = \begin{bmatrix} -\frac{i}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

Theorem 24.20 Let $M \in \mathbb{C}^{n\times n}$ be the Hermitian matrix with eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_n$, then

(1) $\lambda_i \in \mathbb{R}$, $\forall\ i = 1, \cdots, n$.

(2) If $\lambda_i \neq \lambda_j$, then the eigenvectors $\mathbf{x}_i, \mathbf{x}_j$ w.r.t λ_i, λ_j are orthogonal.

Theorem 24.23 (**Schur's decomposition**) For each $n \times n$ matrix A, there exists a unitary matrix U such that $U^H AU$ is upper triangular. i.e., $A = UTU^H$(T is an upper triangular matrix.)

Theorem 24.24 (**Spectral Theorem for Complex Matrix**) If A is Hermitian, then there exists a unitary matrix U that diagonalizes A.

奇异值分解.

Singular Value Decomposition

orthogonal matrix

For any $A \in \mathbb{R}^{m\times n}$, it can be decomposed into

$$A = U\boxed{\Sigma}V^T$$

diagonal-like

where U is a $m \times m$ orthogonal matrix, V is a $n \times n$ orthogonal matrix, Σ is a diagonal-like matrix. $\Sigma = (\tilde{\sigma}_{ij})_{m\times n}$ is defined as $\tilde{\sigma}_{ij} = 0, \text{if} i \neq j$, $\tilde{\sigma}_{ii} = \sigma_i, i = 1, \cdots, \min(m, n)$.

Case 1: If $m \geq n$ (tall matrix), $\Sigma = (\tilde{\sigma}_{ij})_{m\times n}$ is defined as $\tilde{\sigma}_{ij} = 0, \text{if} i \neq j$, $\tilde{\sigma}_{ii} = \sigma_i, i = 1, \cdots, n$, where $\sigma_1 \geq \cdots \geq \sigma_n \geq 0$. If rank of A is r, then $\sigma_1 \geq \cdots \geq \sigma_r > 0 = \sigma_{r+1} = \cdots = \sigma_n = 0$.

Case 2: If $m < n$ (fat matrix), $\Sigma = (\tilde{\sigma}_{ij})_{m\times n}$ is defined as $\tilde{\sigma}_{ij} = 0, \text{if} i \neq j$, $\tilde{\sigma}_{ii} = \sigma_i, i = 1, \cdots, m$, where $\sigma_1 \geq \cdots \geq \sigma_m \geq 0$. If rank of A is r, then $\sigma_1 \geq \cdots \geq \sigma_r > 0 = \sigma_{r+1} = \cdots = \sigma_m = 0$.

If $rank(A) = r$, then for both cases ($m \geq n$ and $m \leq n$), Σ is defined as

$$\Sigma_{m\times n} = \begin{bmatrix} \Sigma_1 & O_1 \\ O_2 & O_3 \end{bmatrix}$$

$\Sigma_1 = \mathrm{diag}(\sigma_1, \cdots, \sigma_r)$, $O_1 = O_{r\times(n-r)}$, $O_2 = O_{(m-r)\times r}$, $O_3 = O_{(m-r)\times(n-r)}$, where $\sigma_1, \cdots, \sigma_r$ are positive numbers.

Theorem 17.10 (**Kernel and Image are Subspaces**) Let L be a linear transformation from V to W, and let S be a subspace of V, then

(1) Ker(L) is a subspace of V.

V中使得变换后为0的元素.

① $L(0)=0$, $0\in ker(V)$

② $L(u+v)=L(u)+L(v)=0+0$, $u+v\in ker(V)$.

③ $L(\alpha u) = \alpha L(u) = \alpha \cdot 0$, $\alpha u \in ker(V)$.

(2) $L(S)$ is a subspace of W.

exercise.

Proof. Skipped. see Steven's book P174-175.

$L(S) = \{ w\in W \mid \exists v\in S, L(v)=w\}$

① $0\in S$, $L(0)=0_W\in L(S)$.

② $u, v\in S$, $u+v\in S$, $L(u+v)\in L(S)$

③ $u\in S$, $\alpha u\in S$, $L(\alpha u)\in L(S)$

Question 1

Prove the following statements about orthogonality.

(a) If matrix A is orthogonal, then its determinants must be either 1 or -1.

(b) If matrices A and B are both orthogonal, then AB is also orthogonal.

(c) If an $n\times n$ matrix A is orthogonal, then A^{-1} definitely exists and must also be orthogonal.

(a) $A^TA=I$ $\quad \det(A)^2=1$, $\det(A)=\pm 1$.

(b) $(AB)^TAB = B^TA^TAB = B^TIB = I$.

(c) ① A is orthogonal, all the column vectors are L.I.

$\therefore$ rank$(A)=n$. A^{-1} exists.

② $(AA^{-1})^T = I \Rightarrow (A^{-1})^TA^T = (A^{-1})^TA^{-1} = I$.

$\therefore$ A^{-1} is also orthogonal.

Question 5. Let A be an $n\times n$ matrix and let $B = A+I$. Is it possible for A and B to be similar? Explain.

✓. if A and B are similar, A and B have same eigenvalues.

However, $Trace(B) = Trace(A) + Trace(I)$

$= Trace(A) + n$.

$\therefore \sum_{i=1}^{r}\lambda_{i(B)} \neq \sum_{i=1}^{r}\lambda_{i(A)}$ so A, B cannot be similar.

Question 6. Let Q be an orthogonal matrix.

(a) Show that if λ is an eigenvalue of Q, then $|\lambda| = 1$.

(b) Show that $|\det(Q)| = 1$.

6. (a) $Qx = \lambda x$. $\quad Q^TQx = \lambda Q^Tx$

$\Rightarrow x = \lambda Q^T x$.

$\|x\| = \|\lambda Q^T x\| = |\lambda| \, \|Q^T x\| = |\lambda| \, \|x\|$

$\therefore |\lambda| = 1$

(b) $|\det(Q)| = \left|\prod_{i=1}^{n} \lambda_i\right| = 1$

Question 9. For each of the following, find a matrix B such that $B^2 = A$.

(a)

$$A = \begin{bmatrix} 2 & 1 \\ -2 & -1 \end{bmatrix}$$

(b)

$$A = \begin{bmatrix} 9 & -5 & 3 \\ 0 & 4 & 3 \\ 0 & 0 & 1 \end{bmatrix}$$

9. $B^2 = A = XDX^{-1}$

(a) $P_A(\lambda) = \begin{vmatrix} 2-\lambda & 1 \\ -2 & -1-\lambda \end{vmatrix} = \lambda^2 - \lambda = 0$.

$\lambda_1 = 1$, $\text{Null}\left(\begin{pmatrix} 1 & 1 \\ -2 & -2 \end{pmatrix}\right) = \text{span}\left\{\begin{bmatrix} 1 \\ -1 \end{bmatrix}\right\}$.

$\lambda_2 = 0$, $\text{Null}\left(\begin{pmatrix} 2 & 1 \\ -2 & -1 \end{pmatrix}\right) = \text{span}\left\{\begin{bmatrix} 1 \\ -2 \end{bmatrix}\right\}$.

$X = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix} \qquad D = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$.

$B = XD^{\frac{1}{2}}X^{-1} = XDX^{-1} = \begin{bmatrix} 2 & 1 \\ -2 & -1 \end{bmatrix}$.

(b) $P_A(\lambda) = \begin{vmatrix} 9-\lambda & -5 & 3 \\ 0 & 4-\lambda & 3 \\ 0 & 0 & 1-\lambda \end{vmatrix} = (9-\lambda)(4-\lambda)(1-\lambda)$

$\lambda_1 = 9$, $\text{Null}\left(\begin{pmatrix} 0 & -5 & 3 \\ 0 & -5 & 3 \\ 0 & 0 & -8 \end{pmatrix}\right) = \text{span}\left\{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right\}$.

$\lambda_2 = 4$, $\text{Null}\left(\begin{pmatrix} 5 & -5 & 3 \\ 0 & 0 & 3 \\ 0 & 0 & -3 \end{pmatrix}\right) = \text{span}\left\{\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}\right\}$.

$\lambda_3 = 1$, $\text{Null}\left(\begin{pmatrix} 8 & -5 & 3 \\ 0 & 3 & 3 \\ 0 & 0 & 0 \end{pmatrix}\right) = \text{span}\left\{\begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}\right\}$.

$X = \begin{bmatrix} 1 & 1 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}$, $D = \begin{bmatrix} 9 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $D^{\frac{1}{2}} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

$B = XD^{\frac{1}{2}}X^{-1} = \begin{bmatrix} 1 & 1 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$

$\left(\left[\begin{array}{ccc|ccc} 1 & 1 & -1 & 1 & 0 & 0 \\ 0 & 1 & -1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right] \Rightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right] \right)$

$= \begin{bmatrix} 3 & 2 & -1 \\ 0 & 2 & -1 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 3 & -1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{bmatrix}$

Question 15. Let A be a diagonalizable $n \times n$ matrix. Prove that if B is any matrix that is similar to A, then B is diagonalizable.

15. $B = S^{-1}AS$.

$$\begin{aligned} P_B(\lambda) &= \det(B - \lambda I) \\ &= \det(S^{-1}AS - \lambda S^{-1}S) \\ &= \det(S^{-1}(AS - \lambda IS) \\ &= \det(S^{-1}) \det(A - \lambda I) \det(S) \\ &= \det(A - \lambda I) \\ &= P_A(\lambda) \end{aligned}$$

$\because$ A is diagonalizable $\quad\quad$ $\therefore$ B is diagonalizable.

7. 二次型

How to classify the type for the quadratic equation with two unknowns:

$$ax^2 + 2bxy + cy^2 + dx + ey + f = 0?$$

Definition 26.1 (**Quadratic Equation with two unknowns**) A quadratic equation in two unknowns is an equation of the form

$$\boxed{ax^2 + 2bxy + cy^2} + \boxed{dx + ey + f} = 0 - - - - - - - - - - - - - - - - (*)$$

(*) can be written in the form

平移 translation

$$\begin{bmatrix} x & y \end{bmatrix}\begin{bmatrix} a & b \\ b & c \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} d & e \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} + f = 0$$

real symmetric matrix (可以规范化)

$$[x\ y]\begin{bmatrix} a & b \\ b & c \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = [ax+by\ \ bx+cy]\begin{bmatrix} x \\ y \end{bmatrix} = ax^2 + 2bx + cy^2.$$

Let

$$\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}, \quad A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$$

$\exists$ a orthonormal matrix Q.
s.t $Q^TAQ = D$
$\Rightarrow A = QDQ^T$

then (*) can be written as

$$\mathbf{x}^T A\mathbf{x} + [d, e]\mathbf{x} + f = 0$$

The term $\mathbf{x}^T A\mathbf{x}$ is called the quadratic form associated with quadratic equation (*). The graph corresponding to (*) is called the **conic section**.

圆锥截面

Example 26.2

$$\overset{z^T}{[x, y]}\overset{A}{\begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}}\overset{z}{\begin{bmatrix} x \\ y \end{bmatrix}} + [0\ \ 8\sqrt{2}]z - 4 = 0.$$

$$3x^2 + 2xy + 3y^2 + 8\sqrt{2}y - 4 = 0$$

Here $A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$ is a real symmetric matrix. Now we can take

by spectral theorem, A can be diagonalized.

$Q = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$ to be an orthogonal matrix such that

$$Q^T AQ = diag(2, 4). = \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}$$

对角矩阵

$$A = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix}\begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix}$$

列向量是单位特征向量
并且互相正交

Let $\begin{bmatrix} x \\ y \end{bmatrix} = Q\begin{bmatrix} \hat{x} \\ \hat{y} \end{bmatrix}$ (changing the coordinate system)

$\hat{z} = Q^T z$, $z = Q\hat{z}$

$$z^TAz = (Q\hat{z})^TAQ\hat{z} = \hat{z}^TQ^TAQ\hat{z} = \hat{z}^TD\hat{z}$$

只有拉伸

$$[\hat{z}_1\ \hat{z}_2]\begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}\begin{bmatrix} \hat{z}_1 \\ \hat{z}_2 \end{bmatrix} + \overset{Q}{[0\ \ 8\sqrt{2}]}\begin{bmatrix} \hat{z}_1 \\ \hat{z}_2 \end{bmatrix} - 4 = 0 \quad [-8\ \ 8]$$

$$A = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix}\begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix}$$

即有旋转
也有拉伸

只有旋转（列向量是单位并且正交）

消除二次项. ✓

Then

$$\Rightarrow \quad 2(\hat{x})^2 + 4(\hat{y})^2 - 8\hat{x} + 8\hat{y} = 4$$

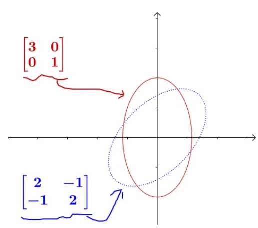
which equivalent to

$$2(\hat{x} - 2)^2 + 4(\hat{y} + 1)^2 = 16$$

i.e.,

$$\frac{(\hat{x} - 2)^2}{8} + \frac{(\hat{y} + 1)^2}{4} = 1$$

运动分解.



$\begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}$

$\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$

which is an ellipse.

Definition 26.3 (**Definite quadratic form and definite matrix**) Let $\mathbf{x} \in \mathbb{R}^n$, $A \in \mathbb{R}^{n\times n}$ be symmetric, then

(1) The quadratic form $f(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ is called **positive definite** if $f(\mathbf{x}) > 0$ for any $\mathbf{x} \neq \mathbf{0}$. And correspondingly, A is called **positive definite matrix**. 正定二次型

(2) The quadratic form $f(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ is called **positive semidefinite** if $f(\mathbf{x}) \geq 0$ for any $\mathbf{x} \neq \mathbf{0}$. And correspondingly, A is called **positive semidefinite matrix**. 半正定二次型.

(3) The quadratic form $f(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ is called **indefinite** if $f(\mathbf{x})$ takes different signs. 不定

(4) The quadratic form $f(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ is called **negative definite** if $f(\mathbf{x}) < 0$ for any $\mathbf{x} \neq \mathbf{0}$. And correspondingly, A is called **negative definite matrix**. 负定二次型

(5) The quadratic form $f(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ is called **negative semidefinite** if $f(\mathbf{x}) \leq 0$ for any $\mathbf{x} \neq \mathbf{0}$. And correspondingly, A is called **negative semidefinite matrix**. 半负定二次型

Theorem 26.5 Let $A \in \mathbb{R}^{n\times n}$ be symmetric, then A is positive definite if only if all eigenvalues are positive.

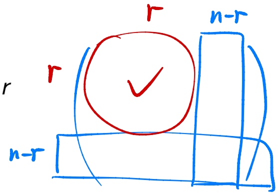
Proof. Since A is symmetric, by spectral theorem for real symmetric matrix, there exists an orthogonal matrix Q such that $Q^{-1}AQ = Q^T AQ = D$, where D is the diagonal matrix. Let $\hat{\mathbf{x}} = Q^T\mathbf{x}$ then $\mathbf{x} = Q\hat{\mathbf{x}}$ and $\mathbf{x}^T A\mathbf{x} = (Q\hat{\mathbf{x}})^T AQ\hat{\mathbf{x}} = \hat{\mathbf{x}}^T Q^T AQ\hat{\mathbf{x}} = \hat{\mathbf{x}}^T D\hat{\mathbf{x}}$. Since Q is invertible and $\hat{\mathbf{x}} = Q^T\mathbf{x}$, thus

$$\mathbf{x}^T A\mathbf{x} > 0,\ \forall \mathbf{x} \neq \mathbf{0} \Leftrightarrow \hat{\mathbf{x}}^T D\hat{\mathbf{x}} > 0,\ \forall \hat{\mathbf{x}} \neq \mathbf{0}$$

Thus, A is positive definite $\Leftrightarrow$ the entries in diagonal elements of D are all positive$\Leftrightarrow$ all eigenvalues of A are positive.

Remark.
1. Let $A \in \mathbb{R}^{n\times n}$ be symmetric, then A is negative definite if only if all eigenvalues are negative.
2. Let $A \in \mathbb{R}^{n\times n}$ be symmetric, then A is indefinite if only if eigenvalues have different signs.

Definition 26.8 (**Leading Principal Submatrix**) Given A, let A_r denoted the matrix formed by deleting the last $n-r$ rows and last $n-r$ columns. A_r is called the leading principle submatrix of A of order r.



Property 26.9 (**Property of Positive Definite Matrix**) If A ($A \in \mathbb{R}^{n\times n}$) is a symmetric positive definite matrix, then all the leading principal submatrices $A_1, A_2, \cdots, A_n$ of A are all positive definite matrices, and thus all leading principal submatrices have positive determinants.

Proof. From the fact that

$$
\begin{aligned}
&[x_1, x_2, \cdots, x_r] A_r [x_1, x_2, \cdots, x_r]^T \\
=&[x_1, x_2, \cdots, x_r, 0, \cdots, 0] A [x_1, x_2, \cdots, x_r, 0, \cdots, 0]^T > 0
\end{aligned}
$$

Theorem 26.12 Let $A \in \mathbb{R}^{n\times n}$ be symmetric positive definite matrix, then

$$A = LD\hat{U} \quad \text{(LDU factorization)}$$

where

$$\hat{U} = L^T$$

$$
L = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ * & 1 & \ddots & 0 \\ * & \ddots & \ddots & 0 \\ * & * & * & 1 \end{bmatrix}, \quad D = \operatorname{diag}\underbrace{(u_{11}, u_{22}, \cdots, u_{nn})}_{>0}
$$

$$
U = \begin{bmatrix} 1 & * & \cdots & * \\ 0 & 1 & \ddots & * \\ 0 & \ddots & \ddots & * \\ 0 & 0 & 0 & 1 \end{bmatrix}
$$

$$A = L \underbrace{\begin{pmatrix} \sqrt{u_{11}} & & & \\ & \sqrt{u_{22}} & & \\ & & \ddots & \\ & & & \sqrt{u_{nn}} \end{pmatrix}}_{L_1} \underbrace{\begin{pmatrix} \sqrt{u_{11}} & & & \\ & \sqrt{u_{22}} & & \\ & & \ddots & \\ & & & \sqrt{u_{nn}} \end{pmatrix} L^T}_{L_1^T}$$

$$= L_1 L_1^T$$

L is a unit lower triangular matrix, and U is a unit upper triangular matrix, D is a diagonal matrix with positive diagonal entries. **Remark. Indeed** $U = L^T$ **and** $A = LDL^T$**.**

Theorem 26.16 Let $A \in \mathbb{R}^{n \times n}$ and A is symmetric, then the following are equivalent

(1) A is positive definite. most important.

(2) The leading principal submatrices $A_1, \cdots, A_n$ all have positive determinants.

(3) $A = LU$, where U is an upper triangular matrix with positive diagonal elements, L is a unit lower triangular matrix.

(4) $A = LDL^T$, where D is a diagonal matrix with positive diagonal elements, L is a unit lower triangular matrix.

(5) $A = LL^T$, where L is a lower triangular matrix with positive diagonal elements.

(6) $A = B^T B$ for some invertible matrix B.