理解矩阵

1. [理解矩阵（一）]
2. [理解矩阵（二）]
3. [理解矩阵（三）]

总结：

- **“空间”是容纳运动的一个对象集合，而变换则规定了对应空间的运动。**
- **矩阵的本质是运动的描述。**
- **所谓变换，其实就是空间里从一个点（元素/对象）到另一个点（元素/对象）的跃迁**
- **矩阵是线性空间里的变换的描述，矩阵与向量相乘，就是实施运动（变换）的过程。**
- **矩阵是线性空间中的线性变换的一个描述。在一个线性空间中，只要我们选定一组基，那么对于任何一个线性变换，都能够用一个确定的矩阵来加以描述。**
- **所谓相似矩阵，就是同一个线性变换的不同的描述矩阵，它们的本质是一样的，所以本征值相同。**
- **矩阵不仅可以作为线性变换的描述，而且可以作为一组基的描述。而作为变换的矩阵，不但可以把线性空间中的一个点给变换到另一个点去，而且也能够把线性空间中的一个坐标系（基）表换到另一个坐标系（基）去。而且，变换点与变换坐标系，具有异曲同工的效果。（运动等价于坐标系的变换，即运动是相对的！）**

Linear Algebra Done Right

Vector Spaces

1. A ***vector space*** is a set V along with an addition on V and a scalar multiplication on V such that the following **properties** hold:

   - **commutativity**

     $u + v = v + u$ for all $u, v \in V$;

   - **associativity**

     $(u + v) + w = u + (v + w)$ and $(ab)v = a(bv)$ for all $u, v, w \in V$ and all $a, b \in F$;

   - **additive identity** (must be unique)

     there exists an element $0 \in V$ such that $v + 0 = v$ for all $v \in V$

   - **additive inverse** (must be unique)

     for every $v \in V$, there exists $w \in V$ such that $v + w = 0$;

   - **multiplication identity**

     $1v = v$ for all $v \in V$;

   - **distributive properties**

     $a(u + v) = au + av$ and $(a + b)u = au + bu$ for all $a, b \in F$ and all $u, v \in V$

   Polynomial is also a vector space

2. **Propositions**:
   - $0v = 0$ for every $v \in V$.
   - $a0 = 0$ for every $a \in F$.
   - $(-1)v = -v$ for every $v \in V$.

   Remember that the $0, -1$ above may not be the regular $0, -1$ as we known, they denote additive inverse and additive identity of the specific operation(pending)

3. **Subspaces**: If U is a subset of V, then to check that U is a subspace of V we need only check that U satisfies the following:
   - **additive identity** : $0 \in U$;
   - **closed under addition** : $u, v \in U$ implies $u + v \in U$;
   - **closed under scalar multiplication** : $a \in F$ and $u \in U$ implies $au \in U$.

   A subspace can be regard as a large space with some properties being limited

4. **Sums and Direct Sums**: Specifically, we say that V is the direct sum of subspaces$U_1, \ldots, U_m$,

   (1) $V = U_1 \oplus \cdots \oplus U_m$,

   (2) each element of V can be written **uniquely** as a sum $u_1 + \cdots + u_m$, where each $u_j \in U_j$.

   Proposition: Suppose that U and W are subspaces of V. Then $V = U \oplus W$ if and only if $V = U + W$ and $U \cap W = \{0\}$. (only with the case of two subspaces).

   $V = U \oplus W \Leftrightarrow V = U + W \& U \cap W = \{0\}$

Finite-Dimensional Vector Spaces

Span and Linear Independence

1. **The set of all linear combinations of $(v_1, \ldots, v_m)$ is called the span of $(v_1, \ldots, v_m)$, denoted $span(v_1, \ldots, v_m)$.**

   $span(v_1, \ldots, v_m) = \{a_1 v_1 + \cdots + a_m v_m : a_1, \ldots, a_m \in F\}$.

   the span of any list of vectors in V is a subspace of V.

2. **Infinite-dimensional** vector spaces are the center of attention in the branch of mathematics called *functional analysis*.

3. A list $(v_1, \ldots, v_m)$ of vectors in V is called **linearly independent** if the only choice of $a_1, \ldots, a_m \in F$ that makes $a_1 v_1 + \cdots + a_m v_m$ equal 0 is $a_1 = \cdots = a_m = 0$.

Bases

A basis of V is a list of vectors in V that is linearly independent and spans V.

1. A list $(v_1, \ldots, v_n)$ of vectors in V is a basis of V if and only if every $v \in V$ can be written uniquely in the form

   $v = a_1 v_1 + \cdots + a_n v_n$

   where $a_1, \ldots, a_n \in F$.

2. Every spanning list in a vector space can be reduced to a basis of the vector space.

3. Every linearly independent list of vectors in a finite-dimensional vector space can be extended to a basis of the vector space.
4. Suppose V is finite dimensional and U is a subspace of V. Then there is a subspace W of V such that $V = U \oplus W$.

Dimension

1. If V is finite dimensional and U is a subspace of V, then $dimU \leq dimV$.
2. If U_1 and U_2 are subspaces of a finite-dimensional vector space, then $dim(U_1 + U_2) = dimU_1 + dimU_2 - dim(U_1 \cap U_2)$.
   - It is same as the number of elements in the union of two finite sets:
   $$n(A + B) = n(A) + n(B) - n(A \cap B) \tag{69}$$
3. Suppose V is finite dimensional and $U1, \ldots, Um$ are subspaces of V such that

   $V = U_1 + \cdots + U_m$

   and

   $dimV = dimU_1 + \cdots + dimU_m$.

   Then $V = U_1 \oplus \cdots \oplus U_m$.

   Recall that direct sum is analogous to disjoint union.
   - In sets:

     If $B = A_1 \cup \ldots A_m$, and $n(B) = n(A_1) + \cdots + n(A_m)$, then the union is a disjoint union.

Linear Maps

Definitions and Examples

A linear map from V to W is a function $T : V \to W$ with the following properties:

- **additivity** : $T(u + v) = Tu + Tv$ for all $u, v \in V$;
- **homogeneity** : $T(av) = a(Tv)$ for all $a \in F$ and all $v \in V$.

For $S, T \in L(V, W)$,

- **associativity**
$$(T_1T_2)T_3 = T_1(T_2T_3) \tag{1}$$
- **identity**
$$TI = T andIT = T \tag{2}$$
- **distributive properties**
$$(S_1 + S_2)T = S_1T + S_2TandS(T_1 + T_2) = ST_1 + ST_2 \tag{3}$$
- Multiplication of linear maps is **not commutative**.
$$ST \neq TS \tag{4}$$

Null Spaces and Ranges

1. $nullT = v \in V : Tv = 0$
    - If $T \in L(V, W)$, then $nullT$ is a subspace of V.
    - Let $T \in L(V, W)$. Then T is **injective（单射）** if and only if $nullT = 0$.

      (A linear map $T : V \to W$ is called injective if whenever $u, v \in V$ and $Tu = Tv$, we have $u = v$.)
2. $rangeT = Tv : v \in V$
    - If $T \in L(V, W)$, then $rangeT$ is a subspace of W.
    - A linear map $T : V \to W$ is called **surjective（满射）** if its range equals W.
3. If V is finite dimensional and $T \in L(V, W)$, then $rangeT$ is a finite-dimensional subspace of W and $dimV = dimnullT + dimrangeT$
4. If V and W are finite-dimensional vector spaces such that $dimV > dimW$, then no linear map from V to W is **injective**.

$$dimnullT = dimV - dimrangeT \geq dimV - dimW > 0 \tag{5}$$

5. If V and W are finite-dimensional vector spaces such that $dimV < dimW$, then no linear map from V to W is **surjective**.

$$dimrangeT = dimV - dimnullT \leq dimV < dimW \tag{6}$$

Conclusions:

1. **A homogeneous system (齐次系统) of linear equations in which there are more variables than equations must have nonzero solutions.** ($dimV > dimW$, 非单射, 核空间维度大于零, 即存在非零解)
2. **An inhomogeneous system (非齐次系统) of linear equations in which there are more equations than variables has no solution for some choice of the constant terms.** ($dimV < dimW$, 非满射, 像空间维度小于W, W中部分值无解)

The Matrix of a Linear Map

$$M(T) = \begin{bmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{bmatrix} \tag{7}$$

Let $T \in L(V, W)$. Suppose that $(v_1, \ldots, v_n)$ is a basis of V and $(w_1, \ldots, w_m)$ is a basis of W.

$$Tv_k = a_{1,k}w_1 + \cdots + a_{m,k}w_m, \tag{8}$$

the matrix of T with respect to the bases $(v_1, \ldots, v_n)$ and $(w_1, \ldots, w_m)$; we denote it by

$$M(T, (v_1, \ldots, vn), (w_1, \ldots, w_m)) \tag{9}$$

$$\begin{array}{c c} & \begin{array}{ccccc} v_1 & \ldots & v_k & \ldots & v_n \end{array} \\ \begin{array}{c} w_1 \\ \vdots \\ w_m \end{array} & \begin{bmatrix} & & a_{1,k} & & \\ & & \vdots & & \\ & & a_{m,k} & & \end{bmatrix} \end{array}$$

$$M(Tv) = \begin{bmatrix} a_{1,1}b_1 + \cdots + a_{1,n}b_n \\ \vdots \\ a_{m,1}b_1 + \cdots + a_{m,n}b_n \end{bmatrix} \tag{10}$$

Invertibility

A linear map $T \in L(V, W)$ is called invertible if there exists a linear map $S \in L(W, V)$ such that ST equals the identity map on V and TS equals the identity map on W.

- A linear map is **invertible** if and only if it is **injective** and **surjective**.
- Two vector spaces are called **isomorphic（同构）** if there is an invertible linear map from one vector space onto the other one.
  - two isomorphic spaces have the same properties;
  - we can think of an invertible linear map as a **relabeling** of the elements of a vector space.
  - Two finite-dimensional vector spaces are isomorphic if and only if they have the **same dimension.** ($dimV = dimW$, 核空间为0, 像空间为W, 既是单设又是满射, 故可逆) (且其齐次方程下的变换只有零解, 故基下的矩阵为列线性无关的方阵, 即为满秩矩阵, 故可逆)
- If V and W are finite dimensional, then $L(V, W)$ is finite dimensional and

$$dimL(V, W) = (dimV)(dimW). \tag{11}$$

- Suppose V is finite dimensional. If $T \in L(V)$, then the following are equivalent:
  - T is invertible;
  - T is injective;
  - T is surjective.

Polynomials

Recall that a function $p : F \to F$ is called a polynomial with coefficients in F if there exist $a_0, \ldots, a_m \in F$ such that

$$p(z) = a_0 + a_1 z + a_2 z_2 + \cdots + a_m z_m \tag{12}$$

Degree

1. If all the coefficients $a_0, \ldots, a_m$ equal 0, then we say that p has degree $-\infty$.
2. A polynomial may have **more than one** degree.
3. A number $\lambda \in F$ is called a **root** of a polynomial $p \in P(F)$ if
$$p(\lambda) = 0 \tag{13}$$
4. Suppose $p \in P(F)$ is a polynomial with degree $m \geq 1$. Let $\lambda \in F$. Then λ is a root of p if and only if there is a polynomial $q \in P(F)$ with degree $m - 1$ such that
$$p(z) = (z - \lambda)q(z) \tag{14}$$
5. Suppose $p \in P(F)$ is a polynomial with degree $m \geq 0$. Then p has at most m distinct roots in F.
6. Suppose $a_0, \ldots, a_m \in F$. If
$$a_0 + a_1 z + a_2 z^2 + \cdots + a_m z^m = 0 \tag{15}$$

for all $z \in F$, then $a_0 = \cdots = a_m = 0$.

It implies that $(1, z, \ldots, z^m)$ is linearly independent in $P(F)$ for every nonnegative integer m.

7. **Division Algorithm:** Suppose $p, q \in P(F)$, with $p \neq 0$. Then there exist polynomials $s, r \in P(F)$ such that

$$q = sp + r \tag{16}$$

and $\deg r < \deg p$

Eigenvalues and Eigenvectors

Invariant Subspaces

1. For $T \in L(V)$ and U a subspace of V, we say that U is **invariant under** T if $u \in U$ implies $Tu \in U$. In other words, U is invariant under T if $T|_U$ is an operator on U.
2. If $T \in L(V)$, $nullT$ and $rangeT$ is invariant under T.
3. **Eigenvalue** : Specifically, a scalar $\lambda \in F$ is called an **eigenvalue** of $T \in L(V)$ if there exists a nonzero vector $u \in V$ such that $Tu = \lambda u$.
4. The equation $Tu = \lambda u$ is equivalent to $(T - \lambda I)u = 0$, so λ is an eigenvalue of T if and only if $T - \lambda I$ is **not injective**. λ is an eigenvalue of T if and only if $T - \lambda I$ is **not invertible**, and this happens if and only if $T - \lambda I$ is **not surjective**.
5. **Eigenvector:** An operator has an eigenvalue if and only if there exists a nonzero vector in its domain that gets sent by the operator to a scalar multiple of itself.
6. Let $T \in L(V)$. Suppose $\lambda_1, \ldots, \lambda_m$ are distinct eigenvalues of T and $v_1, \ldots, v_m$ are corresponding nonzero eigenvectors. Then $(v_1, \ldots, v_m)$ is linearly independent.

Upper-Triangular Matrices

1. Every operator on a finite-dimensional, nonzero, complex vector space has an eigenvalue.
2. Suppose $T \in L(V)$ and $(v_1, \ldots, v_n)$ is a basis of V. Then the following are equivalent:
   - the matrix of T with respect to $(v_1, \ldots, v_n)$ is upper triangular;
   - $Tv_k \in span(v_1, \ldots, v_k)$ for each $k = 1, \ldots, n$;
   - $span(v_1, \ldots, v_k)$ is invariant under T for each $k = 1, \ldots, n$.
3. Suppose $T \in L(V)$ has an upper-triangular matrix with respect to some basis of V. Then T is invertible if and only if all the entries on the **diagonal** of that upper-triangular matrix are **nonzero**.
4. Suppose $T \in L(V)$ has an upper-triangular matrix with respect to some basis of V. Then the **eigenvalues** of T consist precisely of **the entries on the diagonal** of that upper-triangular matrix.

Diagonal Matrices

1. An operator $T \in L(V)$ has a diagonal matrix with respect to some basis of V if and only if V **has a basis consisting of eigenvectors of** T.
2. If $T \in L(V)$ has $dimV$ distinct eigenvalues, then T has a diagonal matrix with respect to some basis of V .
3. Suppose $T \in L(V)$. Let $\lambda_1, \ldots, \lambda_m$ denote the distinct eigenvalues of T. Then the following are equivalent:

- T has a diagonal matrix with respect to some basis of V;
- V has a basis consisting of eigenvectors of T;
- there exist one-dimensional subspaces $U_1, \ldots, U_n$ of V, each invariant under T, such that $V = U_1 \oplus \cdots \oplus U_n$;
- $V = null(T - \lambda_1 I) \oplus \cdots \oplus null(T - \lambda_m I)$;

  > every vector in V is a linear combination of eigenvectors of T. Hence
  >
  > $$V = null(T - \lambda_1 I) + \cdots + null(T - \lambda_m I). \tag{17}$$
  >
  > $null(T - \lambda_m I) = \{x|(T - \lambda_m I)x = 0\}$, where x is the eigenvector of λ_m. Thus eigenvectors are linearly independent. This implies that the sum is a direct sum.

- $dimV = dimnull(T - \lambda_1 I) + \cdots + dimnull(T - \lambda_m I)$.

Invariant Subspaces on Real Vector Spaces

1. Every operator on a *finite-dimensional*, *nonzero*, *real vector space* has an invariant subspace of dimension 1 **or** 2.
2. Every operator on an **odd-dimensional** real vector space has **an eigenvalue**.

Inner-Product Spaces

Inner Products

1. the **complex conjugate** of λ is defined by $\tilde{\lambda} = a - bi$,
2. $|\lambda|^2 = \lambda\tilde{\lambda}$
3. **Properites** of *inner product*:
   - **positivity**
     $$\langle v, v\rangle \geq 0 \quad \forall v \in V; \tag{18}$$
   - **definiteness**
     $$\langle v, v\rangle = 0 \quad iff \quad v = 0; \tag{19}$$
   - **additivity in first slot** (also in second slot)
     $$\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle \quad \forall u, v, w \in V; \tag{20}$$
   - **homogeneity in first slot** (also in second slot)
     $$\langle av, w\rangle = a\langle v, w\rangle \quad \forall a \in F \quad and \quad \forall v, w \in V; \tag{21}$$
   - **conjugate symmetry**
     $$\langle v, w\rangle = \widetilde{\langle w, v\rangle} \quad \forall v, w \in V. \tag{22}$$
4. An ***inner-product space*** is a vector space V along with an inner product on V.

Norms

1. $\|(z1, \ldots, zn)\| = \sqrt{|z1|^2 + \cdots + |zn|^2}$.

2. $\|p\| = \sqrt{\int_0^1 |p(x)|^2 dx}$.

3. Two vectors $u, v \in V$ are said to be orthogonal if $\langle u, v\rangle = 0$.

4. **Pythagorean Theorem**: If u,v are orthogonal vectors in V,then

$$\|u+v\|^2 = \| u \|^2 + \| v \|^2. \tag{23}$$

5. **Cauchy-Schwarz Inequality**: If u,v ∈V, then

$$|\langle u, v\rangle| \leq \| u \| \| v \|. \tag{24}$$

This inequality is an equality if and only if one of u, v is a scalar multiple of the other.

6. **Triangle Inequality**: If $u, v \in V$, then

$$\|u+v\| \leq \| u \| + \| v \|. \tag{25}$$

This inequality is an equality if and only if one of u, v is a nonnegative multiple of the other.

7. **Parallelogram Equality**: If $u, v \in V$, then

$$\|u+v\|^2 + \| u-v \|^2 = 2(\| u \|^2 + \| v \|^2). \tag{26}$$

教材：[矩阵理论与应用]

第一章 线性代数概要与提高

引言

线性代数讨论的两个问题

1. 引入矩阵来解线性方程组
   - 线性方程组简洁地表示为 $Ax = b$
   - 二次型表示为

$$f(x) = f(x_1, x_2, \cdots, x_n) = \sum_{i,j=1}^{n} a_{ij} x_i x_j \tag{27}$$

即为 $f(x) = x^T Ax$. 因此线性方程组解的存在性取决于系数矩阵A与增广矩 阵(A,b)的秩相等与否.

- 将线性方程组Ax = b再次改写为向量形式

$$x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n = b \tag{28}$$

线性方程组的解实际上是向量b关于A的列$\alpha_1, \cdots, \alpha_n$的组合系数.

- 通过研究齐次线性方程组$Ax = 0$的解的结构, 可以知道任意解均可表为任一基础解系的线性组合, 而$Ax = b$如果有解, 则可将其化为相应的齐次线性方程组.

2. 矩阵乘法
   - 理解n阶方阵A的高次幂A^m是理解矩阵乘法的关键(也是理解和化简二次型的关键).

- 利用特征值与特征向量, 一些矩阵可以化为对角形, 即存在可逆矩阵P使得 $P^{-1}AP = D = \operatorname{diag}(\lambda_1, \lambda_2 \cdots, \lambda_n)$, 因此

  $$A^m = PD^mP^{-1} = P\operatorname{diag}(\lambda_1^m, \lambda_2^m \cdots, \lambda_n^m)P^{-1} \tag{29}$$

  特别地, 实对称矩阵可以正交对角化, 即存在正交矩阵Q使得

  $$Q^{-1}AQ = Q^TAQ = \operatorname{diag}(\lambda_1, \lambda_2 \cdots, \lambda_n) \tag{30}$$

  于是利用坐标变换x = Qy即可将实二次型f(x) = xTAx化为标准型

  $$f = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2 \tag{31}$$

矩阵乘法与分块矩阵

乘法证明

1. 第i行第j列元素为1, 其余元素均为0的$m \times n$矩阵称为基本矩阵, 记为$E_{ij}, 1 \le i \le m, 1 \le j \le n$. 任意$m \times n$矩阵$A = (a_{ij})$均能唯一地表示成

$$\begin{aligned} A &= \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \\ &= \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}E_{ij} \end{aligned}$$

因为

$$E_{ij}E_{kl} = \delta_{jk}E_{il}, \quad 1 \le i \le m, 1 \le j, k \le p, 1 \le l \le n \tag{32}$$

其中

$$\delta_{jk} = \begin{cases} 1, & j = k \\ 0, & else \end{cases} \tag{33}$$

于是矩阵的乘法可表为:

$$\begin{aligned} AB &= \left(\sum_{i=1}^{m}\sum_{j=1}^{p} a_{ij}E_{ij}\right)\left(\sum_{i=1}^{p}\sum_{j=1}^{n} b_{ij}E_{ij}\right) \\ &= \sum_{i,j,k,l}(a_{ij}E_{ij})(b_{kl}E_{kl}) = \sum_{i,j,k,l}(a_{ij}b_{kl})(E_{ij}E_{kl}) \\ &= \sum_{i,j,k,l}(a_{ij}b_{kl})(\delta_{jk}E_{il}) = \sum_{i,j,k,l}(\delta_{jk}a_{ij}b_{kl})E_{il} \\ &= \sum_{i=1}^{m}\sum_{j=1}^{n}\left(\sum_{k=1}^{p} a_{ik}b_{kj}\right)E_{ij} \end{aligned}$$

即乘积AB的第i行第j列的元素等于$\sum_{k=1}^{p} a_{ik}b_{kj}$, 这正是矩阵乘法的“左行右列”规则

2. 矩阵乘一个列向量等于该矩阵所有列的线性组合, 组合系数即是该列向量的对应元素. 同理, 一个行向量左乘一个矩阵等于该矩阵所有行的线性组合, 组合系数即是该行向量的对应元素

$$C_j = AB_j, \quad C^i = A^i B \tag{34}$$

   即矩阵AB的第j列是A的列向量的线性组合, 组合系数恰为矩阵B的第j列的相应元素; AB的 第i行是B的行向量的线性组合, 组合系数恰为矩阵A的第i行的相应元素.

3. 方程组$Ax = b$有解 $\Longleftrightarrow$ b是系数矩阵A的列的线性组合 $\Longleftrightarrow$ $r(A) = r(A, b)$, 即系数矩阵的秩等于增广矩阵的秩.
4. 齐次线性方程组$Ax = 0$有非零解 $\Longleftrightarrow$ A的列向量线性相关; 有唯一解(即零解)$\Longleftrightarrow$ A的列向量线性无关.

行列式

n阶方阵$A = (a_{ij})$的行列式记为$|A|$(另一个通用记号是$detA$), 它具有性质$|AB| = |A||B|$.

迹

方阵A的**迹** trA 是A的对角线元素之和 $\sum_{i=1}^{n} a_{ii}$

性质:

命题 1.1.1 设$A = (a_{ij})$, B均为n阶方阵, λ是数, 则

(1) $\mathrm{tr}(A + B) = \mathrm{tr}\, A + \mathrm{tr}\, B$;

(2) $\mathrm{tr}\,(\lambda A) = \lambda(\mathrm{tr}\, A)$;

(3) $\mathrm{tr}\,(AB) = \mathrm{tr}\,(BA)$ (此仅需A, B分别为$m \times n$和$n \times m$矩阵即可);

(4) $\mathrm{tr}\, A^T = \mathrm{tr}\, A$;

(5) $\mathrm{tr}(AA^*) = \sum_{i,j=1}^{n} |a_{ij}|^2$. 特别地, $\mathrm{tr}\,(AA^*) = 0 \iff A = 0$.

秩 逆矩阵

矩阵A的所有不为零的子式的最高阶数称为矩阵A的**秩**, 记为$r(A)$. 约定零矩阵的秩是0.

逆矩阵

对n阶方阵而言, “秩为n”(也称为“**满秩**”), “**非奇异**”与“**可逆**”是等价的三个概念. 可逆矩阵的逆矩阵是唯一的, 记为A^{-1}. 逆矩阵具有下述**性质**:

命题 1.1.2 (1) $A^{-1}=\frac{1}{|A|}\text{adj}\,A$;
(2) $(A^{-1})^{-1}=A$;
(3) $(A^T)^{-1}=(A^{-1})^T$;
(4) $(AB)^{-1}=B^{-1}A^{-1}$;
(5) 若数$\lambda\neq 0$, 则$(\lambda A)^{-1}=\lambda^{-1}A^{-1}$;
(6) $|A^{-1}|=|A|^{-1}$;
(7) 设A是$m\times n$矩阵, P是m阶可逆方阵, Q是n阶可逆方阵, 则

$$r(A)=r(PA)=r(AQ)=r(PAQ),$$

即可逆矩阵与任何矩阵乘积的秩等于该矩阵的秩.

由于矩阵可以看作基下的变换, 因此秩就是**变换后的空间维度**. **列空间**是**变换后基的坐标**, 因此秩也相当于变换后列空间的维数

矩阵秩的另一个极端是1. 一个非零矩阵A的秩为1 $\iff A$是一个非零列矩阵与一个非零行矩阵的乘积, 即存在列向量α,β使得$A=\alpha\beta^T$. 因此秩为1的方阵的高次幂可以如下算出:

$$A^m=\left(\alpha\beta^T\right)^m=\left(\beta^T\alpha\right)^{m-1}\alpha\beta^T=\left(\beta^T\alpha\right)^{m-1}A \tag{35}$$

矩阵的和与乘积的秩

定理 1.1.1 (1) $r(A+B)\leq r(A)+r(B)$;
(2)(Sylvester不等式) 设A,B分别为$m\times p$, $p\times n$矩阵, 则

$$r(A)+r(B)-p\leq r(AB)\leq\min\{r(A),r(B)\}. \tag{1.1.2}$$

余子式 代数余子式 伴随

对任意n阶方阵 $A=(a_{ij})$, 去掉第i行第j列后所剩余的$n-1$阶方阵的行列式称为元素a_{ij} 的**余子式**, 记为M_{ij}. 而$(-1)^{i+j}M_{ij}$称为元素a_{ij}的**代数余子式**, 记为A_{ij}. n阶方阵

$$\begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \cdots & \ldots & \ldots & \cdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix} \tag{36}$$

称为方阵A的**伴随矩阵**, 记为$adjA$

$$A(\text{adj}\,A)=(\text{adj}\,A)A=|A|I \tag{37}$$

分块矩阵

分块对角矩阵 **直和**

$$\begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_s \end{pmatrix} = A_1 \oplus A_2 \oplus \cdots \oplus A_s = \sum_{i=1}^{s} \oplus A_i,$$

线性方程组与 n维线性空间 F^n

线性方程组

如果线性 方程组$x_1\alpha_1 + x_2\alpha_2 + \cdots + x_s\alpha_s = 0$仅有零解, 则称向量组S是线性无关的. 否则就称S是 **线性相关** 的

极大线性无关组: 齐次线性方程组的解都可以由线性无关的$\alpha_1, \alpha_2, \cdots, \alpha_s$的线性表示

$$k_1\alpha_1 + k_2\alpha_2 + \cdots + k_s\alpha_s \tag{38}$$

齐次线性方程组$Ax = 0$的解集的一个极大线性无关组称为该方程组的一个 **基础解系**

定理 1.2.2 (齐次线性方程组的基本定理) 齐次线性方程组$A_{m\times n}x = 0$的任何一个基础解系恰含$n - r(A)$个解向量, 它的全体解为:

$$x = c_1\alpha_1 + c_2\alpha_2 + \cdots + c_{n-r}\alpha_{n-r}$$

其中$c_1, c_2, \cdots, c_{n-r}$是任意常数, $\alpha_1, \alpha_2, \cdots, \alpha_{n-r}$是该方程组的一个基础解系.

定理 1.2.3 (线性方程组的基本定理) 设线性方程组$A_{m\times n}x = b$有解, 则它的全体解为:

$$x = x_0 + c_1\alpha_1 + c_2\alpha_2 + \cdots + c_{n-r}\alpha_{n-r}$$

其中x_0是其任意一个解(称为**特解**), $c_1, c_2, \cdots, c_{n-r}$是任意常数, $\alpha_1, \alpha_2, \cdots, \alpha_{n-r}$是$Ax = 0$的一个基础解系.

初等变换

1. (重排变换) 交换第i行(列)与第j行(列)的初等矩阵为$I - E_{ii} - E_{jj} + E_{ij} + E_{ji}$;
2. (倍乘变换) 给第i行(列)乘以非零数a的初等矩阵为$I + (a-1)E_{ii}$;
3. (倍加变换) 将第j行(列)的a倍加到第i行(列)的初等矩阵为$I + aE_{ij}$.

矩阵的一次行初等变换相当于左乘一个初等矩阵, 而矩阵的一次列初等变换相当于右乘一 个初等矩阵.

目的: 将矩阵化为**Hermite型**

定义 1.2.3 (Hermite 标准形) 设$m \times n$矩阵H的秩为r且满足以下条件:

(1) 它的非零行恰为前r行, 且这r行的第一个非零元(称为该行的先导元素)为 1;

(2) 非零行的先导元素的列标随行标严格递增; 即若设第k行的先导元素出现在第j_k列, 则$j_1 < j_2 < \cdots < j_r$;

(3) 非零行的先导元素所在的列的其它位置的元素为零, 即第j_k列为标准单位列向量e_k;

则称H是 Hermite 标准形或简化行阶梯形.

若$H_A = PA$, 则$Ax = b$与$H_A x = Pb$同解

满秩分解

定义 1.2.4 设$A \neq 0$. 如果列满秩矩阵L与行满秩矩阵R使得$A = LR$, 则称$A = LR$是A的一个满秩分解.

可用于计算小秩矩阵的高次幂

不唯一, 矩阵L与R的秩显然等于矩阵A的秩.

Proof:

$$r(L_{m\times p}) + r(R_{p\times n}) - p \le r(A_{m\times n}) \le \min\{r(L), r(R)\} \tag{39}$$

由于L列满秩, R行满秩, 秩均为p

$$\begin{gathered} p \le r(A) \le p \\ r(A) = p \end{gathered} \tag{40}$$

计算方法:

设矩阵$A = (a_{ij})_{m\times n}$经过行初等变换化为 Hermite 标准形H_A. 由H_A的定义, 可设

$$H_A = \begin{pmatrix} H_r \\ 0 \end{pmatrix},$$

其中$r = r(A)$, H_r为$r \times n$的行满秩的 Hermite 标准形矩阵. 设H_r的第k行的先导元素所在的列标为j_k, 则$j_1 < j_2 < \cdots < j_r$, 且H_r的$j_1, j_2, \cdots, j_r$列组成一个r阶单位矩阵, 故A的$j_1, j_2, \cdots, j_r$列线性无关, 且

$$A = (A_{j_1}, A_{j_2}, \cdots, A_{j_r})H_r. \tag{1.2.2}$$

则公式(1.2.2)是A的一个满秩分解.

例 1.2.4 设A是秩为 1 的n阶方阵, 试求$A^n (n \ge 1)$.

解 设$A = LR$为A的满秩分解, 其中L, R分别为列, 行向量. 则RL为一阶方阵, 且$RL = \operatorname{tr}(RL) = \operatorname{tr}(LR) = \operatorname{tr} A$. 从而当$n \ge 1$时, 有

$$A^n = (LR)^n = L(RL)^{n-1}R = (\operatorname{tr} A)^{n-1}LR = (\operatorname{tr} A)^{n-1}A.$$

特征值与矩阵的相似对角化

对于大秩特别是满秩矩阵求高次幂, 使用对角化方法

满秩矩阵无效. 因此需要更为普遍的方法, 这就是对角化方法. 确切地说, 设A是n阶复矩阵, 若存在对角矩阵$D = \operatorname{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$与可逆矩阵$P = (\alpha_1, \alpha_2, \cdots, \alpha_n)$使得$A = PDP^{-1}$, 则称$A$与$D$相似, 同时称$A$可以相似对角化或简称为可以对角化. (一般称矩阵$P^{-1}AP$与矩阵$A$是相似的矩阵.) 如此则有$A^m = PD^mP^{-1}$, 因此$A$的任意高次幂可以较方便地求出. 由于$AP = PD$, 故$A\alpha_i = \lambda_i\alpha_i$, $1 \le i \le n$. 所以α_i是齐次线性方程组$Ax = \lambda_i x$的非零解, 即数λ_i与非零向量α_i均满足齐次线性方程组

$$Ax = \lambda x \tag{1.3.1}$$

可知, 相似对角矩阵D的对角线元素即为A的特征值.

特征值

几何重数

$n - r(\lambda_i I - A)$为特征值 λ_i的几何重数. 根据定理1.2.2 (齐次线性方程组的基本定理), 特征方程 $(\lambda_i I - A)x = 0$ 的基础解系秩为$n - r(\lambda_i I - A)$, **即为对应特征向量的个数**

代数重数

设$\sigma(A) = \{\lambda_1, \lambda_2, \cdots, \lambda_s\}$, 且

$$|\lambda I - A| = \prod_{i=1}^{s}(\lambda - \lambda_i)^{n_i} \tag{1.3.2}$$

则正整数n_i称为特征值λ_i的**代数重数**.

谱半径

特征值的最大模称为A的**谱半径**, 记为ρ(A)

$$\rho(A) = \max\{|\lambda| : \lambda \in \sigma(A)\} \tag{41}$$

从几何上看, 矩阵A的特征值全部位于以原点为圆心, 谱半径ρ(A)为半径的圆盘内.

特征值性质

定理 1.3.1 (特征值的性质) (1) 矩阵的行列式等于其所有特征值的积, 即$|A|=\prod\limits_{i=1}^{s}(\lambda_i)^{n_i}$;

(2)矩阵的迹等于其所有特征值的和, 即$\operatorname{tr}(A)=\sum\limits_{i=1}^{s}n_i\lambda_i$;

(3) A可逆 $\Longleftrightarrow$ 0不是A的特征值;

(4)设$f(x)$是任意多项式, λ是A的一个特征值, α是属于λ的特征向量, 则$f(\lambda)$是$f(A)$的一个特征值, α是属于$f(\lambda)$的特征向量;

(5)设A可逆且其特征多项式为(1.3.2), 则其逆矩阵的特征多项式为

$$|\lambda I-A^{-1}|=\prod_{i=1}^{s}(\lambda-\lambda_i^{-1})^{n_i},$$

且若α是A的属于特征值λ的特征向量, 则α也是A^{-1}的属于特征值λ^{-1}的特征向量;

(6)任何特征值的几何重数不超过其代数重数;

(7)相似矩阵具有相同的特征多项式(因此具有相同的特征值).

特征向量的性质

定理 1.3.2 (特征向量的性质) (1)属于不同特征值的特征向量线性无关;

(2) n阶矩阵A可以对角化 $\Longleftrightarrow$ A有n个线性无关的特征向量 $\Longleftrightarrow$ $\mathbb{F}^n$有一组由A的特征向量组成的基.

对角化主定理

定理 1.3.3 (对角化主定理) 一个n阶矩阵A可以对角化 $\Longleftrightarrow$ A 的每个特征值的代数重数与几何重数相等. 特别, 若A有n个不同的特征值, 则A可以对角化.

$$A=PDP^{-1} \tag{42}$$

这很好理解, 代数重数表示特征值重根的次数, 几何重数则是其对应的线性无关的特征向量的个数, 如果二者相等, 即说明 A 具有 n 个线性无关的特征向量, 这样 P 才能是可逆矩阵, 对角化才存在.

线性空间

考察n维线性空间$\mathbb{F}^n$可知, 其运算加法“$+$”满足下面五个条件:
(C) **封闭性**: 对任意$\alpha,\ \beta\in\mathbb{F}^n$, $\alpha+\beta\in\mathbb{F}^n$;
(A1) **结合律**: 对任意$\alpha,\ \beta,\ \gamma\in\mathbb{F}^n$, $(\alpha+\beta)+\gamma=\alpha+(\beta+\gamma)$;
(A2) **交换律**: 对任意$\alpha,\ \beta\in\mathbb{F}^n$, $\alpha+\beta=\beta+\alpha$;
(A3) 存在**零向量**: 即存在元素0, 使得对任意$\alpha\in\mathbb{F}^n$, $\alpha+0=\alpha$;
(A4) 存在**负向量**: 对任意$\alpha\in\mathbb{F}^n$, 存在一个向量, 记为$-\alpha$, 使得$\alpha+(-\alpha)=0$.

F^n中的数乘满足下列四个条件(以下将F^n记为V):

(B1) 数乘的结合律: 设$a, b \in \mathbb{F}$, $\alpha \in V$, 有$a(b\alpha) = (ab)\alpha$;
(B2) 数乘关于向量加法的分配律: 设$a \in \mathbb{F}$, $\alpha, \beta \in V$, 有$a(\alpha+\beta) = a\alpha + a\beta$;
(B3) 数乘关于数的加法的分配律: 设$a, b \in \mathbb{F}$, $\alpha \in V$, 有$(a+b)\alpha = a\alpha + b\alpha$;
(B4) 数乘的初始条件: $1 \bullet \alpha = \alpha$, 其中$1 \in \mathbb{F}$.

线性空间定义

定义 1.4.1 设$(V, +)$是一个加群, 如果定义了数域$\mathbb{F}$中的数与V中元素(称为向量)的数乘(记为$\bullet$, 一般省略不写), 则称$(V, +, \bullet)$是数域$\mathbb{F}$上的**线性空间**(或**向量空间**), 简称V是($\mathbb{F}$)线性空间[5]. 特别, 当$\mathbb{F} = \mathbb{R}$或$\mathbb{C}$时, V称为实空间或复空间. 一般地, 数域$\mathbb{F}$称为线性空间V的**基域**. 本书除特别指明, 所有线性空间的基域均是指数域$\mathbb{F}$(所以我们经常略去定语“数域$\mathbb{F}$上的”), 而将$\mathbb{F}$中的元素统称为“数”.

过渡矩阵

矩阵$P = (p_{ij})$称为由α–基到β–基的**过渡矩阵**, 显然, 此时由β–基到α–基的过渡矩阵 为P^{-1}

$$\begin{cases} \beta_1 = p_{11}\alpha_1 + p_{21}\alpha_2 + \cdots + p_{n1}\alpha_n \\ \beta_2 = p_{12}\alpha_1 + p_{22}\alpha_2 + \cdots + p_{n2}\alpha_n \\ \cdots \\ \beta_n = p_{1n}\alpha_1 + p_{2n}\alpha_2 + \cdots + p_{nn}\alpha_n \end{cases} \tag{43}$$

$$(\beta_1, \beta_2, \cdots, \beta_n) = (\alpha_1, \alpha_2, \cdots, \alpha_n) P \tag{44}$$

坐标变换公式

$$\begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{pmatrix} = P \begin{pmatrix} y_1 \\ y_2 \\ \cdots \\ y_n \end{pmatrix}, \quad 或 \quad \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = P^{-1} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

例 1.4.4 求任意$f(x)\in\mathbb{F}[x]_n$在标准基$1,x,x^2,\cdots,x^{n-1}$下的坐标. 向量组$1,x-a,(x-a)^2,\cdots,(x-a)^{n-1}$也构成$\mathbb{F}[x]_n$的一组基, 求$f(x)$在该组基下的坐标. 求从标准基到第二组基的过渡矩阵.

解 由高等数学或数学分析中的Taylor[6] (泰勒)公式可知:

$$f(x)=\sum_{k=0}^{n-1}\frac{f^{(k)}(0)}{k!}x^k,$$

即$f(x)$在标准基下的坐标为

$$(f(0),f'(0),\cdots,\frac{f^{(k)}(0)}{k!},\cdots,\frac{f^{(n-1)}(0)}{(n-1)!})^T.$$

对任意$a\in\mathbb{F}$, 仍由Taylor 公式 可知, $f(x)$在第二组基下的坐标为

$$(f(a),f'(a),\cdots,\frac{f^{(k)}(a)}{k!},\cdots,\frac{f^{(n-1)}(a)}{(n-1)!})^T.$$

由于

$$(x-a)^k=\sum_{r=0}^{k}C_k^r x^r(-a)^{k-r},$$

故从标准基到第二组基的过渡矩阵$P=(p_{ij})$由下式给出:

$$p_{ij}=\begin{cases}C_{j-1}^{i-1}(-a)^{j-i} & i\leq j,\\ 0 & i>j.\end{cases}$$

内积空间与正定二次型

并非每个矩阵都可以对角化, 但任何实对称矩阵却可以以更精细的方式对角化

内积空间

定义

定义 1.5.1 设$\mathbb{F}$是实数域或复数域, V是$\mathbb{F}$上的线性空间. 若对V中任意两个向量α, β, 都定义了$\mathbb{F}$中一个数(α,β)(称为向量α与β的内积), 使得

(1) (共轭对称性) $(\alpha,\beta)=\overline{(\beta,\alpha)}$, 其中$\overline{(\beta,\alpha)}$是复数$(\beta,\alpha)$的共轭复数;

(2) (正定性) $(\alpha,\alpha)\geq 0$, 且等号成立 $\Longleftrightarrow \alpha=0$;

(3) (双线性) $(a\alpha+b\beta,\gamma)=a(\alpha,\gamma)+b(\beta,\gamma)$, 对任意$\alpha,\beta,\gamma\in V,\ a,b\in\mathbb{F}$成立;

则称V为一个**内积空间**.

请注意, 由于内积的共轭对称性, 内积的第三个条件“双线性”仅对第一个变量成立, 而对第二个变量是“共轭双线性”的, 即

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

有限维实内积空间又称为欧几里得空间或**欧氏空间**; 当F = C是复数域, 则内积空间称为**酉空间**(或复内积空间).

内积和范数的性质

命题 1.5.1 (内积与范数的性质) (1) $(a\alpha,\, b\beta) = a\bar{b}(\alpha, \beta)$;
(2) $(\alpha+\beta, \gamma) = (\alpha, \gamma) + (\beta, \gamma)$;
(3) $(0, \alpha) = (\alpha, 0) = 0$;
(4) $\|c\alpha\| = |c|\|\alpha\|$;
(5) (Cauchy-Schwarz [8]不等式)$|(\alpha\,, \beta)| \le \|\alpha\|\|\beta\|$且等号成立 $\Longleftrightarrow$ α与β线性相关;
(6) (三角不等式) $\|\alpha+\beta\| \le \|\alpha\| + \|\beta\|$.

由Cauchy-Schwarz不等式即可一般地定义“角度”, 即设α与β的夹角为θ, 则

$$\theta = \arccos \frac{|(\alpha\,, \beta)|}{\|\alpha\|\|\beta\|} \tag{1.5.2}$$

对内积空间V 的任意一组基$\alpha_1, \alpha_2, \cdots, \alpha_n$, 利用`Gram-Schmidt`正交化方法可求得V 的 一个标准正交基.

投影向量

注意, 如果β是一个非零向量, α是任意向量, 则向量$\frac{(\alpha,\beta)}{\|\beta\|}\beta$恰好是向量$\alpha$在向量$\beta$上的**投影向量**, 一般记为$\mathrm{Proj}_\beta\alpha$. 特别, 若$\beta$是一个单位向量, 则$\mathrm{Proj}_\beta\alpha = (\alpha, \beta)\beta$. 因此, Gram-Schmidt 正交化方法的几何意义是: 设$\alpha_1, \alpha_2 \in \mathbb{R}^3$ 线性无关, 则向量$\alpha \in \mathbb{R}^3$在α_1, α_2确定的平

面上的投影等于它在该平面的一组正交坐标轴上的投影之和. 比如, 公式(1.5.7)中的被减各项恰好是α_k在诸γ_j $(1 \le j \le k-1)$上的投影.

酉矩阵

$Q^*Q = I$

设$\alpha_1, \alpha_2, \cdots, \alpha_n$是$V = \mathbb{R}^n$或$V = \mathbb{C}^n$的一组标准正交基, 则矩阵

$$Q = (\alpha_1, \alpha_2, \cdots, \alpha_n)$$

称为**酉矩阵**. 实的酉矩阵称为**正交矩阵**.

定理 1.5.2 (酉矩阵的刻画) 矩阵Q是酉矩阵 $\Longleftrightarrow$ $Q^*Q = I$. 特别, 实矩阵A是正交矩阵 $\Longleftrightarrow$ $Q^TQ = I$ $\Longleftrightarrow$ $Q^{-1} = Q^T$.

酉矩阵的逆矩阵就是其共轭转置矩阵!

Herimite矩阵

复共轭对称矩阵 $A^* = A$

在第二章我们将看到，当$n=2$时，正交矩阵的几何意义几乎就是平面解析几何中的旋转变换. 由于可以利用旋转变换将中心在原点的有心二次曲线(即椭圆和双曲线)变成标准形式(这是化简二次型的范例)，故利用正交矩阵就可以将实对称矩阵对角化. 需要注意的是，与实对称矩阵相应的复数矩阵不是复对称矩阵，而是复共轭对称矩阵，习惯上称为 **Hermite 矩阵**，即需要满足$A^* = A$. Hermite 矩阵是最重要的一类矩阵，比如量子力学中的物理量均是由 Hermite 矩阵来表达的，而酉空间中的内积则由所谓"正定" Hermite 矩阵完全确定(见下面的定理 1.5.6).

定理 1.5.3 Hermite 矩阵的特征值均为实数，且属于不同特征值的特征向量彼此正交(证明见习题 20).

定理 1.5.4 (Hermite 矩阵的酉对角化) Hermite 矩阵A可以酉对角化，即存在酉矩阵U使得$U^*AU = D$是对角矩阵. 特别，实对称矩阵可以正交对角化.

正定

定义 1.5.3 设$f(x) = x^*Ax$是复二次型，A是 Hermite 矩阵，若对任意非零向量$\alpha \in \mathbb{C}^n$，均有$f(\alpha) = \alpha^*A\alpha > 0$，则称$f(x)$是正定二次型，$A$是正定矩阵. 类似地可以定义半正定，负定，半负定二次型，半正定，负定，半负定矩阵等. 对实二次型和实对称矩阵的定义类似.

Hermite矩阵性质

定理 1.5.5 设A是n阶 Hermite 矩阵，则下列条件等价:
(1) A是正定的;
(2) $f(x) = x^*Ax$是正定二次型;
(3) A的特征值均为正实数;
(4) 存在$m \times n$阶列满秩矩阵M使得$A = M^*M$;
(5) 存在n阶可逆矩阵M使得$A = M^*M$;
(6) 存在n阶可逆矩阵P使得$P^*AP = I$(即A与I合同).

内积与正定矩阵

定理 1.5.6 (内积与正定矩阵) 设V是n维复线性空间，则其上的内积与正定矩阵一一对应. 确切地说，设(α, β)是V上的二元向量函数，则(α, β)是内积 $\Longleftrightarrow$ 存在正定 Hermite 矩阵$A = (a_{ij})$使得 $(\alpha, \beta) = y^*Ax$，其中x与y分别是α与β在某组基下的坐标.

有限维实或复线性空间上的任意 两个内积**本质上是完全相同的**!

第二章 矩阵与线性变换

子空间: 直和与空间分解

线性映射

线性映射$\sigma : x \to Ax$应满足: (加法, 数乘封闭)

(T1) (可加性) $\sigma(x+y)=\sigma(x)+\sigma(y)$;

(T2) (齐次性) $\sigma(ax)=a\sigma(x)$, $a\in\mathbb{F}$.

子空间

定义 2.1.1 (子空间) 设V是一个线性空间, U是V的一个非空子集. 如果U本身关于V的向量加法与数乘作成一个线性空间, 则称U是V的一个线性子空间或子空间.

判别: (检验加法, 数乘封闭性)

定理 2.1.1 (子空间判别法) 设U是线性空间V的一个非空子集. 则U是子空间 $\iff$ 对任意$\lambda\in\mathbb{F}$, $\alpha,\beta\in U$, 有$\alpha+\beta\in U$与$\lambda\alpha\in U$.

注意: 子空间必然包含0向量

子空间的和

一般来说, 两个子空间的并集$U\cup W$不再是子空间 因为有可能不是**凸集**? 所以说空间应该是**凸集**

那么包含U与W的最小的子空间 应为U与W的**凸包** 记作 $U+W$, 称为U与W的**和**

维数定理

设V是线性空间，U与W是V的两个子空间，则

$$dim(U+W)=dimU+dimW-dim(U\cap W) \tag{45}$$

Proof:

取$U\cap W$的一组基$\alpha_1,\alpha_2,\cdots,\alpha_r$. 由于$U\cap W$是$U$与$W$的公共子空间, 故$U\cap W$的基是$U$与$W$的线性无关的向量组, 因此可以扩充成$U$或$W$的基. 设

$$\alpha_1,\alpha_2,\cdots,\alpha_r,\beta_{r+1},\beta_{r+2},\cdots,\beta_s$$

与

$$\alpha_1,\alpha_2,\cdots,\alpha_r,\gamma_{r+1},\gamma_{r+2},\cdots,\gamma_t$$

分别是U与W的基. 则

$$\alpha_1,\alpha_2,\cdots,\alpha_r,\beta_{r+1},\beta_{r+2},\cdots,\beta_s,\gamma_{r+1},\gamma_{r+2},\cdots,\gamma_t$$

是$U+W$的一组基. (为此只需证明该向量组线性无关, 且$U+W$的任何向量均可由这些向量线性表示.) $\square$

直和

欲使子空间U+W的维数最大，必须且只要$U\cap W=0$，亦即U和W重合的部分最小，为零空间。这时称U+W是**直和**，记为$U\oplus W$。因此

$$dim(U \oplus W) = dimU + dimW \tag{46}$$

判别:

定理 2.1.3 (直和的判定) 设U与W是线性空间V的两个子空间, 则下列命题等价:

(1) $U+W$是直和(即$U \cap W = 0$);

(2) 对任意$\alpha \in U+W$, 分解式$\alpha = u+w$, 其中$u \in U,\ w \in W$是唯一的, 即若还有$\alpha = u'+w'$, 则$u=u',\ w=w'$;

(3) 零向量的分解式唯一; 即若$0 = u+w,\ u \in U,\ w \in W$, 则$u=w=0$;

(4) $\dim(U+W) = \dim U + \dim W$.

补空间

$V = U \oplus W$，则W是U的补子空间。 补空间不唯一（反例：二维空间中两个相交在原点的向量）

1. 基的扩充定理
2.
3. 如果是直和，则任意$\gamma \in U+W$可以唯一由U和W中的两个元素表示，并且零向量也表示唯一（反证法），这三个命题等价。

四个子空间

1. 零空间$N(A)$：齐次线性方程组$Ax=0$的解空间
2. 列空间$R(A)$（像空间、值域）：**A的列向量生成的子空间, 以列向量中的极大线性无关组为基**
3. 行空间$R(A^T)$：**A的行向量生成的子空间, 以行向量中的极大线性无关组为基**
4. 左零空间$N(A^T)$：线性方程组$y^TA=0$或$A^Tx=0$的解空间

注意, $N(A)$与$R(A^T)$是$\mathbb{F}^n$的子空间; 而$R(A)$与$N(A^T)$是$\mathbb{F}^m$的子空间, 且有

$$\dim N(A) + \dim R(A^T) = n; \qquad \dim N(A^T) + \dim R(A) = m.$$

矩阵的秩就是行空间和列空间的秩, 因此, 其解空间(即为零空间) 秩为 $n - r(A)$**.**

计算方法:

计算矩阵的四个子空间的方法如下:

首先将A通过行初等变换化为其Hermite标准形H_A. 因为$Ax=0$等价于$H_Ax=0$, 所以$N(A)=N(H_A)$, 同时, 由于行初等变换不改变列向量之间的线性关系, 从而由H_A的列向量的极大线性无关组可得到A的列向量的极大线性无关组, 这就是$R(A)$的一组基; 其次, H_A的行是A的行的线性组合, 于是A的行空间等于H_A的行空间, 即$R(A^T)=R(H_A^T)$. 但由H_A不能同时得出A的左零化空间. 为此, 需要记录将A化为Hermite标准形时的可逆矩阵P(参看下例), 即$PA=H_A$. 设$r(A)=r$, 则H_A的最后$m-r$行均为0, 由矩阵乘法的行向量结构可知, 矩阵P的最后$m-r$行是齐次线性方程组$y^TA=0$的线性无关解, 因此它们恰好是A的左零化空间$N(A^T)$的一组基.

例 2.1.13 设

$$A=\begin{pmatrix}1&1&2\\0&1&1\\1&2&3\end{pmatrix},$$

求A的四个相关子空间.

解 将A化为Hermite标准形, 并记录相应的可逆矩阵P:

$$(A,I_3)=\begin{pmatrix}1&1&2&1&0&0\\0&1&1&0&1&0\\1&2&3&0&0&1\end{pmatrix}\longrightarrow\begin{pmatrix}1&0&1&1&-1&0\\0&1&1&0&1&0\\0&0&0&-1&-1&1\end{pmatrix}=(H_A,P),$$

由此可得$PA=H_A$, 即

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

从而矩阵A的秩$=2$, 而H_A的前两列线性无关, 所以A的前两列线性无关, 即

$$\begin{aligned}&R(A)=\mathrm{Span}\{A_1,A_2\}=\mathrm{Span}\{(1,0,1)^T,(1,1,2)^T\};\\&R(A^T)=R(H_A^T)=\mathrm{Span}\{(H_A^1)^T,(H_A^2)^T\}=\mathrm{Span}\{(1,0,1)^T,(0,1,1)^T\};\\&N(A)=N(H_A)=\mathrm{Span}\{(-1,-1,1)^T\}.\end{aligned}$$

最后, 为了求出$N(A^T)$, 可考察矩阵P, 它的最后一行左乘A为零向量, 故它是$xA=0$的解. 于是

$$N(A^T)=\mathrm{Span}\{(P^3)^T\}=\mathrm{Span}\{(-1,-1,1)^T\}.$$

线性变换

线性变换应满足可加性和齐次性，U到V的线性变换全体记为$Hom(U,V)$

1. **核：**“零点”集$\{\alpha\in U:\sigma(\alpha)=0\}$，记作$Ker(\sigma)$或$\sigma^{-1}(0)$
2. **像：**“函数值”的集合$\{\alpha\in V:\exists\beta\in U,\text{使得}\ \alpha=\sigma(\beta)\}$，记作$Im(\sigma)$或$\sigma(U)$
3. $Ker\sigma$和$Im\sigma$分别是U和V的子空间，其维数记作$\eta(\sigma)$与$r(\sigma)$, 称为 σ 的**零度与秩** **理解：** 线性变换$\sigma=Ax$对应矩阵A，零变换对应零矩阵，恒等变换对应恒等矩阵；变换的核与矩阵零空间$N(A)$有关，像与矩阵的列空间$R(A)$。当选取标准基时，线性变换几乎就是左乘一个矩阵A，进一步，设$A=(a_{ij})$，则

$$\begin{aligned}\sigma(\alpha)&=A(x_1,x_2,\ldots,x_n)^T\\&=(a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n,\ldots,a_{n1}x_1+a_{n2}x_2+\cdots+a_{nn}x_n)^T\\&=(\sigma_1(\alpha),\ldots,\sigma_n(\alpha))\end{aligned}$$

   即线性变换不过是多个多元线性函数
4. **不同基下的矩阵的关系：** 设α和β是U的两组基，P是由$\alpha-$基到$\beta-$基的过渡矩阵；设α'和β'是V的两组基，Q是由$\alpha'-$基到$\beta'-$基的过渡矩阵；设$\sigma\in Hom(U,V)$关于$\alpha-$基和$\alpha'-$基的矩阵为A，关于$\beta-$基和$\beta'-$基的矩阵为B：

$$B=Q^{-1}AP \tag{47}$$

假定U=V且$\alpha-$基等于$\alpha'-$基，$\beta-$基等于β'基，于是:

$$B=P^{-1}AP \tag{48}$$

则A与B**相似**, 即它们不过是同一线性变换在不同基下的矩阵而已, 而矩阵P不过是两个基之间的过渡矩阵.

$$\begin{tikzcd}
(\alpha-\text{基})U \arrow[r, "\sigma", "A"'] \arrow[d, "P^{-1}"'] & V(\beta-\text{基}) \arrow[d, "Q^{-1}"] \\
(\alpha'-\text{基})U \arrow[r, "B", "\sigma"'] & V(\beta'-\text{基})
\end{tikzcd}$$

下面介绍一些特殊的线性变换:

(1) 零变换: 将线性空间V的所有向量均变为$0\,(\in V)$向量的变换, 称为零变换, 记为0(更确切地, 0_V); 即对任意$\alpha\in V$, 有$0(\alpha)=0$.

(2) **恒等变换**: 将线性空间V中任意向量均变为自己的变换, 称为恒等变换或单位变换, 记为I_V(或简单地记为I或1); 即对任意$\alpha\in V$, $I(\alpha)=\alpha$. 恒等变换显然是V到自身的同构(称为V的**自同构**).

(3) **位似**: 设$k\in\mathbb{F}$. 将线性空间V的任意向量α变为$k\alpha$的变换σ称为(伸缩)系数为k的位似, 即$\sigma(\alpha)=k\alpha$. 零变换与恒等变换分别是$k=0$与$k=1$时的位似. 非零位似均是自同构.

(4) **可逆变换**: 设线性变换$\sigma\in\mathrm{Hom}(U,V)$, 如果存在线性变换$\tau\in\mathrm{Hom}(V,U)$使得对任意$\alpha\in U$, $\beta\in V$, 均有$\tau(\sigma(\alpha))=\alpha$, $\sigma(\tau(\beta))=\beta$, 则称$\sigma$是可逆线性变换, τ称为其逆变换. 可以证明, 如果σ的逆变换存在则必定唯一, 其逆记为σ^{-1}, 且$(\sigma^{-1})^{-1}=\sigma$(习题16).

比如, 任何系数为$k(k\neq 0)$的位似均是可逆的, 且其逆变换是系数为k^{-1}的位似. 又如, 例2.2.1中的旋转变换就是线性空间$\mathbb{R}^2$的可逆线性变换, 其逆变换就是顺时针旋转θ角度或逆时针旋转$-\theta$角度.

同构

$\sigma:U\to V$ 既是单射又是满射；两个空间不仅元素对应，运算也对应；并且**维度相同**

定理 2.2.5 (同构定理) 域$\mathbb{F}$上的两个线性空间U与V同构$\iff \dim_{\mathbb{F}}U=\dim_{\mathbb{F}}V$.

同构映射相当于重起名字, 本质上没有改变空间结构

矩阵与线性变换

定理 2.2.6 (矩阵与线性变换) 设V是$\mathbb{F}$上的n维线性空间, $\alpha_1, \alpha_2, \cdots, \alpha_n$是$V$的一组基. 设$M_n(\mathbb{F})$是$\mathbb{F}$上全体$n$阶矩阵组成的线性空间. 对任意$\sigma \in \operatorname{End} V$, 记$A(\sigma)$是$\sigma$在该基下的矩阵. 定义$\operatorname{End} V$到$M_n(\mathbb{F})$的映射$\psi$为

$$\begin{array}{cccc} \psi: & \operatorname{End} V & \longrightarrow & M_n(\mathbb{F}) \\ & \sigma & \longmapsto & A(\sigma) \end{array}$$

则ψ是一个保持运算(加法, 数乘与乘法)的一一映射, 即满足下列条件:

(1) $A(\sigma+\tau) = A(\sigma) + A(\tau)$;
(2) $A(k\sigma) = kA(\sigma), \forall k \in \mathbb{F}$;
(3) $A(\sigma\tau) = A(\sigma)A(\tau)$;
(4) σ可逆 $\Longleftrightarrow$ $A(\sigma)$可逆; 且此时$A(\sigma)^{-1} = A(\sigma^{-1})$;
(5) $A(0) = 0$; $A(I) = I$.

内积空间的正交分解

正交补

在内积空间中, 每个子空间除去有无穷多个补子空间外, 还有唯一一个"**好的**"补子空间, 即所谓 **正交补**.

定理 2.3.1 设V是n维内积空间, U是V的子空间. 则$V = U \oplus U^{\perp}$. 特别, $\dim U^{\perp} = n - \dim U$.

定理 2.3.2 (1) $N(A) = R(A^*)^{\perp}$; (2) $R(A) = N(A^*)^{\perp}$.

投影向量

是β在v上的投影向量$Proj_v \beta$

$$\hat{\beta} = \frac{(\beta, v)}{(v, v)} v \tag{49}$$

最佳近似

定义 2.3.1 (最佳近似) 设U是欧氏空间V的子空间, $\beta \in V$. 如果$\alpha \in U$满足下面的条件:

$$\|\beta - \alpha\| \leq \|\beta - \gamma\|, \quad \forall \gamma \in U.$$

则称α是β在U中的**最佳近似(向量)**, 常记为$\hat{\beta}$(如果不需要强调子空间U).

β在子空间U中的最佳近似(向量)就是U中与β最近(距离最短)的向量, 这样的向量一定存在而且唯一

定理 2.3.3 (最佳近似定理) 设U是欧氏空间V的子空间, $\beta \in V$, $\alpha \in U$. 则α是β在U上的最佳近似向量 $\Longleftrightarrow$ $\beta - \alpha \in U^{\perp}$.

求解:

求向量β在子空间U中的最佳近似向量, 等价于求β在直和分解$V = U \oplus U^{\perp}$下的表达式, 即若$\beta = \alpha + \gamma$, 其中$\alpha \in U, \gamma \in U^{\perp}$, 则$\alpha$就是$\beta$在$U$上的最佳近似向量 $\hat{\beta}$.

命题 2.3.1 设β是欧氏空间V的一个向量, U是V的一个子空间, $\alpha_1, \cdots, \alpha_s$是$U$ 的一个正交基, 则β在U上的最佳近似向量为

$$\hat{\beta} = \mathrm{Proj}_{v_1}\beta + \cdots + \mathrm{Proj}_{v_s}\beta = \frac{(\beta, \alpha_1)}{(\alpha_1, \alpha_1)}\alpha_1 + \cdots + \frac{(\beta, \alpha_s)}{(\alpha_s, \alpha_s)}\alpha_s. \tag{2.3.5}$$

特别, 若$\alpha_1, \cdots, \alpha_s$还是$U$的一个标准正交基, 则$\beta$在$U$上的最佳近似向量为

$$\hat{\beta} = (\beta, \alpha_1)\alpha_1 + (\beta, \alpha_2)\alpha_2 + \cdots + (\beta, \alpha_s)\alpha_s. \tag{2.3.6}$$

根据命题2.3.1即可知欧氏空间中的向量β在任何非零子空间U中的投影向量$\mathrm{Proj}_U\beta$为其最佳近似向量, 即

$$\mathrm{Proj}_U\beta = \hat{\beta}.$$

而命题2.3.1可以重新表述为"向量β在子空间U中的最佳近似就是其在U上的投影向量, 就是其在U的一组正交基的各个基向量上的投影向量之和".

此处应注意: 需要时正交基

正规化方程

对于矛盾方程来说, 对任意 $x \in \mathbb{R}^n$, $Ax \neq b$, 要求其最优解 Ax^0 即是求 b 在矩阵 A 的值空间 $R(A)$ 中的最佳近似向量. 由最佳近似定理可知, $b - Ax^0 \in R(A)^{\perp} = N(A^T)$. 因此向量 $b - Ax^0$ 满足 $A^T(Ax^0 - b) = 0$, 即 x^0满足:

$$A^TAx = A^Tb \tag{50}$$

即为**正规化方程**, 经过变换后,方程总是相容的, 最优解总是存在.

命题 2.3.2 方程组$A^TAx = A^Tb$与方程组$Ax = \mathrm{Proj}_{R(A)}b$同解, 即方程组$Ax = b$的最优解就是方程组$Ax = \mathrm{Proj}_{R(A)}b$的解.

对于非正交基的问题, 通过求解正规化方程来解最佳近似向量

内积空间中的线性变换

本节探讨 线性变换是否会影响长度, 角度和距离

等距变换/正交变换

定义 2.4.1 设V是内积空间, $\sigma \in \mathrm{End}\,V$. 如果$\sigma$保持向量间的距离, 即对任意$\alpha, \beta \in V$, 均有$d(\sigma(\alpha), \sigma(\beta)) = d(\alpha, \beta)$, 则称$\sigma$是**等距变换** 或保距变换.

等距变换必然是可逆变换. 恒等变换是等距变换; $\mathbb{R}^2$中的旋转也是等距变换; 但比例系数的模不等于1 的位似变换均非等距变换.

定理 2.4.1 设V是内积空间, $\sigma \in \mathrm{End}\,V$. 则$\sigma$是等距变换 $\Longleftrightarrow$ σ保持向量的长度 $\Longleftrightarrow$ σ保持内积.

在标准正交基下的等距变换 $\Longleftrightarrow$ 该标准正交基下的变换矩阵为酉矩阵

因此欧氏空间的**等距变换**又称为 **正交变换**；而复内积空间的等距变换也称为 **酉变换**

但是等距变换在一般基下的矩阵未必是酉矩阵

定理 2.4.3 设V是内积空间, $\sigma \in \operatorname{End} V$. 则$\sigma$是等距变换 $\Longleftrightarrow$ σ将标准正交基变为标准正交基.

等距变换保持长度和内积, 从而也保持向量之 间的角度, 因此等距变换不改变图形的形状, 但一般的线性变换则可能将图形变得面目全非

Householder变换

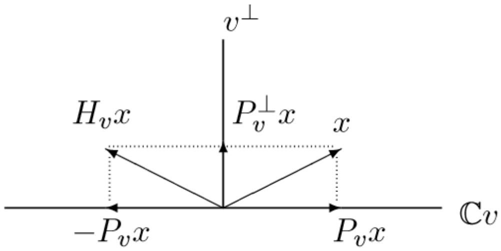
例 2.4.5 (Householder 变换) 设$v \in \mathbb{C}^n$是非零向量, 定义n阶复矩阵H如下:

$$H = I - \frac{2vv^*}{v^*v}. \tag{2.4.3}$$

矩阵H称为 **Householder 矩阵**. 由矩阵H定义的线性变换H_v称为 Householder 变换, 即对任意$x \in \mathbb{C}^n$

$$H_v x = x - \frac{2vv^*}{v^*v}x. \tag{2.4.4}$$

容易看出, Householder变换H_v实际上是关于超平面$v^\perp$的镜面反射



Givens变换

例 2.4.7 (Givens 变换) 由矩阵(2.4.1)决定的旋转称为 Jacobi [21] 旋转或(2 维) Givens 旋转. 一般地, 设$c^2+s^2=1, \theta = \arctan\frac{s}{c}$, 则将$n$阶实矩阵

$$G(i,\, j,\, \theta) = I_n - (1-c)(E_{ii}+E_{jj}) + s(E_{ij}-E_{ji}) \tag{2.4.5}$$

称为 Givens 旋转矩阵(感兴趣的读者可写出该矩阵的形状). 这是一个正交矩阵(习题 44).

对称变换

定义 2.4.2 设σ是欧氏空间V的线性变换. 如果对任意$\alpha, \beta \in V$, 均有

$$(\sigma(\alpha), \beta) = (\alpha, \sigma(\beta)) \tag{2.4.6}$$

则称σ是一个**对称变换**.

定理 2.4.4 欧氏空间的线性变换σ是对称变换 $\Longleftrightarrow$ σ在一组标准正交基下的矩阵是对称矩阵.

正交分解

例 2.4.10 (正交分解) 由定理2.3.1, 任何向量$x \in V$可以唯一分解为如下形式

$$x = P_U x + P_U^{\perp} x,$$

其中$P_U x \in U$, $P_U^{\perp} x \in U^{\perp}$分别为$x$在$U$上的**投影向量** 和**正交投影向量**. 因此, 可将P_U与$P_U^{\perp}$分别理解为内积空间V的向子空间U与$U^{\perp}$ 的**正交投影变换**, 即

$$P_U : x \mapsto P_U x$$

而

$$P_U^{\perp} : x \mapsto P_U^{\perp} x.$$

显然有$P_U^{\perp} = P_{U^{\perp}}$以及$P_U + P_U^{\perp} = I$.

一般, 如果$U = \mathrm{Span}\{v\}$是1维子空间, 则将下标U改记为v, 如$P_U x$记为$P_v x$等.

第三章 特征值与矩阵的Jordan标准形

Schur 三角化定理

Schur酉三角化定理

定理 3.1.1 (Schur (酉)三角化定理) 设$A \in \mathbb{C}^{n\times n}$, 则存在酉矩阵$U$使

$$U^* A U = B, \tag{3.1.1}$$

其中B为一个上三角矩阵.

Schur三角化定理是化简矩阵的基础, 用于计算方阵的任意次幂 A^m

对于任意复方阵,都能找到一个酉矩阵将其化为上三角矩阵

引理 3.1.1 设$A = (a_{ij})$是$\mathbb{F}$上的任意上三角矩阵, $1 \le p < q \le n$. 设$P = I + \alpha E_{pq}$, $\alpha \in \mathbb{F}$, 则$B = (b_{ij}) = P^{-1}AP$是与A的主对角线相同(包括顺序)的上三角矩阵, 且$b_{pq} = a_{pq} + \alpha(a_{pp} - a_{qq})$. 特别地, 若$a_{pp} \ne a_{qq}$, 则可选取适当的$\alpha$使得$b_{pq} = 0$.

分块Schur 三角化定理

将寻求矩阵标准形的问题化为寻找主对角元素均相同的上三角矩阵的标准形

定理 3.1.2 (分块Schur三角化定理) 设n阶复矩阵A的特征多项式$|\lambda I - A| = \prod_{i=1}^{s}(\lambda - \lambda_i)^{n_i}$, 其中$\sigma(A) = \{\lambda_i \,|\, 1 \le i \le s\}$, $\sum_{i=1}^{s} n_i = n$. 则存在可逆矩阵P使得

$$P^{-1}AP = A_1 \oplus A_2 \oplus \cdots \oplus A_s,$$

其中A_i是特征值均为λ_i的n_i阶上三角矩阵.

设A是n阶严格上三角矩阵 (对角线也为0) , 则$A^n = 0$.

Cayley-Hamilton

因此, 可以证明`Cayley-Hamilton`定理, 即设矩阵A的特征多项式为$f(\lambda)$, 则有$f(A) = 0$.

$$f(A) = (A - \lambda_1 I)^{n_1}(A - \lambda_2 I)^{n_2} \cdots (B - \lambda_s I)^{n_s} \tag{51}$$

Sylvester降幂公式

命题 3.1.1 (Sylvester 降幂公式) 设A与B分别是$m \times n$与$n \times m$矩阵, $m \geq n$. 则

$$|\lambda I_m - AB| = \lambda^{m-n}|\lambda I_n - BA|.$$

最小多项式 零化多项式

定义 3.1.1 设A是n阶矩阵, $f(x)$是多项式. 如果$f(A) = 0$, 则称$f(x)$是A的**零化多项式**. 次数最低的首一零化多项式称为A的**最小多项式**, 记为$m_A(x)$或$m(x)$.

性质

1. 设$m(x)$是A的最小多项式, $f(x)$是A的任意零化多项式, 则$m(x)|f(x)$. 特别地, $m(x)||xI - A|$.
2. 设A是任意方阵, $m(x)$是A的最小多项式. 设$\lambda_0 \in C$. 则λ_0是A的特征值 $\Longleftrightarrow m(\lambda_0) = 0$.
3. 矩阵A是幂零矩阵 $\Longleftrightarrow$ A的特征值均为0.
4. **分块对角矩阵的最小多项式等于各个子块的最小多项式的最小公倍式(对照: 分块对角矩阵的特征多项式等于各个子块的特征多项式的乘积)**.
5. 相似矩阵具有相同的最小多项式.
6. **矩阵A可以对角化 ⟺ A的最小多项式没有重根.**
7. 设A为方阵, f(x)是无重因式的多项式. 若f(A) = 0, 则A可以对角化.

线性变换的特征值与特征向量

特征向量是在线性变换下不改变方向的那些非零向量, 而特征值则是特征向量的伸缩系数. 注意线性变换在其每个特征子空间上的限制是一个位似, 因此特征值与特征向量具有简化线性变换的矩阵表示的作用. 确切地说, 如果有限维线性空间 V 有一组基其基向量全部为σ的特征向量, 则σ在这组基下的矩阵是对角矩阵

定理 3.1.5 设V是有限维线性空间, $\sigma \in \mathrm{End}V$, A是σ在某组基下的矩阵. 则
(1) A与σ有完全相同的特征值(即重数也一样);
(2) 设σ的不同特征值为$\lambda_1, \cdots, \lambda_s$, 则$A$可以对角化 $\Longleftrightarrow$ $V = \sum_{i=1}^{s} \oplus V_{\lambda_i}$. 于是$A$可以对角化 $\Longleftrightarrow$ σ可以对角化, 即$\sigma = \sum_{i=1}^{s} \oplus \sigma_i$, 其中$\sigma_i$是$\sigma$在$V_{\lambda_i}$上的限制.

因此, 可以将矩阵的特征值, 特征向量求解问题转化为空间线性变换

Jordan 标准形

根据分块Schur三角化定理, 可将矩阵转化为

$$A = \lambda I + N \tag{52}$$

N为严格上三角, 因为

$$P^{-1}AP = \lambda I + P^{-1}NP \tag{53}$$

故只需研究严格上三角的标准形. 即为Jordan 标准形 J_n

$$J_n = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

命题 3.2.1 (1) $J_n^T J_n = (0) \oplus I_{n-1}$, $J_n J_n^T = I_{n-1} \oplus (0)$;
(2) $J_n e_i = e_{i-1}$, 其中e_i是标准向量, 约定e_0是 0 向量;
(3) 对任意$x \in \mathbb{C}^n$, $(I_n - J_n^T J_n)x = (x^T e_1)e_1$.

幂零矩阵的Jordan标准形定理

设A是n阶严格上三角复矩阵. 则存在可逆矩阵P和正整数$n_1 \geq n_2 \geq \cdots \geq n_m, n_1 + n_2 + \cdots + n_m = n$, 使得

$$P^{-1}AP = J_{n_1} \oplus J_{n_2} \oplus \cdots \oplus J_{n_m} \tag{54}$$

上式右端的矩阵称为 **幂零Jordan矩阵**, 它是矩阵A的Jordan标准形, 一般记为J. 如果不计诸n_j的次序与大小, 则A的Jordan标准形是唯一的.

幂零矩阵相似准则

设M与N是两个n阶幂零矩阵. 则M与N相似 $\iff r(M^k) = r(N^k), \forall k \geq 1$

矩阵$\lambda I_n + J_n$称为n阶$\lambda - Jordan$块, 记为$J_n(\lambda)$.

Jordan标准形定理

设A是n阶严格上三角复矩阵. 则存在可逆矩阵P和正整数$n_1 \geq n_2 \geq \cdots \geq n_m, n_1 + n_2 + \cdots + n_m = n$, 使得

$$P^{-1}AP = J_{n_1}(\lambda_1) \oplus J_{n_2}(\lambda_2) \oplus \cdots \oplus J_{n_m}(\lambda_m) \tag{55}$$

上式右端的矩阵称为**Jordan矩阵**, 它是矩阵A的**Jordan标准形**. 如果不计诸nj的次序与大小, 则A的Jordan标准形是唯一的.

方阵A可以对角化 $\iff$ A的Jordan标准形是对角矩阵

但是根据定义求特征值再求Jordan形太麻烦, 具体计算如下

Jordan标准形计算

1. 对阶数较小的幂零矩阵可以通过其幂零指数和秩来确定其Jordan标准形: **最大Jordan块**
2. 对高阶矩阵, 需要更加精细的方法:

最大Jordan块

设A是严格上三角矩阵, 则其Jordan标准形的**Jordan块的阶数的最大值等于其幂零指数**.

高阶矩阵方法

定理 3.3.1 设n阶严格上三角矩阵A的Jordan标准形J同公式(3.2.1), 其幂零指数为e. 则
(1) $e = \max\{n_i \,|\, 1 \le i \le m\}$;
(2) J中Jordan块的个数m等于A的零度;
(3) 记J中k阶Jordan块的个数为ℓ_k, A^k的零度为η_k, $0 \le k \le m$. 则

$$\ell_k = 2\eta_k - \eta_{k-1} - \eta_{k+1}, \quad 1 \le k \le m. \tag{3.3.1}$$

注 1. 计数公式(3.3.1)来源于数列

$$\eta_0 = 0, \eta_1, \eta_2, \cdots, \eta_{k-1}, \eta_k, \eta_{k+1}, \cdots$$

的二阶差分, 即$\ell_k = (\eta_k - \eta_{k-1}) - (\eta_{k+1} - \eta_k)$.

注 2. 上面定理中的(2)等价于说J中Jordan块的个数等于A的线性无关的特征向量的个数. 这是因为每个Jordan块的秩恰好比其阶数小1, 即零度为1, 因此对应于1个线性无关的特征向量.

盖尔圆定理

盖尔圆用于特征值估计

定理 3.4.1 (盖尔圆盘定理[33]) 设A是n阶矩阵, 则它的特征值λ至少满足下列不等式之一:

$$|\lambda - a_{ii}| \le \sum_{j \ne i}^{n} |a_{ij}|, \quad i = 1, 2, \cdots, n.$$

等价地, $\sigma(A) \subseteq G(A)$, 即A的每个特征值都落在A的某个圆盘之内.

精细圆盘定理

设C是盖尔区域的一个由k个圆盘组成的连通分支, 则C恰好含有k个特征值.

第四章 正规矩阵和矩阵分解

矩阵分解是用于简化矩阵计算, 快速计算矩阵高次幂

我们首先探讨一种特殊的"好矩阵"——**正规矩阵**. 由于根据Schur三角化定理, 任何一个矩阵都可以**酉三角化**, 因此可以**酉对角化**的才算是好矩阵.

正规矩阵

定理4.1.1 矩阵 A可以 **酉对角化** $\Longleftrightarrow AA^* = A^*A$

定义4.1.1 设 $A \in C^{n\times n}$，若 $AA^* = A^*A$，则称 A为 **正规矩阵**

1. 实对称: $A = A^T$
2. 实反对称: $A = -A^T$ (**注意**: 反对称阵必定对角线全为0)
3. 正交: $AA^T = I \Longleftrightarrow A^T = A^{-1}$
4. Hermite: $A = A^*$ (即复对称矩阵)
5. 反Hermite: $A = -A^*$ (即复反对称矩阵)
6. 酉矩阵: $AA^* = I \Longleftrightarrow A^* = A^{-1}$ (即复正交矩阵)

引理4.1.1 设 A 为正规矩阵，若 A 又为三角矩阵，则 A 为对角矩阵

定理4.1.2 设 $A \in C^{n\times n}$，则 A 为正规矩阵 $\Longleftrightarrow$ A 有n个两两正交的单位特征向量

推论4.1.1 正规矩阵属于不同特征值的特征向量是相互正交的

定理4.1.3 设 $A = (a_{ij})_{n\times n}$ 是复矩阵，$\lambda_1, \lambda_2, \ldots, \lambda_n$为 A的 n 个特征值，则

1. （**Schur不等式**）$\sum_{i=1}^{n}|\lambda_i|^2 \leqslant \sum_{i,j=1}^{n}|a_{ij}|^2$
2. A为正规矩阵 $\Longleftrightarrow \sum_{i=1}^{n}|\lambda_i|^2 = \sum_{i,j=1}^{n}|a_{ij}|^2$

> Proof:
>
> 根据Schur定理, $U^*AU = B$, B为上三角
>
> 1. 若 A为正规矩阵, B为对角矩阵, 且$U^*AA^*U = BB^*$
>
> $$\begin{aligned} tr(AA^*) &= \sum_{i,j=1}^{n}|a_{ij}|^2 \\ tr(BB^*) &= \sum_{i=1}^{n}|\lambda_i|^2 \\ tr(BB^*) &= tr(AA^*) \\ So \qquad \sum_{i=1}^{n}|\lambda_i|^2 &= \sum_{i,j=1}^{n}|a_{ij}|^2 \end{aligned}$$

矩阵分解

- 正交三角分解
- 三角分解（方阵）
- 谱分解（必须是可对角化的方阵）
- 奇异值分解（极分解）
- 满秩分解

> 化为Hermite型，P9 定义1.2.4

正交三角分解(UR分解)

> 一种对任何可逆矩阵均存在的理想分解，其原理是Gram-Schmidt的正交化方法

定理 4.4.1 设$A \in C^{n \times n}$, 且A为满秩的, 则存在**唯一**的**酉矩阵U**和**对角线元素都大于零的上三角矩阵R**满足

$$A = UR \tag{56}$$

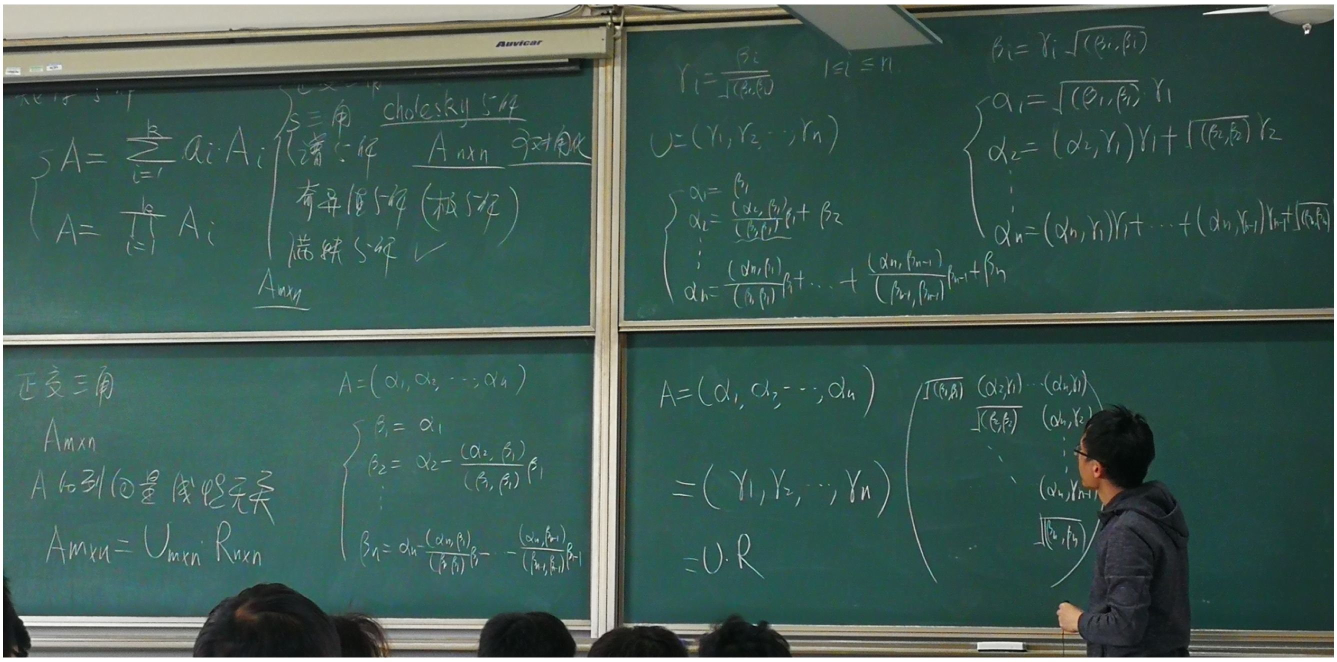
要求与性质

1. $A_{m \times n}$，不一定为方阵
2. A 的列向量线性无关，**列满秩**
3. $A_{m \times n} = U_{m \times n} R_{n \times n}$，则 U 的列向量构成一组标准正交向量组，R为对角线元素大于零的上三角

4. 此分解**唯一**
5. 复空间上为 UR 分解, 实空间为 QR分解

定理证明（P136）

利用施密特正交化，再单位化



证 设α_j为A的第j个列向量，$j=1,2,\cdots,n$. 则$A=(\alpha_1,\alpha_2,\cdots,\alpha_n)$. 因$A$满秩，$\alpha_1$, α_2, $\cdots$, α_n线性无关. 我们知道由下述 Gram-Schmidt 正交化过程可以将α_1, α_2, $\cdots$, α_n化成两两正交的向量η_1, η_2, $\cdots$, η_n:

$$\begin{cases}\eta_1=\alpha_1,\\ \eta_2=\alpha_2-\dfrac{(\alpha_2,\eta_1)}{(\eta_1,\eta_1)}\eta_1,\\ \eta_3=\alpha_3-\dfrac{(\alpha_3,\eta_1)}{(\eta_1,\eta_1)}\eta_1-\dfrac{(\alpha_3,\eta_2)}{(\eta_2,\eta_2)}\eta_2,\\ \cdots\cdots\cdots\cdots\\ \eta_n=\alpha_n-\dfrac{(\alpha_n,\eta_1)}{(\eta_1,\eta_1)}\eta_1-\dfrac{(\alpha_n,\eta_2)}{(\eta_2,\eta_2)}\eta_2-\cdots-\dfrac{(\alpha_n,\eta_{n-1})}{(\eta_{n-1},\eta_{n-1})}\eta_{n-1}.\end{cases}$$

于是

$$\begin{cases}\alpha_1=\eta_1,\\ \alpha_2=\dfrac{(\alpha_2,\eta_1)}{(\eta_1,\eta_1)}\eta_1+\eta_2,\\ \alpha_3=\dfrac{(\alpha_3,\eta_1)}{(\eta_1,\eta_1)}\eta_1+\dfrac{(\alpha_3,\eta_2)}{(\eta_2,\eta_2)}\eta_2+\eta_3,\\ \cdots\cdots\cdots\cdots\\ \alpha_n=\dfrac{(\alpha_n,\eta_1)}{(\eta_1,\eta_1)}\eta_1+\dfrac{(\alpha_n,\eta_2)}{(\eta_2,\eta_2)}\eta_2+\cdots+\dfrac{(\alpha_n,\eta_{n-1})}{(\eta_{n-1},\eta_{n-1})}\eta_{n-1}+\eta_n.\end{cases}$$

令$\beta_i=\frac{1}{\sqrt{(\eta_i,\eta_i)}}\eta_i$, 则$U=(\beta_1,\beta_2,\cdots,\beta_n)$为酉阵, 且

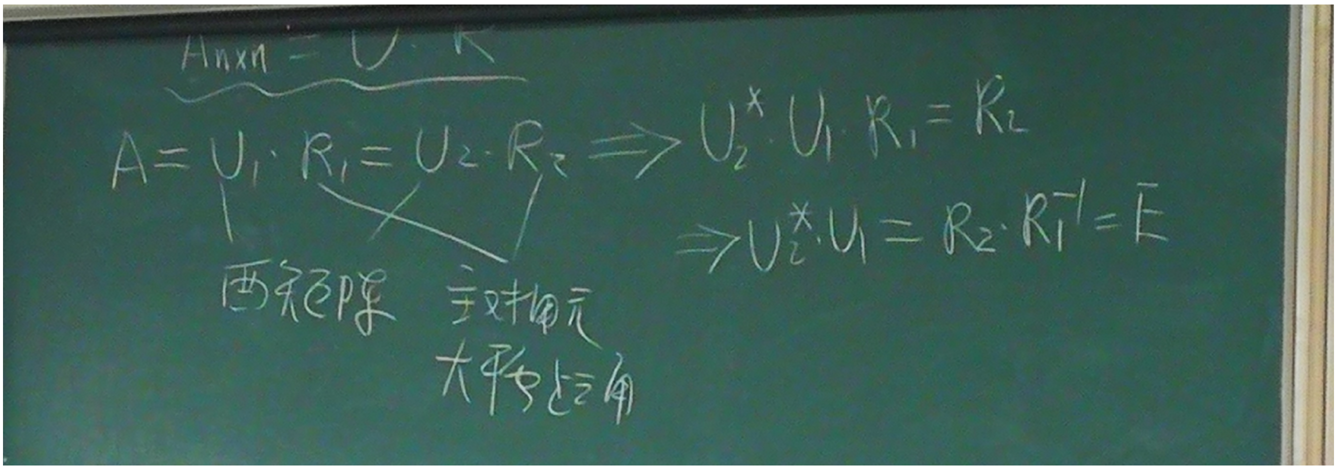
$$\begin{cases}\alpha_1=\sqrt{(\eta_1,\eta_1)}\beta_1,\\ \alpha_2=(\alpha_2,\beta_1)\beta_1+\sqrt{(\eta_2,\eta_2)}\beta_2,\\ \alpha_3=(\alpha_3,\beta_1)\beta_1+(\alpha_3,\beta_2)\beta_2+\sqrt{(\eta_3,\eta_3)}\beta_3,\\ \cdots\cdots\cdots\cdots\\ \alpha_n=(\alpha_n,\beta_1)\beta_1+(\alpha_n,\beta_2)\beta_2+\cdots+(\alpha_n,\beta_{n-1})\beta_{n-1}+\sqrt{(\eta_n,\eta_n)}\beta_n.\end{cases}$$

因而$A=UR$, 其中$U=(\beta_1,\beta_2,\cdots,\beta_n)$, 以及

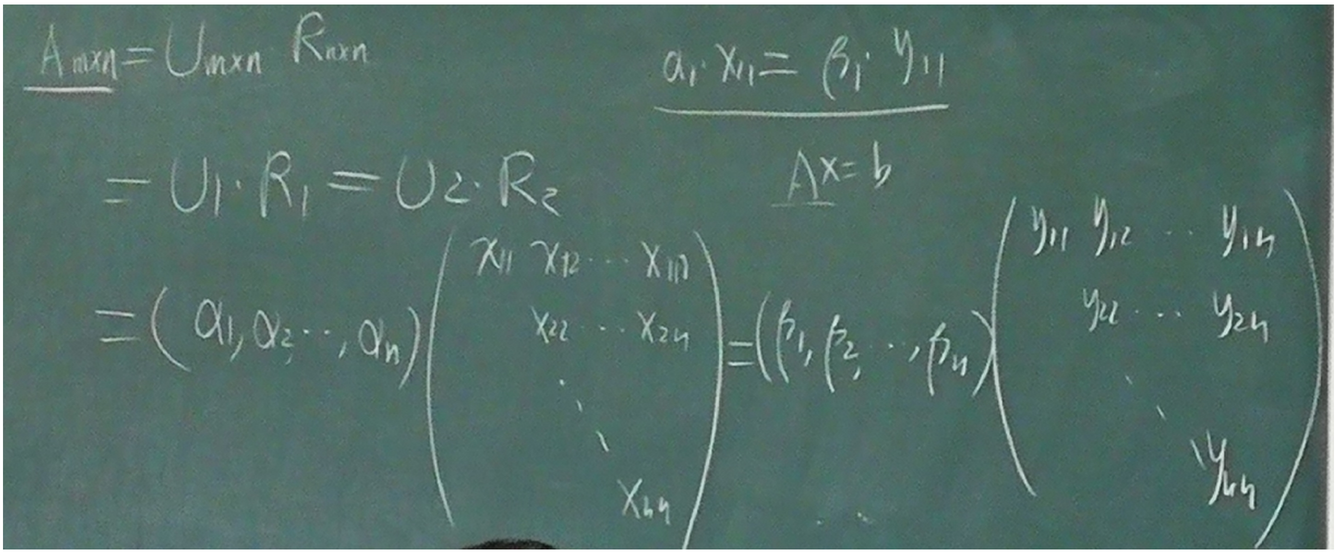
$$R=\begin{pmatrix}\sqrt{(\eta_1,\eta_1)} & (\alpha_2,\beta_1) & \cdots & (\alpha_{n-1},\beta_1) & (\alpha_n,\beta_1)\\ & \sqrt{(\eta_2,\eta_2)} & \cdots & (\alpha_{n-1},\beta_2) & (\alpha_n,\beta_2)\\ & & \ddots & \vdots & \vdots\\ & & & \sqrt{(\eta_{n-1},\eta_{n-1})} & (\alpha_n,\beta_{n-1})\\ & & & & \sqrt{(\eta_n,\eta_n)}\end{pmatrix}.$$

唯一性证明（P137）

反证法，首先证明方阵



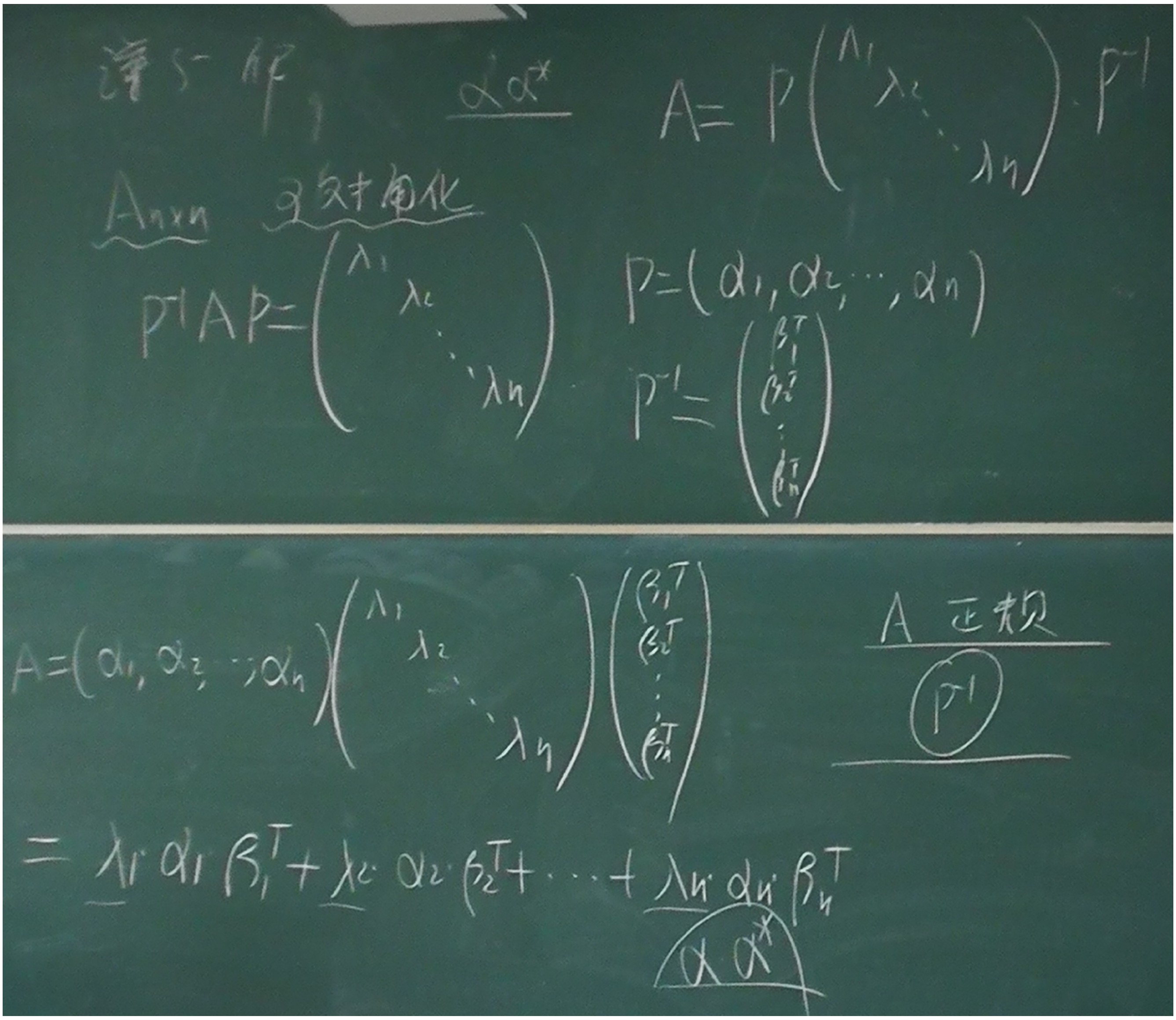
非方阵证明



谱分解或特征值分解

要求与性质

1. $A_{n\times n}$ 方阵
2. 可对角化 $P^{-1}AP = diag(\lambda_1, \ldots, \lambda_n)$



计算方法

1. 计算特征值
2. 计算**单位正交特征向量** (因为是正规矩阵, 因此必存在n个正交特征向量)
3. 求正交投影矩阵 $P_i = \alpha_i \alpha_i^*$
4. $A = \sum_{i=1}^{s} \lambda_i P_i$