Advanced Industrial Mathematics
Fall 2025

Ju Ming

jming@hust.edu.cn
Huazhong University of Science and Technology

September 24, 2025

Chap 1: 线性空间和线性变换

Example 1.0.1

If $x, y \in \mathcal{Z}$,then $x + y \in \mathcal{Z}$;
If $c \in \mathcal{Z}$, then $cx \in \mathcal{Z}$;
If $c \notin \mathcal{Z}$, then in most case, $cx \notin \mathcal{Z}$;
Conclusion: 所有整数构成的集合$\mathcal{Z}$对于加法和(整数)数乘运算封闭

Example 1.0.2

If $\vec{x}, \vec{y} \in \mathcal{R}^3$,then $x + y \in \mathcal{R}^3$;
If $c \in \mathcal{R}$, then $c\vec{x} \in \mathcal{R}^3$;
If $c \in \mathcal{C}$, then $c\vec{x} \in \mathcal{C}^3$; but in most case, $c\vec{x} \notin \mathcal{R}^3$
Conclusion: 即三维实向量空间$\mathcal{R}^3$关于加法和(实数)数 乘运算封闭

1.1 线性空间

Definition 1.1.1

- V 是一个非空集合（如向量、函数集合等）
- F 是一个数域（常用实数域$\mathcal{R}$，或复数域$\mathcal{C}$）

对V 中任意两个元素α, β 定义如下运算：

- 加法运算：

$$\alpha + \beta \in V$$

  ($\alpha + \beta$称为 α与β的和)

- 数乘运算：

$$\lambda \in F, \lambda\alpha \in V$$

  ($\lambda\alpha$称为λ与α的数积)

线性运算满足的条件

- 加法交换律:$\forall\alpha,\beta\in V$, 有
$$\alpha+\beta=\beta+\alpha;$$
- 加法结合律：$\forall\alpha,\beta,\gamma\in V$,有
$$(\alpha+\beta)+\gamma=\beta+(\alpha+\gamma);$$
- 加法零元：$\exists 0(\text{零元})\in V$， $\forall\alpha\in V$, 都有
$$\alpha+0=\alpha$$
- 加法负元：$\forall\alpha\in V$, $\exists-\alpha\in V$, 使得
$$\alpha+(-\alpha)=0$$

线性运算满足的条件(cont'd)

- 乘法分配律 V:$\forall k \in F$ 和 $\forall \alpha, \beta \in V$，有
$$k(\alpha + \beta) = k\alpha + k\beta$$
- 乘法分配律F:$\forall k, l \in F$和 $\forall \alpha \in V$，有
$$(k + l)\alpha = k\alpha + l\alpha$$
- 乘法结合律:$\forall k, l \in F$和 $\forall \alpha \in V$，有
$$(kl)\alpha = k(l\alpha)$$
- 乘法数1：F中的数1,使得$\forall \alpha \in V$, 有
$$1\alpha = \alpha$$

Summary:

1. 加法交换律
2. 加法结合律
3. 加法零元
4. 加法负元
5. 乘法分配律V
6. 乘法分配律F
7. 乘法结合律
8. 乘法数1

如果V关于这两种运算（称为线性运算）封闭， 且满足条件(1)-(8)，那么称V 称为F上的线性空间（或向量空间），记为$V(F)$. 称V中的元素为向量..

- 当F为实数域$\mathcal{R}$ 时，对应的空间$V(\mathcal{R})$称为实线性空间;
- 当F为复数域$\mathcal{C}$ 时，对应的空间$V(\mathcal{C})$称为 复线性空间;

Example 1.1.2

易证:

- $V=F=\mathcal{R}$, 运算取实数间的加法和乘法,$V(F)$是1维实数空间;
- $V=C(\mathcal{R})$($\mathcal{R}$上的连续函数所构成的集合), 运算取一 般函数间的加法和数乘, 则$V(F)$ 是连续函数空间
- $F=\mathcal{R}, V=\mathcal{R}^n$ 加法和数乘取一般向量的加 法和数乘,$V(F)$是n维向量空间;

Example 1.1.3

令 $V \in R^n, \forall k, l \in R$, $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_n)^T, \beta = (\beta_1, \beta_2, \cdots, \beta_n)^T, \gamma = (\gamma_1, \gamma_2, \cdots, \gamma_n) \in R^n$,可以验证$R^n$满足向量空间的加法和乘法规律。

Example 1.1.4

实数域2上次数不超过$n-1$次的关于变量$x \in R$的一切多项式和零多项式所构成的集合

$$\{\sum_{i=0}^{n-1} a_i x^i, a_i \in R\}$$

在通常多项式加法与数乘多项式运算下构成线性空间，称为多项式空间$P_n[x]$

注：次数等于$n-1$ 的多项式集合$\{\sum_{i=0}^{n-1} a_i x^i, a_i \in R, a_{n-1} \neq 0\}$不是线性空间

Theorem 1.1.1

线性空间V有如下性质：

1. V 中的零元素惟一；
2. V 中任一元素的负元素惟一；
3. 设0 为数零，$\mathbf{0}$ 为V中零向量，则
   (i). $0 \cdot \alpha = \mathbf{0}$
   (ii). $k \cdot \mathbf{0} = \mathbf{0}, k \in F$
   (iii). $k \cdot \alpha = \mathbf{0}$, 则$k = 0$或者$\alpha = \mathbf{0}$
   (iv). $(-1)\alpha = -\alpha$

线性组合

Definition 1.1.5

若$\alpha_1, \cdots, \alpha_m \in V$, $k_1, k_2, \cdots, k_m \in F$， 则

$$\beta := k_1\alpha_1 + k_2\alpha_2 + k_m\alpha_m = \sum_{i=1}^{m} k_i\alpha_i$$

是$V(F)$ 中的元素，称为$\alpha_1, \cdots, \alpha_m$ 的一个线性组合，或称β 可由$\alpha_1, \cdots, \alpha_m$线性表出.

Example 1.1.6

$2t^2 + 3t + 4$是$t^2, t, 1$ 的线性组合,也是$t^2 + t, t + 4$的线性组合，但是不是$t, 1$的线性组合， 也不是t^2, t 的线性组合

线性相关与线性无关

Definition 1.1.7
V 中的向量组$\{\alpha_1, \cdots, \alpha_m, m \geq 1\}$ 称 为线性相关的，如果存在一组不全为零的数 $k_1, k_2, \cdots, k_m \in F$，使得

$$k_1\alpha_1 + k_2\alpha_2 + k_m\alpha_m = 0.$$

反之，若上述等式只在$k_1 = k_2 = \cdots = k_m = 0$（系数全为零）时成立，则称向量组$\{\alpha_1, \cdots, \alpha_m, m \geq 1\}$ 是线性无关的.

Theorem 1.1.2

- 当$m > 1$ 时，向量组$\{\alpha_1, \cdots, \alpha_m\}$线性相关的充要条件是，其中至少有一个向量 $\alpha_j (1 \leq j \leq m)$ 可由组中其他向量线性表出.
- 若某向量组线性无关, 则它的任一子向量组必线性无关; 而若某向量组中有一个子向量组线性相关, 那么该向量组必线性相关.
- 单个零向量组$\{0\}$ 是线性相关的, 但单个 非零向量组$\{\alpha, \alpha \neq 0\}$是线性无关的.

线性空间的维数

Definition 1.1.8

线性空间V中，称最大线性无关向量的个 数为V的维数, 记为$dim(V)$.

- 如果在V 中能找到无限多个线性无关的向 量，则称V是无限维的；
- 如果在V 中只能找到最多n 个(n 有限) 线 性无关的向量，则称V 是n 维的； 此时n 维线性空间V 记作V^n.

Example 1.1.9

3 维向量组成的线性空间R^3, 线性无关向量 $e_1 = (1, 0, 0)^T$, $e_2 = (0, 1, 0)^T$, $e_3 = (0, 0, 1)^T$, 且任一向量 $(\alpha_1, \alpha_2, \alpha_3)^T \in R^3$ 可由它们线性表出.

Example 1.1.10

$m \times n$矩阵组成的线性空间$F^{m\times n}$ 是$m \times n$ 维的，$F^{m\times n}$ 中的任一矩阵$A = [a_{ij}]$ 可表示为

$$A = [a_{ij}] = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} E_{ij}$$

其中E_{ij} 定义为第i行第j列元素为1，其余元素为0 的$m \times n$矩阵. 其中，$\{E_{ij}\}_{1\leq i\leq m;1\leq j\leq n}$ 这mn 个矩阵是线性无关的.

Example 1.1.11

全体多项式组成的空间$P(t)$是无限维的，因为其中 $\{1, t, t^2, \cdots, t^m, \cdots\}$ 无限多个都是线性无关的

Example 1.1.12

多项式空间$P_n[t]$是$n+1$维的，因为其中 有$\{1, t, t^2, \cdots, t^n\}$，这$n+1$ 个多项式是线性无 关的, 而且任一次数不大于n 的多项式都可 由这$n+1$个多项式线性表出

- 若V 中能找到m 个线性无关的向量，则

$$dim(V) \geq m$$

- V 中任一向量都可以由m个向量线性表出，则

$$dim(V) \leq m$$

线性空间的基

Definition 1.1.13

V^n中给定顺序的n个线性无关的向量$\alpha_1, \cdots, \alpha_n$ 所组成的向量组称为V^n的一组基(或基底)，记为$\mathcal{B} = \{\alpha_1, \cdots, \alpha_n\}$. $\mathcal{B}$ 中的向量 α_i $(1 \le i \le n)$ 称为第i个基向量.

Example 1.1.14

- R^3中的一组基为$\mathcal{B} = \{e_1, e_2, e_3\} = \{(1,0,0)^T, (0,1,0)^T, (0,0,1)^T\}$
- $F^{m\times n}$中的一组基为

$$\mathcal{B} = \{E_{ij}, 1 \le i \le m, 1 \le j \le n\}$$

- 考虑$V = C, F = C$, 则$\{1\}$ 是一组基, 1 维.
  如果$V = C, F = R$, 则$\{1, i\}$ 是一组基, 2 维.

Theorem 1.1.3

设$\mathcal{B}$ 是V^n 的一组基，则V中任一向量都 可以由$\mathcal{B}$唯一线性表出.

Theorem 1.1.4

- 由于基就是向量集合V的极大线性无关组，从而线性空间的基也不是惟一的.
- n维线性空间中任意n个线性无关的向量构成的向量组都是空间的基.

坐标

Definition 1.1.15

设在V^n中取一组基$\mathcal{B}$，则V^n中任 一向量ξ, 都存在唯一的一组数$x_1; x_2, \cdots, x_n$, 使得

$$\xi = \sum_{i=1}^{n} x_i \alpha_i$$

写成矩阵形式

$$\xi = (\alpha_1, \alpha_2, \cdots, \alpha_n)(x_1, x_2, \cdots, x_n)^T$$

- 矩阵$(\alpha_1, \alpha_2, \cdots, \alpha_n)$仍然记为$\mathcal{B}$
- $X = (x_1, x_2, \cdots, x_n)^T$称为$\xi$在$\mathcal{B}$下的坐标向量（坐标）， x_i称为ξ在$\mathcal{B}$下的第i个坐标

Example 1.1.16

在$P_2(t)$中取$\mathcal{B} = \{1, t, t^2\}$和$\{t+1, t+2, t^2\}$,求$\xi = 2t^2 - t + 1$的坐标

Example 1.1.17

$R^{n\times n}$ 中的任一可逆矩阵P, 其n个列向量 $P_1, P_2, \cdots, P_n \in R^n$ 构成R^n 的基.

- $P_1, P_2, \cdots, P_n$ 线性无关;
- $\forall y \in R^n$, 则

$$y = P(P^{-1}y) = Px = \sum_{i=1}^{n} P_i x_i$$

即y 能被P 的列向量线性表出, 其中这里的坐标x 满足$x = P^{-1}y$.

注: 由R^n(或C^n) 的一组基作为列向量排列成的n阶方阵是可逆的.

- 引入坐标的原因：在V^n中取定一组基$\mathcal{B}$, 则存在一一对应：V^n 中任一元素$\alpha \Leftrightarrow \alpha$ 在$\mathcal{B}$ 下的坐标$x \in F^n$, 关于V^n 的问题可以转化为关于F^n 的问题,

Example 1.1.18

考虑$P_2(t)$ 的基$\mathcal{B} = \{t+1; t+2; t^2\}$.$P_1(t) = 2t^2 - t + 1$, $P_2(t) = t^2 - t + 1$, 考虑 $3P_1 + 2P_2$

Theorem 1.1.5

设$\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$是$n$维线性空间 $V^n(F$ 的一组基， $V^n(F)$中向量 β_i 在该基下坐标为 $X_i, i = 1, 2, 3, \cdots, m$，则$V^n(F)$中向量组 $\{\beta_1, \beta_2, \cdots, \beta_m\}$ 线性相关的充分必要条件是其坐标向量组 $\{X_1, X_2, \cdots, X_m\}$是 F^n 中的线性相关组.

两组基的变换关系

设$\mathcal{B}_\alpha = \{\alpha_1, \alpha_2, \cdots, \alpha_n\}$和$\mathcal{B}_\beta = \{\beta, \beta, \cdots, \beta_n\}$是 $V_n(F)$的两组基, 则由基的定义（空间中的任意向量可以由基底线性表出）,$\exists C_i \in R^n$, 使得

$$\beta_i = \sum_{k=1}^{n} c_{ki}\alpha_k = (\alpha_1, \alpha_2, \cdots, \alpha_n)\begin{pmatrix} c_{1i} \\ c_{2i} \\ \vdots \\ c_{ni} \end{pmatrix} = (\alpha_1, \alpha_2, \cdots, \alpha_n)C_i$$

$i = 1, 2, \cdots, n$. 考虑矩阵的分块计算，事实上有：

$$\beta = \alpha C,$$

这里 $\beta = (\beta_1, \beta_2, \cdots, \beta_n)$, $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_n)$,
$C = (C_1, C_2, \cdots, C_n)$

Definition 1.1.19

设$\mathcal{B}_\alpha = \{\alpha_1, \alpha_2, \cdots, \alpha_n\}$和$\mathcal{B}_\beta = \{\beta, \beta, \cdots, \beta_n\}$是 $V_n(F)$的两组基,若存在 $C \in F^{n\times n}$, 使得

$$\beta = \alpha C$$

, 则称 C是从基底$\mathcal{B}_\alpha$到基底$\mathcal{B}_\beta$的过渡矩阵（基变换矩阵）

Theorem 1.1.6

过渡矩阵是满秩矩阵，故 $\forall \xi \in V_n$, 设 $\xi = \mathcal{B}_\alpha x = \mathcal{B}_\beta y$, 其中$x$和$y$分别是$\xi$在两个基下面的坐标，则有

$$x = Cy\text{，或者 } y = C^{-1}x$$

Example 1.1.20

设R^3 的两组基为$\alpha_1 = (1, 0, -1)^T, \alpha_2 = (2, 1, 1)^T, \alpha_3 = (1, 1, 1)^T$ 和 $\beta_1 = (0, 1, 1)^T, \beta_2 = (-1, 1, 0)^T, \beta_3 = (1, 2, 1)^T$

- 求从基$\{\alpha_1, \alpha_2, \alpha_3\}$ 到基$\{\beta_1, \beta_2, \beta_3\}$的过渡矩阵$C$
- 求向量$\alpha = \alpha_1 + 2\alpha_2 - 3\alpha_3$ 在基 $\{\beta_1, \beta_2, \beta_3\}$ 下的坐标.

Example 1.1.21

设R^3 的两组基为$\alpha_1 = (1, 0, 0)^T, \alpha_2 = (0, 1, 0)^T, \alpha_3 = (0, 0, 1)^T$ 和 $\beta_1 = (-3, -7, 1)^T, \beta_2 = (3, 6, 1)^T, \beta_3 = (-2, -3, 2)^T$ 求在这两组基下坐标相同的所有向量.

Example 1.1.22

设$f_1=1+2x+4x^3, f_2=x+x^2+4x^3, f_3=1+x-3x^2, f_4=-2x^2+x^3$

- 求证： $\{f_1,f_2,f_3,f_4\}$是线性空间$P_4[x]$的一组基
- 求空间的基$\{1,x,x^2,x^3\}$到基$\{f_1,f_2,f_3,f_4\}$的过渡矩阵C,并求向量 $f=1+x+x^2+x^3$在基$\{f_1,f_2,f_3,f_4\}$下的坐标Y

Example 1.1.23

已知$P_2(x)$的两个基是 $\mathcal{B}_\alpha=\{2x+1,x+3,x^2+x\}$ 和 $\mathcal{B}_\beta=\{x+1,x+2,x^2\}$, 求$\mathcal{B}_\alpha$到$\mathcal{B}_\beta$的变换阵

子空间

Definition 1.1.24

设$V_n(F)$为线性空间，W是V的非空子集合．若W元素关于V中加法与数乘向量法运算也构成线性空间，则称W是V的一个子空间．也记作 $W \subset V$

注意子空间需要两点条件

- 集合$W \subset V$
- 运算对W而言也具有封闭性

Definition 1.1.25

- V和$\{0\}$都是V的子空间, 称为平凡子空间
- $\alpha_1, \alpha_2, \cdots, \alpha_r$, $1 \leq r \leq n$, 是V的r个线性无关向量，则集合

$$span\{\alpha_1, \alpha_2, \cdots, \alpha_r\} := \{\sum_{i=1}^{r} k_i\alpha_i, k_i \in F\}$$

是V的一个子空间，称为$\alpha_1, \alpha_2, \cdots, \alpha_r$张成的空间

Example 1.1.26

给定$A \in R^{m\times n}$, 则

$$N(A) := \{x \in R^n : Ax = 0\}$$
$$R(A) := \{y \in R^m : y = Ax, x \in R^n\}$$

分别是R^n和R^m的子空间，分别称为A的零空 间和列空间.(注: 验证子空间主要是证明线性运算的封闭性)

Example 1.1.27

$R^{n\times n}$中所有对称矩阵组成的集合

$$F := \{A \in R^{n\times n} : A^T = A\}$$

是$R^{n\times n}$的一个子空间,

Example 1.1.28

对于$C^{n\times n}$中矩阵A, 定义$A^H=\overline{A^T}=(\bar{A})^T$为$A$的共轭转置. 若在$C^{n\times n}$ 中$A=A^H$, 则称A为Hermite 矩阵, $C^{n\times n}$中所有Hermite 矩阵组成的集合

$$\{A\in C^{n\times n}:A^H=A\}$$

是$C^{n\times n}$的一个子空间. (类似$R^{n\times n}$中对称矩阵)

Theorem 1.1.7

设W是V^n的一个r维子空间, $\mathcal{B}_W = \{\alpha_1, \alpha_2, \cdots, \alpha_r\}$是$W$的一组基，则$\mathcal{B}_W$可以扩充为$V^n$的基，即在$V^n$中一定可以找到 $n-r$个向量$\{\alpha_{r+1}, \alpha_{r+2}, \cdots, \alpha_n\}$, 使 $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$是$V^n$的一组基.

子空间的运算:交 与 和

回忆集合的运算，考虑与子空间的运算的异同

Definition 1.1.29
设空间W_1和W_2是空间V的子空间，则W_1与W_2的“交”与“和”分别定义为：

$$W_1 \cap W_2 := \{x : x \in W_1 \text{ 并且 } x \in W_2\}$$

$$W_1 + W_2 := \{\xi_1 + \xi_2 : \text{ 其中 } \xi_i \in W_i, i = 1, 2\}$$

Theorem 1.1.8
设空间$W_i \subset V, i = 1, 2$,则 $W_1 \cap W_2$ 和 $W_1 + W_2$都是V的子空间

Theorem 1.1.9

设空间$W_i \subset V, i = 1, 2$,则

$$dim(W_1 + W_2) + dim(W_1 \cap W_2) = dim(W_1) + dim(W_2).$$

Example 1.1.30

设$\alpha_1 = (1, 1, 0, 0)^T, \alpha_2 = (0, 1, 1, 0)^T, \alpha_3 = (0, 0, 1, 1)^T, \alpha_4 = (1, 0, 0, 1)^T$, R^4的两个子空间分别是

$$W_1 = span\{\alpha_1, \alpha_2\}, W_2 = span\{\alpha_3, \alpha_4\},$$

求$W_1 + W_2$及$W_1 \cap W_2$的基和维数.

子空间的运算：和空间与直和

Definition 1.1.31

若$W_1 + W_2$ 中任一向量只能唯一的分解为 W_1中的一个向量与W_2中的一个向量之和, 则 $W_1 + W_2$ 称为W_1 与W_2 的直和, 记为$W_1 \oplus W_2$.

- 根据定义, W_1与W_2的直和是W_1 与 W_2的一个特殊和空间, 但和空间不一定是直和.
- 两个线性子空间的交、和与直和，可以推广 到多个子空间的交、和与直和.即

$$\cap_{i=1}^{n} W_i := W_1 \cap W_2 \cap \cdots \cap W_n$$

$$\sum_{i=1}^{n} W_i = W_1 + \cdots + W_n$$

$$\oplus_{i=1}^{n} W_i = W_1 \oplus W_2 \oplus \cdots \oplus W_n$$

Theorem 1.1.10
设W_1 与W_2 是V 的子空间， $W = W_1 + W_2$，则成立以下等价条件

- $W = W_1 \oplus W_2$
- $W_1 \cap W_2 = \{0\}$（W的和表达式唯一）
- 若$\xi_1 + \xi_2 = 0, \xi_1 \in W_1, \xi_2 \in W_2$,则$\xi_1 = \xi_2 = 0$（零向量表达式唯一）
- $dim(W_1 + W_2) = dim(W_1) + dim(W_2)$ *(维数公式)*

Theorem 1.1.11
设V_1 是V_n 的一个子空间, 则必存在V_n 的 子空间V_2, 使得

$$V_1 \oplus V_2 = V_n$$

.

Example 1.1.32

设I_r表示r阶单位矩阵,对n阶方阵 $A = \begin{pmatrix} I_r, 0 \\ 0, 0 \end{pmatrix}, B = \begin{pmatrix} 0, 0 \\ 0, I_{n-r} \end{pmatrix}$.
它们的列 空间为$R(A), R(B)$，证明：$R^n = R(A) \bigoplus R(B)$.

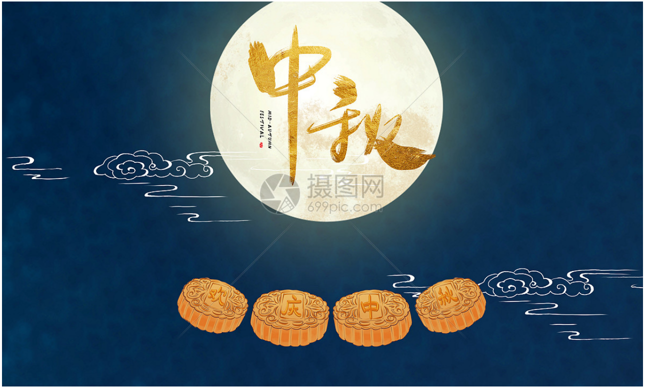

Figure: 中秋快乐

Exe: 矩阵论（杨）：p31， 1-13

1.2 内积空间

Definition 1.2.1

欧氏空间与酉空间 对数域F上的n维线性空间$V_n(F)$，定义的一个从$V_n(F)$中向量到数域F 的二元运算，记为(α,β)，即

$$(\alpha,\beta): V_n(F)\rightarrow F,$$

如果满足

- 对称性：$(\alpha,\beta)=(\beta,\alpha)$,
- 线性性： $(k\alpha,\beta)=k(\alpha,\beta)$, $(\alpha_1+\alpha_2,\beta)=(\alpha_1,\beta)+(\alpha_2,\beta)$
- 正定性: $(\alpha,\alpha)\geq 0$, $(\alpha,\alpha)=0$的充要条件是$\alpha=0$

则称(α,β)是$V_n(F)$的一个内积，并称其中定义了内积的线性空间$[V_n(F);(\alpha,\beta)]$为内积空间.

Example 1.2.2

- n维欧氏空间: $[R^n; (\alpha, \beta) = \alpha^T \beta]$,其中 $\alpha = (\alpha_1, \cdots, \alpha_n)^T$, $\beta = (\beta_1, \cdots, \beta_n)^T$,

$$(\alpha, \beta) = \alpha^T \beta = \sum_{i=1}^{n} \alpha_i \beta_i$$

- L^2空间: $f, g \in L^2(D)$,则

$$(f, g) = \int_D fg$$

Cauchy 不等式

Theorem 1.2.1

$\forall \alpha, \beta \in V_n$,

$$|(\alpha, \beta)|^2 \leq (\alpha, \alpha)(\beta, \beta)$$

Definition 1.2.3

$\forall \alpha \in V_n$, 定义 $\|\alpha\| := (\alpha, \alpha)^{1/2}$， $\|\alpha\|$称为 α的范数

Theorem 1.2.2

Minkovski 不等式

$$\|\alpha + \beta\| \leq \|\alpha\| + \|\beta\|$$

标准正交基

Definition 1.2.4
在内积空间$[V_n(F);(\alpha,\beta)]$中，若一组基$\{\alpha_1,\cdots,\alpha_n\}$满足条件$(\alpha_i,\alpha_j)=\begin{cases}1, i=j\\0, i\neq j\end{cases}$，则称$\{\alpha_1,\cdots,\alpha_n\}$为$V_n(F)$的标准正交基.

Theorem 1.2.3
（*Gram-Schmidt*正交化方法）设$\{\alpha_1,\cdots,\alpha_n\}$是内积空间$[V_n(F);(\alpha,\beta)]$中线性无关的向量组，则由如下方法:

$$\beta_1=\alpha_1,\beta_k=\alpha_k-\sum_{i=1}^{k-1}\frac{(\beta_i,\alpha_k)}{(\alpha_i,\alpha_i)}\beta_i, k=2,\cdots,n$$

所得向量组$\{\beta_1,\cdots,\beta_n\}$是正交向量组.

1.3 线性变换

Definition 1.3.1

映射/变换：两个非空集合A与B间的对应关系f，使得A中的每一个元素a，都在B中总有唯一的一个元素$b = f(a)$与它对应，这种对应称为从A到B的映射(变换)，记作$f : A \to B$.

- Q: 映射与函数的区别是什么?

Definition 1.3.2

T称为由V_n到V_m的变换(映射), 如果T将V_n中的向量映射到V_m中的向量, 记作

$$T : \alpha \in V_n \to \beta = T\alpha \in V_m$$

其中β为α在T下的像, α称为β的原像.

Example 1.3.3

- $T : \alpha = (x_1, x_2)^T \in R^2 \to \beta = (0, x_2)^T \in R^2$
- $T : \alpha = (x_1, x_2)^T \in R^2 \to \beta = (2x_1, 2x_2)^T \in R^2$
- $T : \alpha \in P(t) \to \beta = (\alpha)^2 \in P(t)$

Definition 1.3.4

T为由V_n到V_m的变换, 如果 $\forall k \in F$,$\forall \alpha, \beta \in V_n$, 都有

$$T(\alpha+\beta) = T(\alpha) + T(\beta);\ T(k\alpha) = kT(\alpha);$$

即变换T与线性运算可交换, 则T是线性变换.
当T是V_n到自身的一个线性变换, 则称T是V_n的线性变换.

Example 1.3.5

给定$A \in F^{m\times n}$, 定义由F^n到F^m的变换T 为

$$T : x \in F^n \to y = Ax \in F^m :$$

Example 1.3.6

给定$P \in F^{m\times m}$ 和$Q \in F^{n\times n}$, 定义由$F^{m\times n}$到$F^{m\times n}$的变换T为

$$T : X \in F^{m\times n} \to Y = PXQ \in F^{m\times n}$$

Example 1.3.7

对于$P_n(t)$中的多项式求导运算$\frac{d}{dt}$, 记为D, 即

$$Dp(t) = \frac{dp(t)}{dt}, p(t) \in P_n(t)$$

Example 1.3.8

- 恒等变换I:$I\alpha = \alpha; \forall \alpha \in V$
- 零变换$\mathbf{0}$: $\mathbf{0}\alpha = 0, \forall \alpha \in V$

都是V的线性变换.

Example 1.3.9

定义$T : R^{n\times n} \to R$，使$\forall A \in R^{n\times n}$ $T(A) = det(A)$，证明T不是一个线性变换.

Exe: 判断是否是线性变换:

- $T := (x_1, x_2)^T \in R^2 \to \beta = (0, x_2)^T \in R^2$
- $T := (x_1, x_2)^T \in R^2 \to \beta = (1, x_2)^T \in R^2$
- $T := (x_1, x_2)^T \in R^2 \to \beta = (2x_1, 3x_2)^T \in R^2$
- $T := (x_1, x_2)^T \in R^2 \to \beta = (2x_1, x_2^2)^T \in R^2$

Theorem 1.3.1

V_n 上线性变换T具有如下性质:

- $T(0) = 0$
- $T(-\alpha) = -T(\alpha)$
- $T(\sum k_i\alpha_i) = \sum k_i T(\alpha_i)$
- 若 $\{\alpha_1, \cdots, \alpha_n\}$线性相关, 则$\{T(\alpha_1), \cdots, T(\alpha_n)\}$也线性相关

注: 线性无关的一组向量在T下的像可能是线性相关的, 例如零变换把线性无关的向量都映射为零向量.

Definition 1.3.10

设T_1与T_2都是线性空间$V_n(F)$上的线性变换，定义如下运算：

- 变换的乘积T_1T_2：

$$\forall\alpha\in V_n(F):(T_1T_2)(\alpha)=T_1(T_2(\alpha))$$

- 变换的加法T_1+T_2：

$$\forall\alpha\in V_n(F),(T_1+T_2)(\alpha)=T_1(\alpha)+T_2(\alpha)$$

- 数乘变换kT：

$$\forall\alpha\in V_n(F),(kT)(\alpha)=kT(\alpha)$$

- 可逆变换：对变换T_1，如果存在变换T_2，使

$$T_1T_2=T_2T_2=I(\text{ 恒等变换 })$$

可以证明，上述线性变换运算的结果仍然是$V_n(F)$上的线性变换.

线性变换的矩阵表示

Definition 1.3.11

设T是V_n到V_m的线性变换, 在V_n和V_m 中取基$\mathcal{B}_\alpha = \{\alpha_1, \cdots, \alpha_n\}$和$\mathcal{B}_\beta = \{\beta_1, \cdots, \beta_m\}$, 则$\alpha_j$的像$T(\alpha_j)$可由基$\mathcal{B}_\beta$线性表出:

$$T\alpha_j = \sum_{i=1}^{m} \beta_i a_{ji} = (\beta_1, \cdots, \beta_m) \begin{pmatrix} a_{1i} \\ \vdots \\ a_{mi} \end{pmatrix}$$

将$T\alpha_j$按$j = 1, \cdots, n$的顺序排列, 则有

$$(T\alpha_1, \cdots, T\alpha_m) = (\beta_1, \cdots, \beta_n) \begin{pmatrix} a_{1,1} \cdots a_{1n} \\ \cdots\cdots\cdots \\ a_{m1} \cdots a_{mn} \end{pmatrix}$$

To be cont'd

令$T\mathcal{B}_\alpha=(T\alpha_1,\cdots,T\alpha_m)$,则上式可记为

$$T\mathcal{B}_\alpha=\mathcal{B}_\beta A$$

则这称为线性变换T的矩阵表示, 其中A称为T在基偶$\{\mathcal{B}_\alpha,\mathcal{B}_\beta\}$下的矩阵.
特别地, 若$V_n=V_m$, 并且$\mathcal{B}_\alpha=\mathcal{B}_\beta$,则$T\mathcal{B}_\alpha=\mathcal{B}_\alpha A$,此时称$n$阶方阵$A$为$T$在基$\mathcal{B}_\alpha$下的矩阵.
事实上，T与n阶方阵A是一一对应关系

Example 1.3.12

求$P_{n+1}(t)$ 的线性变换$D=\dfrac{d}{dt}$在基$\mathcal{B}=\{1,t,\cdots,t^n\}$下的矩阵.

Example 1.3.13

求$P_2(t)$ 到$P_3(t)$的线性变换T

$$T(p) = \int_0^t p(s)ds$$

基偶$\{\{1, t, t^2\}; \{1, t, t^2/2.t^3/3\}\}$下的矩阵.

现考虑，由T的矩阵表示，确定T的像的坐标:
设$\alpha \in V_n$, 则可由$\mathcal{B}_\alpha$线性表出,

$$\alpha = \sum_{i=1}^{n} x_i \alpha_i = \mathcal{B}_\alpha x$$

即α在$\mathcal{B}_\alpha$下的坐标为x. 对于线性变换T,

$$T\alpha = \sum_{i=1}^{n} x_i T\alpha_i = T(\mathcal{B}_\alpha)x$$

将$T\mathcal{B}_\alpha$ 用V_m 的基$\mathcal{B}_\beta$表出, 若$T\mathcal{B}_\alpha = \mathcal{B}_\beta A$, 则

$$T\alpha = \mathcal{B}_\beta A x$$

即$T\alpha$在$\mathcal{B}_\beta$下的坐标为Ax.

Theorem 1.3.2

设$\mathcal{B}_\alpha = \{\alpha_1, \cdots, \alpha_n\}$ 和 $\mathcal{B}_\beta = \{\beta_1, \cdots, \beta_m\}$分别是$V_n$ 和V_m 的基, 对于 给定的$m \times n$矩阵$A = [a_{ij}]$, 则存在唯一的从V_n 到V_m 的线性变换T, 使得它在$\{\mathcal{B}_\alpha; \mathcal{B}_\beta\}$下的矩阵是$A$, 即$T\mathcal{B}_\alpha = \mathcal{B}_\beta A$.

Definition 1.3.14

给定V_n和V_m空间的一对基偶$\{\mathcal{B}_\alpha, \mathcal{B}_\beta\}$, 存在 如下一一对应:

线性变换$T: V_n \to V_m \iff m \times n$矩阵$A$;

Question:

- 如果取V_n 和V_m 的另一组基偶$\{\mathcal{B}_{\alpha'}, \mathcal{B}_{\beta'}\}$, 那么T对应另外一个矩阵$B$, 那么矩阵$A$和 B之间有什么关系?
- 怎样选取V_n和V_m的基, 使得T的矩阵表示尽可能简单?

设n阶方阵P是基$\mathcal{B}_\alpha$到$\mathcal{B}_{\alpha'}$的变换矩阵，而m阶方阵Q 是基$\mathcal{B}_\beta$到$\mathcal{B}_{\beta'}$的变换矩阵:

$$\mathcal{B}_{\alpha'} = \mathcal{B}_\alpha P; \mathcal{B}_{\beta'} = \mathcal{B}_\beta Q$$

设$m \times n$矩阵A, B 分别是T在基偶 $\{\mathcal{B}_\alpha, \mathcal{B}_\beta\}$和$\{\mathcal{B}_{\alpha'}, \mathcal{B}_{\beta'}\}$下的矩阵:

$$T\mathcal{B}_\alpha = \mathcal{B}_\beta A, \; T\mathcal{B}_{\alpha'} = \mathcal{B}_{\beta'} B$$

可以推出

$$\mathcal{B}_\beta AP = T\mathcal{B}_\alpha P = T\mathcal{B}_{\alpha'} = \mathcal{B}_{\beta'} B = \mathcal{B}_\beta QB$$

$$\mathcal{B}_\beta (AP - QB) = 0$$

因$\mathcal{B}_\beta$是基，则$AP = QB$

$$\Rightarrow A = QBP^{-1}, B = Q^{-1}AP$$

- 如果$A = QBP^{-1}$ 或$B = Q^{-1}AP$, 其中P, Q 为可 逆方阵, 那么称A和B是相抵(或等价)的. 如上 证明了, 一个V_n到V_m的线性变换在不同基偶下 的矩阵是相抵的.
- 假如$V_m = V_n, \mathcal{B}_\alpha = \mathcal{B}_\beta, \mathcal{B}_{\alpha'} = \mathcal{B}_{\beta'}$, 那么$Q = P$, 则有 $A = PBP^{-1}$. 此时方阵A与B是相似的. 即一个Vn 到自身的线性变换在不同基下的矩阵是相似的.
- 之前的第二个问题等价于: 与A相抵(或相似)的最简单的矩阵是什么?

Example 1.3.15

设R^3上线性变换T为

$$T((x_1, x_2, x_3)^T) = (x_1 + 2x_2 + x_3, x_2 - x_3, x_1 + x_2)^T$$

求T在基

$$\alpha 1 = (1, 0, 1)^T, \alpha_2 = (0, 1, 1)^T, \alpha_3 = (1, -1, 1)^T$$

下的矩阵B

线性变换的核与值域

Definition 1.3.16

设T是从V_n到V_m的线性变换, 则

$$N(T) := \{\forall \alpha \in V_n : T\alpha = 0\}$$

$$R(T) := \{\beta \in V_m : \beta = T\alpha, \alpha \in V_n\}$$

分别称为T的核和T的值域.

:

- $N(T) \subset V_n(F)$, 也被称为T的零空间, 其维数称为T的零度, 记作$nullT$
- $R(T) \subset V_m$, 也被称为T的值空间, 其维数称为T的秩, 记作$rankT$

Theorem 1.3.3

设T是从V_n 到V_m 的线性变换, 则

$$nullT + rankT = n$$

不变子空间

Definition 1.3.17

设T是线性空间$V_n(F)$上的线性变换，W 是$V_n(F)$的子空间，如果$\forall\alpha \in W$，有$T(\alpha) \in W$ ，即值域$T(W) \subset W$，则称W是T的不变子空间.

Example 1.3.18

对R^3上正交投影$P(x) = x - (x, u)u, u$ 为单位向量，R^3的子空间$W = Lu$和$W^\perp = \{x : (x, u) = 0\}$都是$P$的不变子空间.

Example 1.3.19

- T的核$N(T)$和T的值域$R(T)$都是T的不 变子空间
- T的不变子空间的交空间与和空间也是T的 不变子空间.

2.1 线性变换的对角矩阵表示

- 问题引入的背景：

$$T\mathcal{B}_\alpha = \mathcal{B}_\alpha A,$$

如果A是对角阵，则要满足什么条件？即考虑

$$(T\alpha_1, \cdots, T\alpha_n) = (\alpha_1, \cdots, \alpha_n)\begin{pmatrix} \lambda_1, & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}$$

Definition 2.1.1

设T为线性空间V_n上的线性变换，如果存在$\xi \in V_n(F)$和数$\lambda \in F$,$\xi \neq 0$,使得

$$T(\xi) = \lambda\xi$$

则称数λ为T的特征值，向量ξ为线性变换T的对应于特征值λ的特征向量.

Theorem 2.1.1

线性变换T的特征值问题与对应矩阵A的特征值问题是一一对应的. 对于给定的基$\mathcal{B}$,

$$T\xi = \lambda_0 \xi \iff Ax = \lambda_0 x$$

Theorem 2.1.2

线性变换T在不同的基下对应的矩阵是不同的，但是矩阵彼此相似:

- 相似的矩阵具有相同的特征值（T的特征值是由T决定的，和基与变换矩阵的选择无关）；
- 相似矩阵具有不同的特征向量,但其线性无关性相同;
- 相似矩阵有相同的特征多项式, 而T的特征值与A的特征值一样，因此我们可以把A的特征多项式

$$f(\lambda) := det(\lambda I_n - A) = \lambda^n + b_{n-1}\lambda^{n-1} + \cdots + b_1\lambda + b_0$$

称为T的特征多项式, 于是T的特征值就是T的特征多项式的根.

求T的特征对的步骤

- $T\mathcal{B} = \mathcal{B}A$
- $f(\lambda) = det(\lambda I - A)$的根$\lambda_i(i = 1, \cdots, n)$. ($T$的特征根)
- 求$(\lambda_i I - A)Y_i = 0$的非零解，即为λ_i对应的有关A的特征向量Y_i
- $X_i = \mathcal{B}Y_i$,X为λ_i对应的有关T的特征向量

Example 2.1.2

$P_2(t)$ 的线性变换 T 定义为

$$Tp(t) = p(t) + (t+1)\frac{d}{dt}p$$

求 T 的特征值和特征向量.

Theorem 2.1.3

T 关于不同特征值的特征向量线性无关.

Theorem 2.1.4

设$\lambda_1, \cdots, \lambda_k$ 是T的不同特征值. 而 $X_{i1}, X_{i2}, \cdots, X_{ir_i}, (1 \leq i \leq k)$ 是T 关于λ_i的r_i 个线性无关特征向量, 则向量组

$$\{X_{11}, X_{12}, \cdots, X_{1r_1}, X_{21}, X_{22}, \cdots, X_{2r_2}, \cdots, X_{k1}, X_{k2}, \cdots, X_{kr_k}\}$$

线性无关.

线性变换的特征子空间

Definition 2.1.3
对于T的任一特征值λ_0，T关于λ_0的所有特征向量,再添上零向量组成一个集合

$$V_{\lambda_0} := \{\alpha \in V_n : T\alpha = \lambda_0 \alpha\}$$

容易验证，对于任意$\alpha, \beta \in V_{\lambda_0}$和$k, l \in F$，有

$$T(k\alpha + l\beta) = kT(\alpha) + lT(\beta) = k\lambda_0\alpha + l\lambda_0\beta = \lambda_0(k\alpha + l\beta)$$

即V_{λ_0} 是$V_n(F)$的一个子空间，称为T关于λ_0的特征子空间.
$dimV_{\lambda_0}$称为λ_0的几何重数.

Theorem 2.1.5
如果$\alpha \in V_{\lambda_0}$, 则$T(T\alpha) = \lambda_0(T\alpha)$, 特征子空 间$V_{\lambda_0}$是$T$的一个不变子空间.

Definition 2.1.4
如果$\lambda_1, \lambda_2, \cdots, \lambda_s$ 是T的所有不同的特征值, 则 T的特征多项式$f(\lambda)$可以表示为

$$f(\lambda) = (\lambda - \lambda_1)^{n_1}(\lambda - \lambda_2)^{n_2} \cdots (\lambda - \lambda_s)^{n_s}$$

其中$n_1 + n_2 + \cdots + n_s = n$,而$n_i, (1 \le i \le s)$ 称 为特征值λ_i的代数重数.

- Q:几何重数与代数重数的关系?

Theorem 2.1.6

设$\lambda_1, \lambda_2, \lambda_s$ 是T的所有不同特征值. 对 任一$\lambda_i, (1 \le i \le s)$, 都有

$$dim V_{\lambda_i} \le n_i$$

即任何特征值的几何重数不大于其代数重数.

可对角化

如果$W_i, (1 \le i \le s)$都是T的不变子空间，且有

$$V_n = W_1 \oplus W_2 \oplus \cdots \oplus W_s,$$

其中每个W_i的维数有$dim W_i = n_i$，则有 $n_1 + n_2 + \cdots + n_s = n$.
现在每个W_i中取一个基 $\{\alpha_{i1}, \cdots, \alpha_{in_i}\}, (i = 1, 2, \cdots, s)$ 并把它们顺序排列为V_n的基

$$\mathcal{B} = \{\alpha_{11}, \cdots, \alpha_{1n_1}, \alpha_{21}, \cdots, \alpha_{2n_2}, \cdots, \alpha_{s1}, \cdots, \alpha_{sn_s}\}$$

， 若T在$\{\alpha_{i1}, \cdots, \alpha_{in_i}\}$下的矩阵是对角阵， 那么$T$在$\mathcal{B}$下的矩阵是对角块矩阵

$$\begin{pmatrix} A_{11} & & & \\ & A_{22} & & \\ & & \cdots & \\ & & & A_{ss} \end{pmatrix} := diag\{A_{11}, A_{22}, \cdots, A_{ss}\}$$

其中$A_{ii}, (1 \le i \le s)$是n_i阶方阵.

反之, 如果T 在$\mathcal{B}$下的矩阵是上述的对角块矩阵, 则

$$W_i = span\{\alpha_{i1}, \cdots, \alpha_{in_i}\}, (i = 1, 2, \cdots, s)$$

是T的不变子空间, 且V_n是这些不变子空间的直和.

Definition 2.1.5

T称为是可对角化的, 如果存在V_n 的基$\mathcal{B}$, 使T在$\mathcal{B}$ 下的矩阵是对角矩阵.

Theorem 2.1.7

V_n上的线性变换T是可对角化的充分必要 条件是下列等价条件之一成立:

- T有n个线性无关的特征向量
- $dim V_{\lambda_i} = n_i, 1 \leq i \leq s$. (每个特征值$\lambda_i$的几何重数等于代数重数)
- $V_{\lambda_1} \oplus V_{\lambda_2} \oplus \cdots \oplus V_{\lambda_s} = V_n$

Corollary 2.1.6

若T有n个不同的特征值，则T可对角化. 于是, 对于$A \in C^{n\times n}$, 若A的特征多项式没有重根, 则A必可对角化

Corollary 2.1.7

若线性变换T可对角化

- T在基B下的矩阵A相似于一个对角矩阵。 即存在可逆矩阵P, 使得$P^{-1}AP$为对角矩阵,对角矩阵的对角线元素为A 的特征值，且P的每个列向量都是A的特征向量.
- T在不同基下的对应的矩阵都可对角化，因为这些矩阵彼此相似，且相似于同一个对角矩阵(允许交换行列顺序)。

Example 2.1.8

证明$P_2(t)$的线性变换$D = \frac{d}{dt}$是不可对角化的.

Example 2.1.9

R^3的线性变换T定义为

$$T\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 4 & 0 & 1 \\ 2 & 3 & 2 \\ 1 & 0 & 4 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

问T是否可对角化?

Example 2.1.10

证明矩阵

$$A = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

在实数域上是不可对角化的, 但在复数域上是可对角化的.

2.2 Jordan 矩阵

Definition 2.2.1

Jordan 块

$$J(\lambda)=\begin{pmatrix}\lambda & 1 & & \\ & \lambda & 1 & \\ & & \ddots & 1 \\ & & & \lambda\end{pmatrix}$$

的r阶方阵称为一个r阶Jordan块. 由若干个Jordan 块$J_i(\lambda_i)$构成的准对角矩阵

$$J=\begin{pmatrix}J_1(\lambda_1) & & & \\ & J_2(\lambda_2) & & \\ & & \ddots & \\ & & & J_m(\lambda_m)\end{pmatrix}$$

称为Jordan矩阵.

Theorem 2.2.1

- *Jordan*矩阵是准对角矩阵，呈上三角阵形，主对角线上是它的全部特征值
- *Jordan*矩阵就是相似标准形，也就是说若在复有限维空间中考虑．每个线性变换T都有一个*Jordan*矩阵表示，每个复方阵都相似于一个*Jordan*矩阵

Theorem 2.2.2

（*Jordan*标准形存在定理）在实（复）数域上，每个n阶方阵A都相似于一个*Jordan*矩阵，即存在可逆矩P，使得

$$P^{-1}AP = J_A = \begin{pmatrix} J_1(\lambda_1) & & & \\ & J_2(\lambda_2) & & \\ & & \ddots & \\ & & & J_s(\lambda_s) \end{pmatrix}$$

其中

$$J_i(\lambda_i) = \begin{pmatrix} J_{i1}(\lambda_i) & & & \\ & J_{i2}(\lambda_i) & & \\ & & \ddots & \\ & & & J_{it_i}(\lambda_i) \end{pmatrix}$$

，

J_{ij}是n_j阶Jordan块，$\sum_{j=1}^{t_i} n_j = k_i$.$J_i(\lambda_i)$是$k_i \times k_i$阶Jordan矩阵，$\sum_{i=1}^{s} k_i = n$.若不计较Jordan块的排列次序，则每个方阵的Jordan标准形J_A是惟一的.

- R^n或C^n线性空间上的线性变换T，都存在一组基，使T在该基下矩阵为Jordan矩阵，从而Jordan矩阵就是线性变换矩阵最简形式

Jordan标准形的求法

1. 求A的特征多项式

$$f_A(\lambda) = |\lambda I - A| = (\lambda - \lambda_1)^{k_1}(\lambda - \lambda_2)^{k_2}\cdots(\lambda - \lambda_s)^{k_s}$$

   其中$\lambda_i, 1 \leq i \leq s$互不相同。

2. 对λ_i, 由$(A - \lambda_i I)X = 0$, 求A的线性无关的特征向量$\alpha_i, 1 \leq i \leq t_i$. λ_i的几何重 数$dimV_{\lambda_i} = t_i$决定$J_i(\lambda_i)$中有t_i个Jordan 块.

3. 若λ_i的代数重数等于几何重数: $k_i = dimV_{\lambda_i}$, λ_i对应的Jordan矩阵为k_i阶对角矩阵.
   若$dimV_{\lambda_i} < k_i$, 则在V_{λ_i}中选择适当特征向量α_i, 求Jordan链$\{\alpha_i, \beta_1, \cdots, \beta_{n_i}\}$, 其中

$$\begin{cases} (A - \lambda_i I)\alpha_i = 0 \\ (A - \lambda_i I)\beta_1 = \alpha_i \\ (A - \lambda_i I)\beta_k = \beta_{k-1}, 2 \leq k \leq n_i \end{cases}$$

   $\beta_1, \cdots, \beta_{n_i}$称为广义特征向量, 从而得到了$J_A$的结构.

4. 所有Jordan链构成矩阵P，必有

$$P^{-1}AP = J_A$$

Example 2.2.2

$$A = \begin{pmatrix} -3 & 3 & -2 \\ -7 & 6 & -3 \\ 1 & -1 & 2 \end{pmatrix}$$

求可逆矩阵P和Jordan矩阵J_A，使$P^{-1}AP = J_A$

Example 2.2.3

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

求可逆矩阵P和Jordan矩阵J_A，使$P^{-1}AP = J_A$

Example 2.2.4

设$P_3[x]$上线性变换T在基$\{1, x, x^2\}$下的矩阵为A，求$P_3[x]$的基$\{\xi_1, \xi_2, \xi_3\}$，使T在此基下的矩阵为Jordan矩阵．其中

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

2.3 最小多项式

Recall：设$A \in F6n \times n$的特征多项式为

$$
\begin{aligned}
f_A(\lambda) := |\lambda I - A| &= \lambda^n + b_{n-1}\lambda^{n-1} + \cdots + b_1\lambda + b_0 \\
&= (\lambda - \lambda_1)\cdots(\lambda - \lambda_n)
\end{aligned}
$$

这个n次多项式在复数域有n个根$\lambda_i, i = 1, \cdots, n$ (其中重根进行重复计数, 例如$(\lambda - 1)^2(\lambda - 2)$有 3个根,$\lambda_1 = \lambda_2 = 1, \lambda_3 = 2$)

- 方阵的迹trA(对角元之和) 等于特征值之和.

$$
tr(A) = a_{11} + a_{22} + \cdots + a_{nn} = \sum_{i=1}^{n} \lambda_i = -b_{n-1}
$$

- 方阵的行列式的值等于所有特征值的乘积.

$$
(-1)^n b_0 = \lambda_1\lambda_2\cdots\lambda_n = det(A)
$$

(自己证明)

Definition 2.3.1

若
$$g(\lambda) = a_m\lambda^m + a_{m-1}\lambda^{m-1} + \cdots + a_1\lambda + a_0$$
是一个多项式，方阵$A \in F^{n\times n}$, 则称
$$g(A) = a_mA^m + a_{m-1}A^{m-1} + \cdots + a_1A + a_0I$$
为A的一个矩阵多项式.显然，$g(A)$也是n阶方阵。

Example 2.3.2

设 $g(\lambda) = \lambda^2 - 3\lambda + 4, h(\lambda) = \lambda^2 - 3\lambda + 2$
$$A = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$$
求 $g(A)$和$h(A)$

Theorem 2.3.1

设$A \in F^{n\times n}$，$g(A)$是A的矩阵多项式，则有如下结果：

1. 若(λ_0, X)是A的特征值与特征向量，则$(g(\lambda_0), X)$是$g(A)$的特征值与特征向量.
2. 若$A \sim B$,则 $g(A) \sim g(B)$
3. 如果A为准对角矩阵，则$g(A)$也是准对角矩阵．而且

$$\begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_k \end{pmatrix}, A_k\text{是方块阵}$$

则

$$g(A) = \begin{pmatrix} g(A_1) & & & \\ & g(A_2) & & \\ & & \ddots & \\ & & & g(A_k) \end{pmatrix}$$

4. $p(\lambda)$和$q(\lambda)$是两个多项式，则$p(A)q(A) = q(A)p(A)$

Example 2.3.3

设
$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$
求$g(A) = 2A^5 - 3A^4 - A^3 + 2A - I$的特征值与特征向量

Theorem 2.3.2

设$A \in F^{m\times n}$,$B \in F^{n\times m}(m \geq n)$，$f_{AB}$和$f_{BA}$分别是$AB$与$BA$的特征多项式，则有
$$f_{AB}(\lambda) = \lambda^{m-n} f_{BA}(\lambda)$$
即m阶方阵AB与n阶方阵BA的非零特征值相同。特别地，若$m = n$，则$\lambda_{AB} = \lambda_{BA}$

Example 2.3.4

求镜像变换的Householder 矩阵

$$H = I - 2\omega\omega^T$$

的特征值及它的迹和行列式

Example 2.3.5

若方阵A的所有特征值的模都小于1(即$|\lambda_A| < 1$), 则方 阵$I - A$是可逆的.

Definition 2.3.6

设A是一个n阶方阵, $g(\lambda)$是一个多项式, 如果$g(A) = 0$, 则称$g(\lambda)$是A 的零化多项式.

- Q: A的零化多项式是否唯一?

Theorem 2.3.3

(Cayley-Hamilton): 设n阶方阵A的特征 多项式为

$$f(\lambda) = det(\lambda I - A) = \lambda^n + b_{n-1}\lambda^{n-1} + \cdots + b_1\lambda + b_0$$

则$f(A) = O$, 即A的特征多项式为其零化多项式.

- Q: 如何求$g(A)$?
- (1).利用Jordan型化简, （2）.利用CH定理

Example 2.3.7

设

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

求(1).$g(A) = 2A^5 - 3A^4 - A^3 + 2A - I$;(2).$A^{-1}$

最小多项式

Definition 2.3.8

A的零化多项式中, 次数最低的首一多项式 称为A的最小(零化)多项式, 记为$m_A(\lambda)$.

Theorem 2.3.4

A的最小多项式$m_A(\lambda)$可整除A的任何零化多项式$g(\lambda)$, 且 $m_A(\lambda)$ 是唯一的. (最小多项式)

Corollary 2.3.9

A的最小多项式$m_A(\lambda)$必能整除A的特征 多项式$f(\lambda)$,即 $m_A(\lambda)|f(\lambda)$

Theorem 2.3.5

λ_0是A的特征值, 其充分必要条件是λ_0 是A的最小多项式$m_A(\lambda)$的根.

Corollary 2.3.10

设矩阵A的特征多项式

$$f(\lambda) = (\lambda - \lambda_1)^{n_1} \cdots (\lambda - \lambda_s)^{n_s},$$

则$n_1 + \cdots + n_s = n$,则必有

$$m_A(\lambda) = (\lambda - \lambda_1)^{k_1} \cdots (\lambda - \lambda_s)^{k_s},$$

且$k_i \leq n_i, i = 1, \cdots, s$.若$A$的代数重数都为1,则

$$m_A(\lambda) = f(\lambda)$$

Example 2.3.11

$$A = \begin{pmatrix} 2 & 1 & & & \\ & & 2 & & \\ & & & 1 & 1 \\ & & & -2 & 4 \end{pmatrix}$$

求$m_A(\lambda)$.

Example 2.3.12

$$A = \begin{pmatrix} A_1 & \\ & A_2 \end{pmatrix}, \text{ 求 } m_A(\lambda),$$

其中

$$A_1 = \begin{pmatrix} 2 & 1 & & & \\ & -2 & & & \\ & & -2 & 1 & \\ & & & -2 & 1 \\ & & & & -2 \end{pmatrix}, A_2 = \begin{pmatrix} 5 & & \\ & 5 & 1 \\ & & 5 \end{pmatrix}$$

- 若 A为对角块矩阵，即 $A = diag\{A_1, \cdots, A_m\}$, 则 $m_A(\lambda)$ 是 $m_A(\lambda_1), \cdots, m_{A_m}(\lambda_m)$的最小公倍式

Theorem 2.3.6

线性变换多项式$f(T) = 0$等价于矩阵多项式$f(A) = 0$,其中A是线性变换T下的矩阵. 事实上，$m_T(\lambda) = m_A(\lambda)$

Theorem 2.3.7

设变换T的特征多项式为

$$f(\lambda) = (\lambda - \lambda_1)^{n_1} \cdots (\lambda - \lambda_s)^{n_s}$$

又T的$Jordan$标准形中关于特征值λ_i 的$Jordan$块的最高阶数为k_i，则T的最小多项式

$$m_T(\lambda) = (\lambda - \lambda_1)^{k_1} \cdots (\lambda - \lambda_s)^{k_s}$$

- 矩阵A的特征子空间

$$V_{\lambda_0} := \{x \in R^n : Ax = \lambda_0 x\}$$

显然 $V_{\lambda_0} = N(A - \lambda_0 I)$, 因此

$$dim V_{\lambda_0} = n - rank(A - \lambda_0 I)$$

Example 2.3.13

$$A = \begin{pmatrix} 0 & 0 & 0 & 4 \\ 1 & 0 & 0 & -4 \\ 0 & 1 & 0 & -3 \\ 0 & 0 & 1 & 4 \end{pmatrix}$$

求$m_A(\lambda)$

Lemma 2.3.14

$A \in F^{m\times n}, B \in F^{n\times s}$,则

$$rank(A) + rank(B) - n \le rank(AB) \le min\{rank(A), rank(B)\}$$

Theorem 2.3.8

n阶方阵A可对角化当且仅当A的最小多项式$m_A(\lambda)$没有重根.

Example 2.3.15

设$A_i, (i = 1, \cdots, k)$ 都是n阶方阵，求证：若

$$A_1 A_2 \cdots A_k = O$$

则

$$rank(A_1) + \cdots + rank(A_k) \le (k-1)n$$

Example 2.3.16

A为n阶方阵，且$g(\lambda) = \lambda^3 - 3\lambda^2 - \lambda + 3$，设$g(A) = 0$,求证：$A$必可对角化

矩阵的分解

- 常见的形式:
  1. $A = A_1 + A_2 + \cdots + A_k$, 矩阵的和
  2. $A = A_1 \cdot A_2 \cdots A_k$, 矩阵的积
- 分解的作用:
  1. 理论上, 揭示矩阵的特性
  2. 应用上, 简化矩阵的计算
- 主要技巧
  1. 各种标准形的理论和计算方法
  2. 矩阵的分块运算和初等变换

- 常见的标准形:
  1. 等价标准型: $A_{mn} = P_{mn} \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} Q_{nn}$
  2. 相似标准形: $A = PJP^{-1}$ 特别的, 当A为对称矩阵时, J 为对角阵, 且存在正交矩阵C 使得$A = CJC^T$.
- 主要内容:(三角分解),满秩分解,Schur 分解,UR 分解, QR 分解,奇异值分解

矩阵的三角分解

Definition 3.0.1

$A \in F^{n\times n}$

- 若$L, U \in F^{n\times n}$分别是下三角矩阵和上三角矩阵，$A = LU$ ，则称A可作LU分解.
- 若$L, V \in F^{n\times n}$分别是对角线元素为1 的下三角矩阵和上三角矩阵，D为对角矩 阵. $A = LDV$，则称A可作LDV分解.

Example 3.0.2

设

$$A = \begin{pmatrix} 2 & 2 & 3 \\ 4 & 7 & 7 \\ -2 & 4 & 5 \end{pmatrix}$$

求A的LU分解和LDV分解

设A 的LU 分解为 $A = LU$ ，则

$$AX = b \iff LUX = b \iff \begin{cases} LY = b \\ UX = Y \end{cases}$$

上式都是系数矩阵为三角形矩阵，是易于用回代法求解的线性方程组．先 用自上往下的回代法得Y．再代入求解X（用自下往上的回代法），即 可得到原方程组$AX = b$ 的解X

Example 3.0.3

设

$$A = \begin{pmatrix} 1 & -3 & 7 \\ 2 & 4 & -3 \\ -3 & 7 & 2 \end{pmatrix}, b = \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix}$$

用LU分解求解线性方程组$AX = b$.

矩阵的满秩分解

Definition 3.0.4

对秩为r 的矩阵$A \in F^{m\times\Theta n}$, 若存在两个秩 为$r$ 的矩阵 $B \in F^{m\times r}$, $C \in F^{r\times n}$, 使得$A = BC$, 则称为A的满秩分解

Theorem 3.0.1

任何非零矩阵$A \in F^{m\times n}$ 都有满秩分解

- 满秩分解的求法: 利用初等变换
  1. : 等价标准形求法(行列变换), 求两个逆
  2. : 阶梯型求法(行变换), 只求一个逆矩阵
  3. : 求列的极大无关组(行变换), 不用求逆

Example 3.0.5

求A 的满秩分解

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

Example 3.0.6

求A 的满秩分解

$$A = \begin{pmatrix} 0 & 1 & 0 & -1 & 5 & 6 \\ 0 & 2 & 0 & 0 & 0 & -14 \\ 2 & -1 & 2 & -4 & 0 & 1 \\ -2 & 1 & -2 & 2 & 10 & 25 \end{pmatrix}$$

可逆矩阵的UR 分解

Theorem 3.0.2

$A \in C^{n\times n}$为可逆方阵, 存在酉矩阵Q和主 对角线上元素皆正的上三角矩阵R, 使$A = UR$. (这称为A 的UR 分解)

Example 3.0.7

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

求A的UR 分解

列满秩矩阵的QR 分解

Theorem 3.0.3

$A \in C^{n\times r}$为列满秩矩阵, 则有$A = QR$, 其 中$Q \in C^{n\times r}$ 的列向量为标准正交的向量组和 $R \in C^{r\times r}$ 为主对角线上元素皆正的上三角矩阵, *(这称为A的QR分解)*

Shur 分解

- 已知：实对称矩阵A 正交相似于对角阵.
- 问题：对于复方阵，何时可以酉相似于对角阵?

Theorem 3.0.4

$A \in C^{n\times r}$, 存在酉矩阵U和上三角矩阵T, 使得

$$U^H A U = T = \begin{pmatrix} \lambda_1 & * & * & * \\ & \lambda_2 & \cdots & * \\ & & \ddots & \vdots \\ & & & \lambda_n \end{pmatrix}$$

其中λ_i为A的特征值。

正规矩阵

Definition 3.0.8

若矩阵$A \in C^{n\times n}$, 满足$A^HA = A^AH$,则称A是一个正规矩阵

Example 3.0.9

下列矩阵都是正规矩阵

- 对角矩阵;
- 对称与反对称矩阵: $A^T = A, A^T = -A$;
- Hermite 与反Hermite 矩阵: $A^H = A, A^H = -A$;
- 正交矩阵与酉矩阵, $A^TA = AA^T = I_n, A^HA = AA^H = I_n$.

Example 3.0.10

设A为正规矩阵, B酉相似于A , 证明B也是正规矩阵

Theorem 3.0.5

$A \in C^{n\times n}$是正规矩阵的充分必要条件是A酉相似于对角矩阵，即存在酉矩阵$U \in C^{n\times n}$, 使得

$$U^H AU = \begin{pmatrix} \lambda_1 & & & \\ & \lambda 2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}$$

Corollary 3.0.11

$A \in C^{n\times n}$是正规矩阵的充分必要条件是A有n个线性无关的特征向量构成空间C^n 的标准正交基.

Example 3.0.12

证明Hermite矩阵的特征值是实数，而且属于不同特征值的特征向量是正 交的.

矩阵的奇异值分解（SVD）

Theorem 3.0.6

$A \in C^{m\times n}, A^HA \in C^{n\times n}$ 和 $AA^H \in C^{m\times m}$ 为*Hermite* 矩阵

1. $rankA = rankAA^H = rankA^HA$
2. A^HA和 AA^H的非零特征值相等
3. A^HA 和 AA^H 半正定.
   A秩为n, A^HA 正定; A 秩为m, AA^H 正定.
   从而A^HA 和AA^H 的特征值为非负实数:

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$$

Definition 3.0.13

若$A \in C^{m\times n}$, 秩为r, 设A^HA 的特征值$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > 0$,
$\lambda_{r+1} = \cdots = \lambda_n = 0$, 则矩阵$A$的奇异值为
$\sigma_i = \sqrt{\lambda_i}, i = 1, 2, \cdots, r$

Theorem 3.0.7

矩阵A 的奇异值具有如下性质：

- $A \in C^{n\times n}$为正规矩阵时，A的奇异值为A的特征值的模
- $A \in C^{n\times n}$为正定的$Hermite$矩阵时，A的奇异值等于A的特征值．
- 若存在酉矩阵$U \in C^{m\times m}$, $V \in C^{n\times n}$，矩阵$B \in C^{m\times n}$，使$UAV = B$，则称A和B 酉等价，酉等价的矩阵A和B有相同的奇异值．

奇异值分解是酉等价型的分解:

Theorem 3.0.8

$A \in C^{m\times n}$,存在酉矩阵$U \in C^{m\times m}$, $V \in C^{n\times n}$, 使得 $A = U\Sigma V^H$,

其中 $\Sigma = \begin{pmatrix} \Delta_r & O \\ O & O \end{pmatrix}_{m\times n}$, $\Delta_r = \begin{pmatrix} \sigma_1 & & & \\ & \sigma_1 & & \\ & & \ddots & \\ & & & \sigma_r \end{pmatrix}$

- 奇异值分解常适用于解决矩阵秩相关问题
- A的奇异值分解依赖于A^HA 的酉相似分解

Example 3.0.14

求矩阵

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

和

$$B = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

的奇异值分解

矩阵U, V 的空间性质

Theorem 3.0.9

$V = [v_1, v_2, \cdots, v_r, \cdots, v_n] = [V1, V2] \in C^{n\times n}$ 的列向量是C^n 的标准正交基*(右奇异向量)*

- $V_2 \in C^{n\times(n-r)}$的列向量是$N(A)$标准正交基$(AV_2 = 0)$

$U = [u_1, u_2, \cdots, u_r, \cdots, u_n] = [U1, U2] \in C^{m\times m}$ 的列向量是C^m 的标准正交基*(左奇异向量)*

- $U_1 \in C^{m\times r}$ 的列向量是$R(A)$标准正交基 $(A = U_1\Delta_r V_1^H, U_1 = AV_1\Delta_r^{-1})$

Theorem 3.0.10

$A = \sigma_1 u_1 v_1^H + \sigma_2 u_2 v_2^H + \cdots + \sigma_r u_r v_r^H$ *(按照奇异值分解展开可得)*

满秩矩阵与单侧逆

Definition 4.1.1

(满秩矩阵与单侧逆)设$A \in C^{m \times n}$，若存在矩阵$B \in C^{n \times m}$，使得

$$BA = I_n$$

则称A是左可逆的，称B为A的一个左逆矩阵，记为A_L^{-1} 若存在矩阵$C \in C^{n \times m}$，使得

$$AC = I_m$$

则称A是右可逆的，称C为A的一个右逆矩阵，记为A_R^{-1}

Theorem 4.1.1

设$A \in C^{m \times n}$，则下列条件是等价的：

- A是左可逆的
- A的零空间$N(A) = \{0\}$
- $m \geq n$，$rank(A) = n$，即A是列满秩的
- $A^H A$ 是可逆的

Theorem 4.1.2

设$A \in C^{m \times n}$，则下列条件是等价的:

- A是右可逆的
- A的列空间$R(A) = C^m$
- $m \leq n$，$rank(A) = n$，即A是行满秩的
- AA^H 是可逆的

单侧逆与解线性方程组

Theorem 4.1.3

设$A \in C^{m \times n}$是左可逆的,$B \in C^{n \times m}$是A 的一个左逆矩阵，则线性方程组 $AX = b$有形如$X = Bb$ 解的充分必要条件是

$$(I_m - AB)b = 0$$

若上式成立，则方程组有惟一解

$$X = (A^H A)^{-1} A^H b$$

Theorem 4.1.4

设$A \in C^{m \times n}$是右可逆的,则线性方程组$AX = b$ 对任何$b \in C^m$ 都有解，且 对A的任意一个右逆矩阵A_R^{-1}, $X = A_R^{-1} b$是其解．特别地，$X = A^H(AA^H)^{-1} b$是方程 组$AX = b$ 的一个解．

4.2 广义逆矩阵

1. 减号广义逆

Definition 4.2.1

设$A \in C^{m\times n}$，若存在矩阵$G \in C^{n\times m}$，使得 $AGA = A$ ， 则称G为A 的一个减号广义逆或$\{1\}$–逆

A的全部减号广义逆的集合记为$A\{1\}$, $A\{1\}$的元素用$A_1^-, A_2^-, \cdots$表示.

Theorem 4.2.1

设$A \in C^{m\times n}$,$rank(A) = r$, 若存在可逆阵$P \in C^{m\times m}$ 和 $Q \in C^{n\times n}$, 使得

$$PAQ = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$$

则$G \in A\{1\}$的充分必要条件是

$$G = Q \begin{pmatrix} I_r & U \\ V & W \end{pmatrix} P$$

其中$U \in C^{r\times(m-r)}$, $V \in C^{(n-r)\times r}$, $W \in C^{(n-r)\times(m-r)}$是任意的.

Example 4.2.2

设

$$A = \begin{pmatrix} 0 & -1 & 3 & 0 \\ 2 & -4 & 1 & 5 \\ -4 & 5 & 7 & -10 \end{pmatrix}$$

求A的减号广义逆.

Theorem 4.2.2

设$A \in C^{m \times n}$ ，则A的减号广义逆惟一的必要充分条件是$m = n$ ，且A^{-1} 存在.

Theorem 4.2.3

设$A \in C^{m \times n}$,$\lambda \in C$ ，则A^{-}满足以下性质:

1. $rank(A) \leq rank(A^{-})$
2. AA^{-}与$A^{-}A$都是幂等矩阵，且$rank(A) = rank(AA^{-}) = rank(A^{-}A)$
3. $R(AA^{-}) = R(A), N(A^{-}A) = N(A)$

Theorem 4.2.4

设$A \in C^{m \times n}, A^{-} \in A\{1\}$, 则当方程组$Ax = b$ 有解时，其通解可表示为

$$x = A^{-}b + (I_n - A^{-}A)z$$

，这里z 是任意的n维列向量.

2. Moore-Penrose 广义逆（加号广义逆）

Definition 4.2.3

设$A \in C^{m\times n}$，若存在矩阵$G \in C^{n\times m}$,使得

- $AGA = A$
- $GAG = G$
- $(AG)^H = AG$
- $(GA)^H = GA$

则称G为A的Moore-Penrose 广义逆或加号广义逆，简称为A的M-P 逆. A的任意M- P 逆记为A^+

Theorem 4.2.5

设$A \in C^{m\times n}$存在M-P广义逆，则A 的M-P逆是惟一的

Theorem 4.2.6

任意矩阵$A \in C^{m\times n}$都存在M-P 广义逆A^+．设$rank(A) = r$， A 的一个满秩分解为

$$A = BC, B \in C^{m\times r}, C \in C^{r\times n}, rank(B) = rank(C) = r$$

则

$$A^+ = C^H(CC^H)^{-1}(B^HB)^{-1}B^H$$

Example 4.2.4

求矩阵

$$A = \begin{pmatrix} 1 & 0 & -1 & 1 \\ 0 & 2 & 2 & 2 \\ -1 & 4 & 5 & 3 \end{pmatrix}$$

的M-P逆A^+

Theorem 4.2.7

设矩阵$A \in C^{m\times n}$, $rank(A) = r$ ，A的奇异值分解为

$$A = U \begin{pmatrix} \Delta & 0 \\ 0 & 0 \end{pmatrix} V^H$$

则

$$A^+ = V \begin{pmatrix} \Delta^{-1} & 0 \\ 0 & 0 \end{pmatrix} U^H$$

其中Δ为对角线由A的正奇异值所构成的对角矩阵，$\Delta \in C^{r\times r}$.

Example 4.2.5

求矩阵

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

的M-P逆A^+

.

注意:
(1). $(AB)^+ \neq B^+A^+$. (2). $(A^+)^k \neq (A^k)^+$

Theorem 4.2.8

设$A \in C^{m\times n}$,$\lambda \in C^n$, 则A^+满足以下性质:

1. $(A^+)^+ = A$
2. $(A^+)^H = (A^H)^+$
3. $(\lambda A)^+ = \lambda^+ A$,其中 $\lambda^+ = \begin{cases} \frac{1}{\lambda}, \lambda \neq 0 \\ 0, \lambda = 0 \end{cases}$
4. $A \in C^{m\times n}$是列满秩的，则
$$A^+ = (A^H A)^{-1} A^H$$
若A是行满秩的，则
$$A^+ = A^H (AA^H)^{-1}$$
5. 若A有满秩分解$A = BC$ ，则$A^+ = C^+B^+$

Example 4.2.6

设$A \in C^{m\times n}$,证明

1. $rank(A) = rank(A^+)$
2. $rank(A^+A) = rank(AA^+) = rank(A)$

最小二乘法

Definition 4.2.7

设$A \in C^{m \times n}, b \in C^m$, 如果存在$u \in C^n$ ，使得

$$\|Au - b\| \leq \|Ax - b\|, \forall x \in C^n$$

则称u是方程组$Ax = b$ 的一个最小二乘解.
设x_0是$Ax = b$的最小二乘解，如果对于$Ax = b$的每一个最小二乘解u ，都有

$$\|x_0\| \leq \|u\|$$

则称x_0是最佳最小二乘解（或称按范数最小的最小二乘解，或称最佳逼近解）.

Theorem 4.2.9

设设$A \in C^{m \times n}, b \in C^m$, ，则$x_0 = A^+ b$是线性方程组$Ax = b$ 的最佳的最小二乘解.

Example 4.2.8

求不相容的线性方程组

$$\begin{pmatrix} 1 & 0 & -1 & 1 \\ 0 & 2 & 2 & 2 \\ -1 & 4 & 5 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 4 \\ 1 \\ 2 \end{pmatrix}$$

最佳的最小二乘解.

- 数据的最佳拟合函数问题：设由观察得到关于物理量x和y的一组数据为$(x_1, y_1), (x_2, y_2), \cdots, (x_m, y_m)$. 我们希望通过这些数据，找出函数$y = f(x)$，使得它能最好的反映x和y之间的依赖关系.

Example 4.2.9

设有一组实验数据： $(1,2),(2,3),(3,5),(4,7)$从数据点的走向看接近一条直线，实验者希望使用直线$y=\beta_0+\beta_1 x$拟合数据点，求最佳拟合直线

Example 4.2.10

设一个质点运动的轨道是椭圆，观察到的位置坐标是$(1,1),(0,2),(-1,1),(-1,-2)$,点在同一平面上，求拟合观察点的最佳标准椭圆方程