Why the matrix analysis

1. 从数学分析引申得到的线性代数的论题
   - ✓多元微积分、复变量、微分方程、最优化和逼近理论等
2. 解决实的复的的线性代数的方法
   - ✓极限、连续和、幂级数等
3. 离散信号分析的最有效的数学工具

本课程的主要内容

1. 矩阵理论：如线性空间、线性变换、内积空间、正交投影、Jordan标准型、范数理论等；
2. 矩阵分析方法：如矩阵函数的微积分、广义逆矩阵、矩阵分解、特征值和奇异值估计、矩阵直积运算等；
3. 特殊矩阵：介绍信号处理中常用的特殊矩阵如Toeplitz矩阵、Hankel矩阵、Hilbert矩阵等
4. 矩阵分析方法在信号处理中的应用

矩阵分析与应用

- 参考书：
  - 《矩阵论》第二版 程云鹏主编 西北工业大学出版社 2004年8月
  - 《矩阵分析与应用》 张贤达 清华大学出版社 2004年9月
  - “Matrix Analysis”, Roger A. Horn 机械工业出版社影印版
  - 《矩阵计算》，G.H.戈卢布等，科学出版社
- 编程工具
  - Matlab、C

线性空间

- 线性空间
- 线性变换与矩阵
- 线性子空间

集合与元素

集合:是指一些对象的总体

元素:这些对象称为集合的元素

- 整数集
- 线性方程组的解集
- 由某个平面上所有的点构成的点集

用S表示集合,a是S的元素

$$a \in S$$

a不是S的元素

$$a \notin S$$

集合的表示

1.列举全部元素

如 $N=\{1,3,5,7,9\}$

2.给出集合中的元素的性质

M是具有某些性质的全部元素所组成的集合

$M=\{a|a$所具有的性质$\}$

单位圆上的所有点 $N=\{(a,b)\,|\,a^2+b^2=1\}$

所有正整数 $N_0=\{n\,|\,n \text{ is integral}\}$

集合的运算

- 子集 $\forall a \in A, \exists a \in B \Rightarrow A \subseteq B$ or $B \supseteq A$
- 真子集 $A \subset B$ $A \supset B$
- 相等 $A = B$
- 交 $A \cap B = \{x \mid x \in A, 且 x \in B\}$
- 并 $A \cup B = \{x \mid x \in A, 或 x \in B\}$
- 和集 $A + B = \{x + y \mid x \in A, y \in B\}$

数环和数域

数环：设Z是一个非空数集，且其中任意两个数的和、差和积仍属于Z，则称Z是一个数环

- ✓任何数环都含有0元素
- ✓若 $\forall a \in Z, \exists -a \in Z$

数域：关于四则运算封闭的数的集合

- ✓任何数域都含有元素0和元素1
- ✓若 $\forall a \in P, \exists 1/a \in P$
- ✓典型数域:复数域C;实数域R;有理数域Q
- ✓任意数域K都包括有理数域Q

设非空集合V，一个数域K，如果V满足：

Ⅰ 在V中定义一个封闭的加法 $x, y, z \in V$, $k, l \in K$

加法交换率 $x + y = y + x$

加法结 x $z)$ $) +$

零向

Ⅱ 在V中

数对元素

元素对数分配率

数因子结合率 $= (k$

单位向量 1 x

V中满足8条性质，
且为封闭的加法
和数乘运算，
统称线性运算

则称V是数域K上的线性空间，当K是实数域时，称V为实线性空间；当K是复数域时，称V为复线性空间

例1 实系数，次数不超过n的一元多项式的集合

$$P[x]_n=\{a_nx^n+\cdots+a_1x+a_0\mid a_n,\cdots,a_1,a_0\in\mathbf{R}\}$$

例2 常系数二阶齐次线性微分方程的解集

$$y''-3y'+2y=0$$

$$Y=\{ae^{2x}+be^{x}\mid a,b\in\boldsymbol{R}\}$$

例3 在所有n阶实矩阵的集合 $\boldsymbol{R}^{n\times n}$

线性空间的基本性质

1. 零元素是唯一的

假设有零元素 $\mathbf{0}_1$ 和 $\mathbf{0}_2$，有 $\mathbf{0}_1 = \mathbf{0}_1 + \mathbf{0}_2 = \mathbf{0}_2$

2. 任一元素的负元素是唯一的

假设$\boldsymbol{x}$有负元素x_1 和 x_2 ，有

$$x_1 = x_1 + \mathbf{0} = x_1 + (x + x_2) = (x_1 + x) + x_2$$

$$= (x + x_1) + x_2 = \mathbf{0} + x_2 = x_2$$

线性空间的基本性质

1. 零元素是唯一的
2. 任一元素的负元素是唯一的
3. 设 $k,0,1 \in K$，$\boldsymbol{x},\boldsymbol{0},-\boldsymbol{x} \in V$，有

① $0\boldsymbol{x} = \boldsymbol{0}$

② $(-1)\boldsymbol{x} = -\boldsymbol{x}$

③ $k\boldsymbol{0} = \boldsymbol{0}$

④ 若 $k\boldsymbol{x} = \boldsymbol{0}$，则 $k = 0$ 或 $\boldsymbol{x} = \boldsymbol{0}$ 。

定义：如果 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_r (r \geq 1)$ 是线性空间V中的一组向量，$k_1, k_2, \cdots, k_r$ 是数域$\boldsymbol{K}$中的数，那么向量 $\boldsymbol{x} = k_1\boldsymbol{x}_1 + k_2\boldsymbol{x}_2 + \cdots + k_r\boldsymbol{x}_r$

称 $\boldsymbol{x}$ 为是向量 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_r$ 的线性组合，或者称向量 $\boldsymbol{x}$ 是 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_r$ 的线性表示

如果 $k_1, k_2, \cdots, k_r$ 不全为零，且使

$$k_1\boldsymbol{x}_1 + k_2\boldsymbol{x}_2 + \cdots + k_r\boldsymbol{x}_r = \boldsymbol{0}$$

则称向量组 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_r$ 线性相关，否则称为线性无关

无关向量组的个数称为向量组的维数

例4：在 $\mathbf{R}^n$ 中，有两个向量组

$$\begin{cases} \boldsymbol{\varepsilon}_1 = (1,0,\cdots,0), \\ \boldsymbol{\varepsilon}_2 = (0,1,\cdots,0), \\ \qquad\vdots \\ \boldsymbol{\varepsilon}_n = (0,0,\cdots,1). \end{cases} \qquad \begin{cases} \boldsymbol{\varepsilon}_1' = (1,1,\cdots,1,1), \\ \boldsymbol{\varepsilon}_2' = (0,1,\cdots,1,1), \\ \qquad\vdots \\ \boldsymbol{\varepsilon}_n' = (0,0,\cdots,0,1). \end{cases}$$

分别考察

$$k_1\boldsymbol{\varepsilon}_1 + k_2\boldsymbol{\varepsilon}_2 + \cdots + k_r\boldsymbol{\varepsilon}_r = \mathbf{0} \qquad k_1\boldsymbol{\varepsilon}_1' + k_2\boldsymbol{\varepsilon}_2' + \cdots + k_r\boldsymbol{\varepsilon}_r' = \mathbf{0}$$

$$\begin{vmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 \end{vmatrix} \neq 0 \qquad \begin{vmatrix} 1 & 1 & \cdots & 1 \\ 0 & 1 & \cdots & 1 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 \end{vmatrix} \neq 0$$

两个方程的系数行列式均不为零，因此方程只有零解

所以两个向量组均为线性无关组

例5：讨论下列 $\mathbf{R}^{2\times 2}$ 的矩阵组的线性相关性

$$A_1=\begin{bmatrix} a & 1 \\ 1 & 1 \end{bmatrix}, A_2=\begin{bmatrix} 1 & a \\ 1 & 1 \end{bmatrix}, A_3=\begin{bmatrix} 1 & 1 \\ a & 1 \end{bmatrix}, A_4=\begin{bmatrix} 1 & 1 \\ 1 & a \end{bmatrix}$$

设有一组数 k_1,k_2,k_3,k_4 使

$$k_1A_1+k_2A_2+k_3A_3+k_4A_4=\mathbf{0}$$

$$\begin{cases} ak_1+k_2+k_3+k_4=0 \\ k_1+ak_2+k_3+k_4=0 \\ k_1+k_2+ak_3+k_4=0 \\ k_1+k_2+k_3+ak_4=0 \end{cases} \qquad \begin{vmatrix} a & 1 & 1 & 1 \\ 1 & a & 1 & 1 \\ 1 & 1 & a & 1 \\ 1 & 1 & 1 & a \end{vmatrix}=(a-1)^3(a+3$$

当 $a\neq 1, a\neq -3$ 线性无关，否则线性相关

设V是**R**上全体实函数构成的线性空间,讨论V中元素组 t, e^t, e^{2t} 的线性相关性

设有一组数 k_1, k_2, k_3 使元素组的线性组合以及一阶和二阶导数为零

$$\begin{aligned} tk_1 + e^t k_2 + e^{2t} k_3 &= 0 \\ k_1 + e^t k_2 + 2e^{2t} k_3 &= 0 \\ k_1 + e^t k_2 + 4e^{2t} k_3 &= 0 \end{aligned}$$

$$\begin{vmatrix} t & e^t & e^{2t} \\ 1 & e^t & 2e^{2t} \\ 0 & e^t & 4e^{2t} \end{vmatrix} = e^{3t}(2t-3) \neq 0$$

因此 t, e^t, e^{2t} 线性无关

1. 单个向量x线性相关的充要条件是$x=0$。两个以上的向量线性相关的充要条件是其中有一个向量是其余向量的线性组合。

2. 如果向量组 $x_1, x_2, \cdots, x_r$ 线性无关，而且可以被向量组 $y_1, y_2, \cdots, y_s$ 线性表出，那么 $r \le s$

✓两个等价的线性无关的向量组，必含有相同数量的向量

3. 如果向量组 $x_1, x_2, \cdots, x_r$ 线性无关，但向量组$x_1, x_2, \cdots, x_r, y$ 线性相关，那么y可以由 $x_1, x_2, \cdots, x_r$ 线性表出，而且表示法唯一

定义：线性空间V中线性无关向量组所含向量最大个数n称为V的维数，记做 $\dim V = n$

- 维数为n的线性空间称为数域K上的n维线性空间，记为 V^n。

- 当 $n \to \infty$ ，称为无限维线性空间

定义：数域K上的线性空间V， $\boldsymbol{x}_1,\cdots,\boldsymbol{x}_r\,(r\geq 1)$ 是属于V的任意r个向量，如果满足

1. $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_r$ 线性无关
2. V中的任一向量$\boldsymbol{x}$都是 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_r$ 的线性组合

则称 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_r$ 是V的一个基或基底，并称 $\boldsymbol{x}_i\,(i=1,2,\cdots,r)$ 为基向量。

✓线性空间的维数就是基中所含基向量个数

称线性空间V^n的一个基 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n$ 为V^n的一个坐标系，设向量 $\boldsymbol{x}\in V^n$，在该基下的线性表示为 $\boldsymbol{x}=\xi_1\boldsymbol{x}_1+\xi_2\boldsymbol{x}_2+\cdots+\xi_n\boldsymbol{x}_n$

则称$\xi_1,\xi_2,\cdots,\xi_n$ 是$\boldsymbol{x}$在该坐标系下的坐标或分量，记为 $(\xi_1,\xi_2,\cdots,\xi_n)^T$

定理:设 $x_1, x_2, \cdots, x_n$是V^n的一个基,$x \in V^n$,则x可以唯一的表示为 $x_1, x_2, \cdots, x_n$ 的线性组合

- 线性空间的维数与所考虑的数域有关

✓如果把复数域K看作是自身上的线性空间，那么它是一维的，数1 就是基；

✓把复数域K看作是实数域上的线性空间，那么就是二维的，数1 与i 就是一组基.

基变换与坐标变换

- 设 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n$ 是 V^n 的旧基，$\boldsymbol{y}_1,\boldsymbol{y}_2,\cdots,\boldsymbol{y}_n$ 是新基。新基可以用旧基表示出来

$$\begin{cases} \boldsymbol{y}_1 = c_{11}\boldsymbol{x}_1 + c_{21}\boldsymbol{x}_2 + \cdots + c_{n1}\boldsymbol{x}_n \\ \boldsymbol{y}_2 = c_{12}\boldsymbol{x}_1 + c_{22}\boldsymbol{x}_2 + \cdots + c_{n2}\boldsymbol{x}_n \\ \quad\vdots \\ \boldsymbol{y}_n = c_{1n}\boldsymbol{x}_1 + c_{2n}\boldsymbol{x}_2 + \cdots + c_{nn}\boldsymbol{x}_n \end{cases}$$

$$(\boldsymbol{y}_1,\boldsymbol{y}_2,\cdots,\boldsymbol{y}_n) = (\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n)\begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{bmatrix} \triangleq (\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n)\boldsymbol{C}$$

称C为由旧基改变为新基的过渡矩阵

过渡矩阵的性质

1）过渡矩阵都是可逆矩阵；反过来，任一可逆矩阵都可看成是两组基之间的过渡矩阵.

证明：若 $x_1,x_2,\cdots,x_n$； $y_1,y_2,\cdots,y_n$ 为V的两组基，

且由基 $x_1,x_2,\cdots,x_n$到$y_1,y_2,\cdots,y_n$ 的过渡矩阵为C，

即 $(y_1,y_2,\cdots,y_n)=(x_1,x_2,\cdots,x_n)C$ ①

又由基 $y_1,y_2,\cdots,y_n$到$x_1,x_2,\cdots,x_n$ 也有一个过渡矩阵，

设为B，即 $(x_1,x_2,\cdots,x_n)=(y_1,y_2,\cdots,y_n)B$ ②

比较① 、②两个等式，有

$$(y_1,y_2,\cdots,y_n)=(y_1,y_2,\cdots,y_n)BC$$
$$(x_1,x_2,\cdots,x_n)=(x_1,x_2,\cdots,x_n)CB$$

$\because$ $x_1,x_2,\cdots,x_n$; $y_1,y_2,\cdots,y_n$ 都是线性无关的,

$\therefore$ $CB=BC=I.$ 即，C是可逆矩阵, 且$C^{-1}=B$.

反过来，设 $C=(c_{ij})_{n\times n}$ 为K上任一可逆矩阵，

任取V的一组基 $x_1,x_2,\cdots,x_n$,

令$y_j=\sum_{i=1}^{n}c_{ij}x_i$, $j=1,2,\cdots,n$

于是有， $(y_1,y_2,\cdots,y_n)=(x_1,x_2,\cdots,x_n)C$

由C可逆，有 $(x_1,x_2,\cdots,x_n)=(y_1,y_2,\cdots,y_n)C^{-1}$

即，$x_1,x_2,\cdots,x_n$也可由 $y_1,y_2,\cdots,y_n$ 线性表出.

$\therefore$ $x_1,x_2,\cdots,x_n$与$y_1,y_2,\cdots,y_n$等价.

故 $y_1,y_2,\cdots,y_n$ 线性无关，从而也为V的一组基.

并且C就是 $x_1,x_2,\cdots,x_n$到$y_1,y_2,\cdots,y_n$ 的过渡矩阵.

2）若由基$x_1,x_2,\cdots,x_n$到基$y_1,y_2,\cdots,y_n$ 过渡矩阵为C，

则由基 $y_1,y_2,\cdots,y_n$到基$x_1,x_2,\cdots,x_n$ 过渡矩阵为C^{-1}.

3）若由基 $x_1, x_2, \cdots, x_n$ 到基 $y_1, y_2, \cdots, y_n$ 过渡矩阵为 C，由基 $y_1, y_2, \cdots, y_n$ 到基 $z_1, z_2, \cdots, z_n$ 过渡矩阵为 B，则由基 $x_1, x_2, \cdots, x_n$ 到基 $z_1, z_2, \cdots, z_n$ 过渡矩阵为 CB.

事实上，若 $(y_1, y_2, \cdots, y_n) = (x_1, x_2, \cdots, x_n)C$

$$(z_1, z_2, \cdots, z_n) = (y_1, y_2, \cdots, y_n)B$$

则有，$(z_1, z_2, \cdots, z_n) = ((x_1, x_2, \cdots, x_n)C)B$

$$= (x_1, x_2, \cdots, x_n)CB$$

定义：V为数域P上n维线性空间 $x_1, x_2, \cdots, x_n$;

$y_1, y_2, \cdots, y_n$ 为V中的两组基，且

$$(y_1, y_2, \cdots, y_n) = (x_1, x_2, \cdots, x_n)\begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{pmatrix}$$

设 $x \in V$ 且x在基 $x_1, x_2, \cdots, x_n$ 与基 $y_1, y_2, \cdots, y_n$

下的坐标分别为 $(\xi_1, \xi_2, \cdots, \xi_n)$ 与 $(\eta_1, \eta_2, \cdots, \eta_n)$，

即,

$$x=(x_1,x_2,\cdots,x_n)\begin{pmatrix}\xi_1\\ \xi_2\\ \vdots\\ \xi_n\end{pmatrix} \text{与} \ x=(y_1,y_2,\cdots,y_n)\begin{pmatrix}\eta_1\\ \eta_2\\ \vdots\\ \eta_n\end{pmatrix}$$

则
$$\begin{pmatrix}\xi_1\\ \xi_2\\ \vdots\\ \xi_n\end{pmatrix}=\begin{pmatrix}c_{11} & c_{12} & \cdots & c_{1n}\\ c_{21} & c_{22} & \cdots & c_{2n}\\ \cdots & \cdots & \cdots & \cdots\\ c_{n1} & c_{n2} & \cdots & c_{nn}\end{pmatrix}\begin{pmatrix}\eta_1\\ \eta_2\\ \vdots\\ \eta_n\end{pmatrix} \qquad (*)$$

或
$$\begin{pmatrix}\eta_1\\ \eta_2\\ \vdots\\ \eta_n\end{pmatrix}=\begin{pmatrix}c_{11} & c_{12} & \cdots & c_{1n}\\ c_{21} & c_{22} & \cdots & c_{2n}\\ \cdots & \cdots & \cdots & \cdots\\ c_{n1} & c_{n2} & \cdots & c_{nn}\end{pmatrix}^{-1}\begin{pmatrix}\xi_1\\ \xi_2\\ \vdots\\ \xi_n\end{pmatrix} \qquad (**)$$

称(*)或(**)为向量x在基变换C下的坐标变换公式.

例： 在V^n中，求由基$x_1,x_2,\cdots,x_n$到基$y_1,y_2,\cdots,y_n$的过渡矩阵及由基$y_1,y_2,\cdots,y_n$到基$x_1,x_2,\cdots,x_n$的过渡矩阵．其中

$$x_1=(1,0,\cdots,0),x_2=(0,1,\cdots,0),\cdots,x_n=(0,\cdots,0,1)$$

$$y_1=(1,1,\cdots,1),y_2=(0,1,\cdots,1),\cdots,y_n=(0,\cdots,0,1)$$

并求向量$\alpha=(a_1,a_2,\cdots,a_n)$在基$y_1,y_2,\cdots,y_n$下的坐标.

解：$\because$ $$\begin{cases} y_1 = x_1 + x_2 + \cdots + x_n \\ y_2 = \qquad x_2 + \cdots + x_n \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ y_n = \qquad\qquad\qquad x_n \end{cases}$$

$$\because (y_1, y_2, \cdots, y_n) = (x_1, x_2, \cdots, x_n)\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 1 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 1 & 1 & \cdots & 1 \end{pmatrix}$$

而，$$(x_1, x_2, \cdots, x_n) = (y_1, y_2, \cdots, y_n)\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 1 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 1 & 1 & \cdots & 1 \end{pmatrix}^{-1}$$

$$= (y_1, y_2, \cdots, y_n)\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ -1 & 1 & 0 & \cdots & 0 \\ 0 & -1 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}$$

故，由基 $x_1,x_2,\cdots,x_n$ 到基 $y_1,y_2,\cdots,y_n$ 的过渡矩阵为

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 1 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 1 & 1 & \cdots & 1 \end{pmatrix}$$

由基 $y_1,y_2,\cdots,y_n$ 到基 $x_1,x_2,\cdots,x_n$ 的过渡矩阵为

$$\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ -1 & 1 & 0 & \cdots & 0 \\ 0 & -1 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}$$

$\alpha=(a_1,a_2,\cdots,a_n)$在基 $x_1,x_2,\cdots,x_n$下的坐标就是

$$(a_1,a_2,\cdots,a_n)$$

设α在基 $y_1,y_2,\cdots,y_n$下的坐标为 $(b_1,b_2,\cdots,b_n)$，则

$$\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ -1 & 1 & 0 & \cdots & 0 \\ 0 & -1 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2-a_1 \\ \vdots \\ a_n-a_{n-1} \end{pmatrix}$$

所以α在基$y_1,y_2,\cdots,y_n$下的坐标为

$$(a_1,a_2-a_1,\cdots,a_n-a_{n-1})$$

例： 在V^4中，求由基 y_1, y_2, y_3, y_4 到基 x_1, x_2, x_3, x_4 的过渡矩阵，其中

$y_1 = (1, 2, -1, 0)$　　$x_1 = (2, 1, 0, 1)$

$y_2 = (1, -1, 1, 1)$　　$x_2 = (0, 1, 2, 2)$

$y_3 = (-1, 2, 1, 1)$　　$x_3 = (-2, 1, 1, 2)$

$y_4 = (-1, -1, 0, 1)$　　$x_4 = (1, 3, 1, 2)$

解：设 $e_1=(1,0,0,0),\ \ e_2=(0,1,0,0),$

$$e_3=(0,0,1,0),\ \ e_4=(0,0,0,1)$$

则有

$$(y_1,y_2,y_3,y_4)=(e_1,e_2,e_3,e_4)\begin{pmatrix}1&1&-1&-1\\2&-1&2&-1\\-1&1&1&0\\0&1&1&1\end{pmatrix}$$

或

$$(e_1,e_2,e_3,e_4)=(y_1,y_2,y_3,y_4)\begin{pmatrix}1&1&-1&-1\\2&-1&2&-1\\-1&1&1&0\\0&1&1&1\end{pmatrix}^{-1},$$

$$(x_1,x_2,x_3,x_4)=(e_1,e_2,e_3,e_4)\begin{pmatrix}2&0&-2&1\\1&1&1&3\\0&2&1&1\\1&2&2&2\end{pmatrix}$$

从而有

$$(x_1,x_2,x_3,x_4)=\left((y_1,y_2,y_3,y_4)\begin{pmatrix}1&1&-1&-1\\2&-1&2&-1\\-1&1&1&0\\0&1&1&1\end{pmatrix}^{-1}\right)\begin{pmatrix}2&0&-2&1\\1&1&1&3\\0&2&1&1\\1&2&2&2\end{pmatrix}$$

$$=(y_1,y_2,y_3,y_4)\left(\begin{pmatrix}1&1&-1&-1\\2&-1&2&-1\\-1&1&1&0\\0&1&1&1\end{pmatrix}^{-1}\begin{pmatrix}2&0&-2&1\\1&1&1&3\\0&2&1&1\\1&2&2&2\end{pmatrix}\right)$$

$$=(y_1,y_2,y_3,y_4)\begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

$\therefore$ 由基 y_1,y_2,y_3,y_4 到基 x_1,x_2,x_3,x_4 的过渡矩阵为

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

例：已知 $R^{2\times 2}$的两组基：

$$E_{11}=\begin{pmatrix}1&0\\0&0\end{pmatrix},E_{12}=\begin{pmatrix}0&1\\0&0\end{pmatrix},E_{21}=\begin{pmatrix}0&0\\1&0\end{pmatrix},E_{22}=\begin{pmatrix}0&0\\0&1\end{pmatrix};$$

$$F_{11}=\begin{pmatrix}1&0\\0&0\end{pmatrix},F_{12}=\begin{pmatrix}1&1\\0&0\end{pmatrix},F_{21}=\begin{pmatrix}1&1\\1&0\end{pmatrix},F_{22}=\begin{pmatrix}1&1\\1&1\end{pmatrix}$$

求由基 $E_{11},E_{12},E_{21},E_{22}$到$F_{11},F_{12},F_{21},F_{22}$ 的过渡矩阵，

并求矩阵 $A=\begin{pmatrix}-3&5\\4&2\end{pmatrix}$在基 $F_{11},F_{12},F_{21},F_{22}$下的矩阵.

解：

$$\because \begin{cases} F_{11} = E_{11} \\ F_{12} = E_{11} + E_{12} \\ F_{21} = E_{11} + E_{12} + E_{21} \\ F_{22} = E_{11} + E_{12} + E_{21} + E_{22} \end{cases}$$

$$\therefore \quad (F_{11}, F_{12,} F_{21}, F_{22}) = (E_{11}, E_{12,} E_{21}, E_{22}) \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

又 $A = -3E_{11} + 5E_{12,} + 4E_{21} + 2E_{22}$

设A在基 $F_{11}, F_{12,} F_{21}, F_{22}$下的坐标为 (a_1, a_2, a_3, a_4),

则 $$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} -3 \\ 5 \\ 4 \\ 2 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -3 \\ 5 \\ 4 \\ 2 \end{pmatrix} = \begin{pmatrix} -8 \\ 1 \\ 2 \\ 2 \end{pmatrix}$$

即A在基 $F_{11}, F_{12}, F_{21}, F_{22}$下的坐标为$(-8, 1, 2, 2)$.

作业

- P26: 10、11、12

矩阵分析与应用

第二讲

线性子空间

2006-12-1

本讲主要内容

- 线性子空间的定义
- 线性子空间的性质
- 线性子空间的交
- 线性子空间的和
- 子空间交与和的有关性质

线性子空间

设V_1是数域K上的线性空间V上一个非空子集合，且对已有的线性运算满足以下条件：

1. 如果 $x, y \in V_1$ ，则 $x + y \in V_1$ ；
2. 如果 $x \in V_1, k \in K$ ， $kx \in V_1$

则称V_1是V的线性子空间或子空间

- 线性子空间也是线性空间
- 非零线性空间的平凡子空间：线性空间自身以及零空间
- 线性子空间的维数小于等于线性空间的维数

线性子空间V_1也是线性空间

证明：必要性由定义直接得出

充分性：各运算律已在V中定义，我们只需证明

$$\exists \boldsymbol{0} \in \boldsymbol{V}_1 \qquad \forall x \in V_1, \exists -x \in V_1$$

实际上，$\boldsymbol{0} = 0 \bullet \boldsymbol{x} \in \boldsymbol{V}_1$

$$\forall \boldsymbol{x} \in V_1, -1 \in \boldsymbol{K} \qquad -\boldsymbol{x} = (-1)\,\boldsymbol{x} \in \boldsymbol{V}_1$$

所以线性子空间V_1也是线性空间

V_1是数域K上的线性空间V上一个非空子空间

$\Longleftrightarrow$ $\forall \boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{V}_1, \forall k, l \in \boldsymbol{K}$ 有 $\exists k\boldsymbol{x} + l\boldsymbol{y} \in \boldsymbol{V}_1$

- 充分性:设 $k=l=1$，$\forall \boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{V}_1 \Rightarrow \boldsymbol{x} + \boldsymbol{y} \in \boldsymbol{V}_1$

  取$l=0$，$\forall \boldsymbol{x} \in \boldsymbol{V}_1, \forall k \in \boldsymbol{K} \Rightarrow k\boldsymbol{x} \in \boldsymbol{V}_1$

- 必要性: $\forall \boldsymbol{x} \in \boldsymbol{V}_1, \forall k \in \boldsymbol{K} \Rightarrow k\boldsymbol{x} \in \boldsymbol{V}_1$ （数乘封闭）

  $\forall \boldsymbol{y} \in \boldsymbol{V}_1, \forall l \in \boldsymbol{K} \Rightarrow l\boldsymbol{y} \in \boldsymbol{V}_1$ （数乘封闭）

  故 $k\boldsymbol{x} + l\boldsymbol{y} \in \boldsymbol{V}_1$ （加法封闭）

#

n元齐次线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{s1}x_1 + a_{s2}x_2 + \cdots + a_{sn}x_n = 0 \end{cases}$$

的全部解向量所成集合V对于通常的向量加法和数量乘法构成的线性空间是n维向量空间 R^n 的一个子空间，称V为方程组的解空间

- 方程组的解空间W的维数＝n－秩(A)，
- 方程组的一个基础解系就是解空间V的一组基

n元齐次线性方程组

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=0 \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=0 \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_{s1}x_1+a_{s2}x_2+\cdots+a_{sn}x_n=0 \end{cases}$$

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1s}x_s=c_{1,s+1}x_{s+1}+\cdots+c_{1n}x_n \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2s}x_s=c_{2,s+1}x_{s+1}+\cdots+c_{2n}x_n \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_{s1}x_1+a_{s2}x_2+\cdots+a_{ss}x_s=c_{s,s+1}x_{s+1}+\cdots+c_{sn}x_n \end{cases}$$

变形后，用n-s组数表示自由未知量 $x_{s+1},\cdots,x_n$

$(1,0,\cdots,0)$ $(0,1,\cdots,0)$ $\cdots$ $(0,0,\cdots,1)$

得到n-s个解向量

$(\gamma_{11},\gamma_{12},\cdots,\gamma_{1s}1,0,\cdots,0)$ $\cdots$ $(\gamma_{s1},\gamma_{s2},\cdots,\gamma_{ss},0,0,\cdots,1)$

这个解向量组就是方程组的解空间的基

判断$\mathbf{R}^n$的下列子集合哪些是子空间：

$$V_1=\{(x_1,x_2,\cdots,x_n)\,|\,x_1+x_2+\cdots+x_n=0,x_i\in R\}$$

$$V_2=\{(x_1,x_2,\cdots,x_n)\,|\,x_1+x_2+\cdots+x_n=1,x_i\in R\}$$

$$V_3=\{(x_1,x_2,\cdots,x_{n-1},0)\,|\ \ x_i\in R,i=1,2,\cdots,n-1\}$$

若为$\mathbf{R}^n$的子空间，求出其维数与一组基．

解：V_1、V_3是$\mathbf{R}^n$的子空间，V_2不是$\mathbf{R}^n$的子空间．

事实上，V_1是n元齐次线性方程组

$$x_1+x_2+\cdots+x_n=0 \qquad ①$$

的解空间．所以，维V_1＝n－1，①的一个基础解系

$\eta_1=(1,-1,0,\cdots,0)$, $\eta_2=(1,0,-1,0,\cdots,0)$, $\cdots\cdots$,

$\eta_{n-1}=(1,0,\cdots,0,-1)$ 就是V_1的一组基.

而在V_2中任取两个向量 x, y ，设

$$x=(x_1,x_2,\cdots,x_n),\quad y=(y_1,y_2,\cdots,y_n)$$

则 $x+y=(x_1+y_1,x_2+y_2,\cdots,x_n+y_n)$

但是 $(x_1+y_1)+(x_2+y_2)+\cdots+(x_n+y_n)$

$$=(x_1+x_2+\cdots+x_n)+(y_1+y_2+\cdots+y_n)=1+1=2$$

$\therefore\quad x+y\notin V_2$， 故$V_2$不是$R^n$的子空间.

下证V_3是R^n的子空间.

首先 $\mathbf{0}=(0,0,\cdots,0)\in V_3$, $\therefore V_3\neq\varnothing$

其次, $\forall x,y\in V_3,\forall k\in K,$

设 $x=(x_1,x_2,\cdots,x_{n-1},0)$, $y=(y_1,y_2,\cdots,y_{n-1},0)$

则有 $x+y=(x_1+y_1,x_2+y_2,\cdots,x_{n-1}+y_{n-1},0)\in V_3$

$$kx=(kx_1,kx_2,\cdots,kx_{n-1},0)\in V_3$$

故，V_3为R^n的一个子空间，且维V_3＝n－1 ，

$$\varepsilon_i=(0,\cdots,0,\underset{i}{1},0\cdots,0),\ \ i=1,2,\cdots,n-1$$

就是V_3的一组基.

设 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n$ 是数域K上的线性空间V的一组向量，所有可能的线性组合的集合

$$\boldsymbol{V}_1 = \{k_1\boldsymbol{x}_1 + \cdots + k_n\boldsymbol{x}_n\} \quad k_i \in K, i = 1, 2, \cdots, n$$

称为由$\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n$ 生成(或张成)的子空间,记为

$$L(\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n) = \{k_1\boldsymbol{x}_1 + \cdots + k_n\boldsymbol{x}_n\}$$

如果 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_m \quad m < n$ 是线性无关组，

则 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_m$ 是生成的子空间的基

零子空间就是零元素生成的子空间

例：设V为数域K上的线性空间， $x_1, x_2, \cdots, x_m \in V$

$$V_1 = \{k_1x_1 + k_2x_2 + \cdots + k_mx_m \mid k_i \in P, i = 1,2,\cdots,m\}$$

则V_1关于V的运算作成V的一个子空间.

即 $x_1, x_2, \cdots, x_m$ 的一切线性组合所成集合.

证： $\forall x \in$ … $\Rightarrow x = k_1x_1 + k_2x_2 + \cdots + k_mx_m$

… ${}_2x_2 + \cdots + l_mx_m$

$\forall k, l \in$ … $x_1 + \cdots + (kk_m + ll_m)x_m$

$\therefore kx + ly \in V_1$

因此，V_1关于V的运算作成V的一个子空间

扩基定理

定理：设V_1为 n 维线性空间 V的一个 m 维子空间，

$x_1, x_2, \cdots, x_m$ 为V_1的一组基，则这组向量必定可扩充

为V的一组基．即在V中必定可找到$n-m$个向量

$x_{m+1}, x_{m+2}, \cdots, x_n$ ，使 $x_1, x_2, \cdots, x_n$为V的一组基．

证明：对$n-m$作数学归纳法．

当 $n-m=0$时，即 $n=m$，

$x_1, x_2, \cdots, x_m$ 就是V的一组基．　定理成立．

假设当$n-m=k$时结论成立．

下面我们考虑 $n-m=k+1$ 的情形.

既然 $x_1,x_2,\cdots,x_m$ 还不是V的一组基，它又是线性无关的，那么在V^n中必定有一个向量 x_{m+1}不能被 $x_1,x_2,\cdots,x_m$ 线性表出，把它添加进去，则

$x_1,x_2,\cdots,x_m,x_{m+1}$ 必定是线性无关的.

由定理， 子空间 $L(x_1,x_2,\cdots,x_{m+1})$ 是$m+1$维的.

因 $n-(m+1)=(n-m)-1=(k+1)-1=k$,

由归纳假设，$L(x_1,x_2,\cdots,x_{m+1})$的基 $x_1,x_2,\cdots,x_m,x_{m+1}$ 可以扩充为整个空间V^n的一组基．由归纳原理得证.

■矩阵的值域

设 $\mathbf{A}\in\mathbf{C}^{m\times n}$ 的n个列向量为 $\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_n$ 则

$$R(\mathbf{A})=L\left(\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_n\right)=\left\{y\mid y=\mathbf{A}\boldsymbol{x},\boldsymbol{x}\in\mathbf{C}^n\right\}$$

是 $\mathbf{C}^n$ 的子空间，称为矩阵A的值域，或列空间

■矩阵的零空间（核空间）

设 $\mathbf{A}\in\mathbf{C}^{m\times n}$ 的n个列向量为 $\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_n$ 则

$$N(\mathbf{A})=\left\{\boldsymbol{x}\mid\mathbf{A}\boldsymbol{x}=0,\boldsymbol{x}\in\mathbf{C}^n\right\}$$

是 $\mathbf{C}^n$ 的子空间，称为矩阵A的零空间，其维数为A的零度，记为 $n(\mathbf{A})$

- 已知 $\mathbf{A}=\begin{bmatrix}1 & 0 & 1\\ 0 & 1 & 1\end{bmatrix}$求A的秩和零度

显然A的秩为2，即 $\mathrm{rank}\mathbf{A}=2$

又由 $\mathbf{A}\boldsymbol{x}=0$ ，可以解得

可以得到 $x=\begin{pmatrix}1 & 1 & -1\end{pmatrix}^T t$ $n(\mathbf{A})=1$

同样可以得到 $\mathrm{rank}\mathbf{A}^T=2,$ $n\left(\mathbf{A}^T\right)=0$

- 若 $\mathbf{A}\in\boldsymbol{R}^{m\times n}$ ，有

$$\mathrm{rank}\mathbf{A}+n(\mathbf{A})=n \quad n(\mathbf{A})-n\left(\mathbf{A}^T\right)=n-m$$

定理： 设V_1、V_2为线性空间V的子空间，则集合

$$V_1 \cap V_2 = \{a \mid a \in V_1 \text{且} a \in V_2\}$$

也为V的子空间，称之为V_1与V_2的交空间.

事实上，$\because 0 \in V_1, 0 \in V_2$，$\therefore 0 \in V_1 \cap V_2 \neq \varnothing$

任取 $x, y \in V_1 \cap V_2$，即 $x, y \in V_1$, 且$x, y \in V_2$,

则有 $x + y \in V_1, x + y \in V_2$，$\therefore$ $x + y \in V_1 \cap V_2$

同时有 $kx \in V_1, kx \in V_2$，$\therefore$ $kx \in V_1 \cap V_2$，$\forall k \in K$

故 $V_1 \cap V_2$为V的子空间.

定理： 设V_1、V_2为线性空间V的子空间，则集合

$$V_1+V_2=\{x_1+x_2\,|\,x_1\in V_1,x_2\in V_2\}$$

也为V的子空间，称之为V_1与V_2的和空间.

事实上，$\because 0\in V_1, 0\in V_2,\quad \therefore 0=0+0\in V_1+V_2\neq\varnothing$

任取 $x,y\in V_1+V_2$， 设 $x=x_1+x_2, y=y_1+y_2$，

其中，$x_1,y_1\in V_1, x_2,y_2\in V_2$，则有

$$x+y=(x_1+x_2)+(y_1+y_2)\ \in V_1+V_2$$

$$kx=k(x_1+x_2)=kx_1+kx_2\in V_1+V_2,\quad \forall k\in K$$

- 子空间的交满足交换率与结合率

$$V_1 \cap V_2 = V_2 \cap V_1,$$

$$(V_1 \cap V_2) \cap V_3 = V_1 \cap (V_2 \cap V_3)$$

- 子空间的和满足交换率与结合率

$$V_1 + V_2 = V_2 + V_1,$$

$$(V_1 + V_2) + V_3 = V_1 + (V_2 + V_3)$$

注意：

V的两子空间的并集未必为V的子空间．例如

$$V_1=\{(a,0,0)\mid a\in R\},\ V_2=\{(0,b,0)\mid b\in R\}$$

皆为R^3的子空间，但是它们的并集

$$V_1\cup V_2=\{(a,0,0),(0,b,0)\mid a,b\in R\}$$

$$=\{(a,b,0)\mid a,b\in R\text{ 且}a,b\text{中至少有一是}0\}$$

并不是R^3的子空间．因为它对R^3的运算不封闭，如

$$(1,0,0),\ (0,1,0)\in V_1\cup V_2$$

但是 $(1,0,0)+(0,1,0)=(1,1,0)\notin V_1\cup V_2$

子空间的交与和的有关性

包含在V1和V2中的最大的子空间

1、设 V_1,V_2,W 为线性空间V的子空

1）若 $W\subseteq V_1, W\subseteq V_2$，则 $W\subseteq V_1\cap V_2$.

2）若 $V_1\subseteq W, V_2\subseteq W$，则 $V_1+V_2\subseteq W$.

2、设 V_1,V_2 为线性空间V的子空间，则以下三条件等价:

包含V1和V2中的最小的子空间

1) $V_1\subseteq V_2$

2) $V_1\cap V_2=V_1$

3) $V_1+V_2=V_2$

3、 $x_1, x_2, \cdots, x_s; y_1, y_2, \cdots, y_t$ 为线性空间V中两组向量，则 $L(x_1, x_2, \cdots, x_s) + L(y_1, y_2, \cdots, y_t)$

$$= L(x_1, x_2, \cdots, x_s, y_1, y_2, \cdots, y_t)$$

4、维数公式（定理1.6）

设 V_1, V_2 为线性空间V的两个子空间，则

$$\dim V_1 + \dim V_2 = \dim(V_1 + V_2) + \dim(V_1 \cap V_2)$$

或 $$\dim(V_1 + V_2) = \dim V_1 + \dim V_2 - \dim(V_1 \cap V_2)$$

证：设 $\dim V_1 = n_1$，$\dim V_2 = n_2$，$\dim(V_1 \cap V_2) = m$

取 $V_1 \cap V_2$ 的一组基 $x_1, x_2, \cdots, x_m$

由扩基定理，它可扩充为 V_1 的一组基

$$x_1, x_2, \cdots, x_m, y_1, y_2, \cdots, y_{n_1 - m}$$

它也可扩充为 V_2 的一组基

$$x_1, x_2, \cdots, x_m, z_1, z_2, \cdots, z_{n_2 - m}$$

即有 $V_1 = L(x_1, x_2, \cdots, x_m, y_1, y_2, \cdots, y_{n_1 - m})$

$$V_2 = L(x_1, x_2, \cdots, x_m, z_1, z_2, \cdots, z_{n_2 - m})$$

所以，有

$$V_1+V_2=L(x_1,x_2,\cdots,x_m,y_1,y_2,\cdots,y_{n_1-m},z_1,z_2,\cdots,z_{n_2-m})$$

下证 $x_1,x_2,\cdots,x_m,y_1,y_2,\cdots,y_{n_1-m},z_1,z_2,\cdots,z_{n_2-m}$

线性无关． 假设有等式

$$k_1x_1+\cdots+k_mx_m+p_1y_1+\cdots+p_{n_1-m}y_{n_1-m}$$

$$+q_1z_1+\cdots+q_{n_2-m}z_{n_2-m}=0$$

令 $x=k_1x_1+\cdots+k_mx_m+p_1y_1+\cdots+p_{n_1-m}y_{n_1-m}$

$$=-q_1z_1-\cdots-q_{n_2-m}z_{n_2-m}$$

则有 $x \in V_1$且$x \in V_2$，于是$x \in V_1 \cap V_2$，

即 x可被 $x_1, x_2, \cdots, x_m$ 线性表出

令 $x = l_1 x_1 + l_2 x_2 + \cdots + l_m x_m$，

则 $l_1 x_1 + l_2 x_2 + \cdots + l_m x_m = -q_1 z_1 - \cdots - q_{n_2 - m} z_{n_2 - m}$

即 $l_1 x_1 + l_2 x_2 + \cdots + l_m x_m + q_1 z_1 + \cdots + q_{n_2 - m} z_{n_2 - m} = 0$

由于 $x_1, x_2, \cdots, x_m, z_1, z_2, \cdots, z_{n_2 - m}$ 线性无关，得

$$l_1 = l_2 = \cdots\cdots = l_m = q_1 = \cdots\cdots = q_{n_2 - m} = 0,$$

因而 $x = 0$ 从而有

$$k_1x_1+\cdots+k_mx_m+p_1y_1+\cdots+p_{n_1-m}y_{n_1-m}=0$$

由于 $x_1,x_2,\cdots,x_m,y_1,y_2,\cdots,y_{n_1-m}$ 线性无关，得

$$k_1=k_2=\cdots=k_m=p_1=\cdots=p_{n_1-m}=0$$

所以， $x_1,x_2,\cdots,x_m,y_1,y_2,\cdots,y_{n_1-m},z_1,z_2,\cdots,z_{n_2-m}$

线性无关． 因而它是 V_1+V_2 的一组基．

$$\therefore \dim(V_1+V_2)=m+(n_1-m)+(n_2-m)=n_1+n_2-m$$

$$=\dim V_1+\dim V_2-\dim(V_1\cap V_2)$$

#

注意：从维数公式中可以看到，子空间的和的维数往往比子空间的维数的和要小.

例如，在$\mathbf{R}^3$中，设子空间

$$V_1 = L(\varepsilon_1,\varepsilon_2),\ V_2 = L(\varepsilon_2,\varepsilon_3)$$

其中，$\varepsilon_1 = (1,0,0),\ \varepsilon_2 = (0,1,0),\ \varepsilon_3 = (0,0,1)$

则，$\dim V_1 = 2,\qquad \dim V_2 = 2$

但，$V_1 + V_2 = L(\varepsilon_1,\varepsilon_2) + L(\varepsilon_2,\varepsilon_3) = L(\varepsilon_1,\varepsilon_2,\varepsilon_3) = R^3$

$$\dim(V_1 + V_2) = 3$$

由此还可得到，$\dim(V_1 \cap V_2) = 1$，$V_1 \cap V_2$ 是一直线.

推论：设 V_1,V_2为n维线性空间V的两个子空间，若 $\dim V_1+\dim V_2>n$，则 V_1,V_2必含非零的公共向量. 即 $V_1\cap V_2$中必含有非零向量.

证：由维数公式有

$$\dim(V_1\cap V_2)=\dim V_1+\dim V_2-\dim(V_1+V_2)$$

又 V_1+V_2是V的子空间，∴ $\dim(V_1+V_2)\leq n$

若 $\dim V_1+\dim V_2>n$, 则 $\dim(V_1\cap V_2)>0.$

故$V_1\cap V_2$中含有非零向量.

例1、在 R^n 中，用 V_1,V_2 分别表示齐次线性方程组

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=0 \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{s1}x_1+a_{s2}x_2+\cdots+a_{sn}x_n=0 \end{cases} \quad ①$$

与

$$\begin{cases} b_{11}x_1+b_{12}x_2+\cdots+b_{1n}x_n=0 \\ b_{21}x_1+b_{22}x_2+\cdots+b_{2n}x_n=0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ b_{t1}x_1+b_{t2}x_2+\cdots+b_{tn}x_n=0 \end{cases} \quad ②$$

的解空间，则 $V_1\cap V_2$ 就是齐次线性方程组③

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{s1}x_1 + a_{s2}x_2 + \cdots + a_{sn}x_n = 0 \\ b_{11}x_1 + b_{12}x_2 + \cdots + b_{1n}x_n = 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ b_{t1}x_1 + b_{t2}x_2 + \cdots + b_{tn}x_n = 0 \end{cases} \qquad ③$$

的解空间.

证：设方程组①,②,③分别为

$$AX = 0, \qquad BX = 0, \qquad \begin{pmatrix} A \\ B \end{pmatrix} X = 0$$

设W为③的解空间，任取 $X_0 \in V$，有

$$\begin{pmatrix} A \\ B \end{pmatrix} X_0 = 0,$$ 从而 $$\begin{pmatrix} AX_0 \\ BX_0 \end{pmatrix} = 0,$$ 即

$$AX_0 = BX_0 = 0. \quad \therefore \quad X_0 \in V_1 \cap V_2$$

反之，任 取，$X_0 \in V_1 \cap V_2$，则有

$$AX_0 = BX_0 = 0,$$ 从而 $$\begin{pmatrix} AX_0 \\ BX_0 \end{pmatrix} = \begin{pmatrix} A \\ B \end{pmatrix} X_0 = 0,$$

$$\therefore \quad X_0 \in V$$

故 $V = V_1 \cap V_2.$

设V_1,V_2为线性空间V的两个子空间，由维数公式

$$\dim V_1+\dim V_2=\dim(V_1+V_2)+\dim(V_1\cap V_2)$$

有两种情形：

1) $\dim(V_1+V_2)<\dim V_1+\dim V_2$

此时 $\dim(V_1\cap V_2)>0,$

即，$V_1\cap V_2$必含非零向量.

2) $\dim(V_1+V_2)=\dim V_1+\dim V_2$

此时 $\dim(V_1\cap V_2)=0,$

$V_1\cap V_2$不含非零向量，即 $V_1\cap V_2=\{0\}$

定义 设V_1,V_2为线性空间V的两个子空间，若和 V_1+V_2 中每个向量 x 的分解式

$$x=x_1+x_2,\qquad x_1\in V_1, x_2\in V_2$$

是唯一的，和 V_1+V_2就称为**直和**，记作 $V_1\oplus V_2$.

① 分解式 $x=x_1+x_2$ 唯一的，意即

若有 $x=x_1+x_2=y_1+y_2$， $x_1,y_1\in V_1, x_2,y_2\in V_2$

则 $x_1=y_1, x_2=y_2$.

② 分解式唯一的，不是在任意两个子空间的和中都成立.

例2、设$R^{2\times 2}$的两个子空间为

$$V_1=\left\{A \mid A=\begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix}, x_1-x_2+x_3-x_4=0\right\}$$

$$V_2=L(B_1,B_2), B_1=\begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix}, B_2=\begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$$

求

1. 将 V_1+V_2 表示成生成子空间
2. 求 V_1+V_2 的基和维数
3. 求 $V_1\cap V_2$ 的基和维数

解：先将V_1表示为生成子空间

齐次线性方程 $x_1 - x_2 + x_3 - x_4 = 0$ 的基础解系

$\alpha_1 = [1\ \ 1\ \ 0\ \ 0]^T$ $\alpha_2 = [0\ \ 1\ \ 1\ \ 0]^T$ $\alpha_3 = [0\ \ 0\ \ 1\ \ 1]^T$

所以V_1的一个基为

$$A_1 = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, A_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, A_3 = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}$$

于是$V_1 = L(A_1, A_2, A_3)$，从而有

$$V_1 + V_2 = L(A_1, A_2, A_3, B_1, B_2)$$

矩阵组A_1, A_2, A_3, B_1, B_2在简单基$E_{11}, E_{12}, E_{21}, E_{22}$下的坐标依次为

$$\alpha_1, \alpha_2, \alpha_3, \beta_1 = [1\ \ 0\ \ 2\ \ 3]^T, \beta_2 = [1\ \ -1\ \ 0\ \ 1]^T$$

该向量组的一个最大无关组为 $\alpha_1, \alpha_2, \alpha_3, \beta_2$

矩阵组的一个最大无关组为 A_1, A_2, A_3, B_2

构成了$V_1 + V_2$的一个基，且$\dim(V_1 + V_2)=4$

设 $A \in V_1 \cap V_2$ 则有数k_1, k_2, k_3,与数l_1, l_2,使得

$$A = k_1A_1 + k_2A_2 + k_3A_3 = l_1B_1 + l_2B_2$$

$$k_1A_1 + k_2A_2 + k_3A_3 - l_1B_1 - l_2B_2 = 0$$

比较上式等号两端矩阵的对应元素的值

$$k_1 - l_1 - l_2 = 0 \qquad k_1 + k_2 + l_2 = 0$$

$$k_2 + k_3 - 2l_1 = 0 \qquad k_3 - 3l_1 - l_2 = 0$$

方程组的通解为 $(k_1, k_2, k_3, l_1, l_2)^T = k(1, -1, 3, 1, 0)$

可得 $A = l_1B_1 + l_2B_2 = k\begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix}$

所以$V_1 \cap V_2$的基为$\begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix}$，且维数为1

例如，$\mathbf{R}^3$的子空间

$$V_1 = L(\varepsilon_1, \varepsilon_2),\ V_2 = L(\varepsilon_2, \varepsilon_3),\ V_3 = L(\varepsilon_3)$$

这里，$\varepsilon_1 = (1,0,0),\ \varepsilon_2 = (0,1,0),\ \varepsilon_3 = (0,0,1)$

在和$V_1 + V_2$中，向量的分解式不唯一，如

$$(2,2,2) = (2,3,0) + (0,-1,2) = (2,1,0) + (0,1,2)$$

所以和 $V_1 + V_2$不是直和.

而在和$V_1 + V_3$中，向量(2,2,2)的分解式是唯一的，

$$(2,2,2) = (2,2,0) + (0,0,2)$$

事实上，对 $\forall \alpha = (a_1, a_2, a_3) \in V_1 + V_3$，

都只有唯一分解式：$\alpha = (a_1, a_2, 0) + (0, 0, a_3)$.

故 $V_1 + V_2$是直和.

直和判断定理：V_1+V_2 是直和的充要条件是

$$V_1 \cap V_2 = L(0)$$

证："$\Leftarrow$" 若 $x_1+x_2=0,\ x_1\in V_1,\ x_2\in V_2$.

则有 $x_1=-x_2\in V_1\cap V_2=\{0\}$

$\therefore x_1=x_2=0,$ 即V_1+V_2 是直和.

"$\Rightarrow$" 任取 $x\in V_1\cap V_2,$

$0=x+(-x),\ x\in V_1,\ -x\in V_2.$

由于V_1+V_2是直和，零向量分解式唯一，

$\therefore\ x=-x=0.$ 故 $V_1\cap V_2=\{0\}.$

推论：和 V_1+V_2 是直和

$$\Leftrightarrow \quad \dim(V_1+V_2)=\dim V_1+\dim V_2$$

证：由维数公式

$$\dim V_1+\dim V_2=\dim(V_1+V_2)+\dim(V_1\cap V_2)$$

有， $\dim(V_1+V_2)=\dim V_1+\dim V_2$

$$\Leftrightarrow \quad \dim(V_1\cap V_2)=0$$

$$\Leftrightarrow \quad V_1\cap V_2=\{0\}$$

$\Leftrightarrow \quad V_1+V_2$ 是直和.

总之，设 V_1,V_2 为线性空间V的子空间，则下面四个条件等价:

1） V_1+V_2是直和

2） 零向量分解式唯一

3） $V_1\cap V_2=\{0\}$

4） $\dim(V_1+V_2)=\dim V_1+\dim V_2$

设 $x_1, x_2, \cdots, x_r$； $y_1, y_2, \cdots, y_s$分别是线性子空间 V_1, V_2 的一组基，则

V_1+V_2是直和 $\Leftrightarrow x_1, x_2, \cdots, x_r, y_1, y_2, \cdots, y_s$ 线性无关.

证：由题设，$V_1 = L(x_1, x_2, \cdots, x_r)$， $\dim V_1 = r$

$$V_2 = L(y_1, y_2, \cdots, y_s), \ \dim V_2 = s$$

$$\therefore \quad V_1+V_2 = L(x_1, x_2, \cdots, x_r, y_1, y_2, \cdots, y_s).$$

若$x_1, x_2, \cdots, x_r, y_1, y_2, \cdots, y_s$线性无关，

则它是 V_1+V_2的一组基. 从而有

$$\dim(V_1+V_2)=r+s=\dim V_1+\dim V_2$$

$\therefore$ V_1+V_2 是直和.

反之, 若 V_1+V_2 直和，则

$$\dim(V_1+V_2)=\dim V_1+\dim V_2=r+s$$

从而 $x_1,x_2,\cdots,x_r,y_1,y_2,\cdots,y_s$ 的秩为 $r+s$.

所以 $x_1,x_2,\cdots,x_r,y_1,y_2,\cdots,y_s$ 线性无关.

定义：

$V_1,V_2,\cdots,V_s$ 都是线性空间V的子空间，若和

$\sum_{i=1}^{s} V_i = V_1 + V_2 + \cdots + V_s$ 中每个向量 x 的分解式

$$x = x_1 + x_2 + \cdots + x_s,\ \ x_i \in V_i, i = 1,2,\ldots,s$$

是唯一的，则和 $\sum_{i=1}^{s} V_i$ 就称为直和，记作

$$V_1 \oplus V_2 \oplus \cdots \oplus V_s$$

判定方法：

设 $V_1, V_2, \cdots, V_s$ 都是线性空间 V 的子空间，则下面

四个条件等价:

1）$W = \sum_{i=1}^{s} V_i$ 是直和

2）零向量分解式唯一，即

$x_1 + x_2 + \cdots + x_s = 0$, $x_i \in V_i$，必有 $x_i = 0$, $i = 1, 2, \ldots, s$

3）$V_i \cap \sum_{j \neq i} V_j = \{0\}, i = 1, 2, \cdots, s$

4）$\dim W = \sum_{i=1}^{s} \dim V_i$

作业

- P26：10、11、12

本讲主要内容

- 集合的映射
- 线性变换
- 线性变换的简单性质
- 线性变换的运算

映射

设S、S'是给定的两个非空集合，如果有 一个对应法则 σ，通过这个法则 σ对于S中的每一个元素a，都有S'中一个唯一确定的元素a'与它对应, 则称 σ为S到S'的一个映射，记作：$\sigma: S \to S'$ 或 $S \xrightarrow{\sigma} S'$

称 a'为 a 在映射 σ下的象，而 a' 称为a在映射 σ下的原象，记作 $\sigma(a)=a'$ 或 $\sigma: a \mapsto a'$.

- S到S自身的映射，也称为S到S的变换

- 关于S到S'的映射σ

1）S与S'可以相同，也可以不同

2）对于S中每个元素a，需要有S'中一个唯一确定的元素a'与它对应

3）一般，S'中元素不一定都是S中元素的像

4）S中不相同元素的像可能相同

5）两个集合之间可以建立多个映射

- 若$\forall a \neq a' \in S$，都有$\sigma(a) \neq \sigma(a')$，则称为单射
- 若$\forall b \in S'$，都存在$a \in S$，使得$\sigma(a) = b$，则称为满射
- 如果既是单射又是满射，则称为双射，或称一一对应

例　判断下列M到M'对应法则是否为映射

1）$M=\{a, b, c\}$、$M'=\{1, 2, 3, 4\}$

σ：$\sigma(a)=1$，$\sigma(b)=1$，$\sigma(c)=2$　（是）

δ：$\delta(a)=1, \delta(b)=2, \delta(c)=3, \delta(c)=4$　（不是）

τ：$\tau(b)=2$，$\tau(c)=4$　（不是）

2）$M=Z$，$M'=Z^{+}$，

σ：$\sigma(n)=|n|, \forall n \in Z$　（不是）

τ：$\tau(n)=|n|+1, \forall n \in Z$　（是）

3） $S=R^{n\times n}$， $S'=K$， （K为数域）

σ： $\sigma(A)=|A|, \forall A\in R^{n\times n}$ （是）

4） $S=K$， $S'=R^{n\times n}$， （K为数域）

τ： $\tau(a)=aE, \forall a\in K$ （E为n级单位矩阵） （是）

5） S、S'为任意两个非空集合，a_0是S'中的一个固定元素.

σ： $\sigma(a)=a_0, \forall a\in M$ （是）

6） $S=S'=P[x]$

σ： $\sigma(f(x))=f'(x), \forall f(x)\in P[x]$ （是）

- 设 σ_1，σ_2都是集合S 到集合S'的两个映射，若对S 的每个元素a 都有 $\sigma_1(a)=\sigma_2(a)$ 则称它们相等，记作 $\sigma_1=\sigma_2$
- 设 σ，τ是集合S 到S_1，集合S_1到S_2的映射，映射的乘积 $\tau\sigma$ 定义为

$$(\tau\sigma)(\alpha)=\tau(\sigma(\alpha)),\quad \alpha\in \boldsymbol{S}$$

- 设 σ，τ，μ是集合S 到S_1,S_1到S_2,S_2到S_3的映射，则映射的乘积满足结合律，但不满足交换律

$$(\tau\sigma)\mu=\tau(\sigma\mu) \qquad\qquad \tau\sigma\neq\sigma\tau$$

- 设 σ 都是集合 S 到集合 S' 的一一对应映射，

1.若 $\forall a \in S, \exists \sigma(a) \in S'; \forall b \in S', \exists a \in S, \text{st.} \sigma(a) = b$

2.若 $\forall a, b \in S$,且 $a \neq b$ 有 $\sigma(a) \neq \sigma(b)$

或者 $\sigma(a) = \sigma(b)$,就有 $a = b$

就称 σ 是集合 S 到集合 S' 的同构映射，且称集合 S 到集合 S' 是同构的

- 由不高于n次的实系数多项式构成的空间与实数域上n+1维的全体向量构成的空间同构，比如

$$a_0 + a_1 x + a_2 x^2 + a_3 x^3 \leftrightarrow (a_0, a_1, a_2, a_3)$$

■设V为数域K上的线性空间，若变换 $T:V\to V$

满足： $\forall x,y\in V$， $k\in K$

$$T(x+y)=T(x)+T(y)$$

$$T(kx)=kT(x)$$

或者

$$T(kx+ly)=k(Tx)+l(Ty)$$

事实上， $\forall x,y\in V$， $\forall k\in K$,

$$K(x+y)=k(x+y)=kx+ky=K(x)+K(y),$$

$$K(mx)=kmx=mkx=mK(x).$$

由数k决定的数乘变换：$K:V\to V$， $x\mapsto kx$， $\forall x\in V$

例1. $V=R^2$(实数域上二维向量空间)，把V中每一向量绕坐标原点旋转 θ 角，就是一个线性变换，用 T_θ 表示，即

$$T_\theta: R^2 \to R^2,\ \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x' \\ y' \end{pmatrix}$$

这里，$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$

易验证：$\forall x, y \in R^2, \forall k \in R$

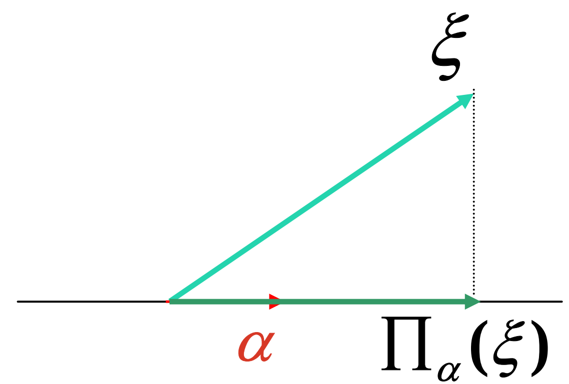
$$T_\theta(x+y) = T_\theta(x) + T_\theta(y)$$

$$T_\theta(kx) = kT_\theta(y)$$

例2． $V=R^3$，$\alpha\in V$ 为一固定非零向量，把V中每一个向量 ξ 变成它在 α 上的内射影是V上的一个线性变换. 用 $\prod_\alpha$ 表示，即

$$\prod_\alpha : R^3 \to R^3,\ \ \xi \mapsto \frac{(\alpha,\xi)}{(\alpha,\alpha)}\alpha,\ \ \forall \xi\in R^3$$

这里 $(\alpha,\xi),(\alpha,\alpha)$ 表示内积.



易验证： $\forall \xi,\eta\in R^3, \forall k\in R$

$$\prod_\alpha(\xi+\eta)=\prod_\alpha(\xi)+\prod_\alpha(\eta)$$

$$\prod_\alpha(k\xi)=k\prod_\alpha(\xi)$$

例3．线性空间 R^n 中，求微分是一个线性变换，用D表示，即

$$D:V\to V,\quad D(f(x))=f'(x),\ \forall f(x)\in V$$

证明：$\forall f(x),g(x)\in P_n$ 和 $\forall k,l\in R$ 有

$$D\left(f(t)+g(t)\right)=\left(f(t)+g(t)\right)'=f'(t)+g'(t)$$
$$=D\left(f(t)\right)+D\left(g(t)\right)$$

$$D\left(kf(t)\right)=\left(kf(t)\right)'=kf'(t)\quad =kD\left(f(t)\right)$$

因此D是一个线性变换.

例4. 闭区间 $[a,b]$上的全体连续函数构成的线性空间 $C(a,b)$ 上的变换

$$J: C(a,b) \to C(a,b),\quad J(f(t)) = \int_a^t f(x)dx$$

是一个线性变换.

证明：$\forall f(x), g(x) \in P_n$ 和 $\forall k, l \in R$ 有

$$\begin{aligned} J(kf(t)+lg(t)) &= \int_a^t (kf(u)+lg(u))du \\ &= k\int_a^t f(u)du + l\int_a^t g(u)du \\ &= kJ(f(t)) + lJ(g(t)) \end{aligned}$$

因此J是一个线性变换.

线性变换的简单性质

1．T 为V的线性变换，则

$$T(0)=0,\ \ T(-x)=-T(x).$$

2．线性变换保持线性组合及关系式不变，即

若 $x=k_1x_1+k_2x_2+\cdots+k_rx_r,$

则 $T(x)=k_1T(x_1)+k_2T(x_2)+\cdots+k_rT(x_r).$

3．线性变换把线性相关的向量组的变成线性相关的向量组. 即

若 $x_1,x_2,\cdots,x_r$ 线性相关，则 $T(x_1),T(x_2),\cdots,T(x_r)$ 也线性相关.

事实上，若有不全为零的数 $k_1,k_2,\cdots,k_r$ 使

$$k_1x_1+k_2x_2+\cdots+k_rx_r=0$$

则由2即有，$k_1T(x_1)+k_2T(x_2)+\cdots+k_rT(x_r)=0.$

注意：3的逆不成立，即 $T(x_1),T(x_2),\cdots,T(x_r)$ 线性相关，$x_1,x_2,\cdots,x_r$未必线性相关.

事实上，线性变换可能把线性无关的向量组变成线性相关的向量组. 如零变换.

练习：下列变换中，哪些是线性变换？

1．在 R^3中，$T\left(x_1, x_2, x_3\right) = (2x_1, x_2, x_2 - x_3)$. √

2．在$P[x]_n$中，$T\left(f(x)\right) = f^2(x)$. ×

3．在线性空间V中，$T\left(\xi\right) = \xi + \alpha$，$\alpha \in V$ 非零固定. ×

4．在 $P^{n\times n}$中，$T\left(X\right) = AX$，$A \in P^{n\times n}$ 固定. √

5．复数域C看成是自身上的线性空间，$T(x) = \overline{x}$. ×

6．C看成是实数域R上的线性空间，$T(x) = \overline{x}$. √

线性变换的运算

一、线性变换的和

二、线性变换的数量乘法

三、线性变换的乘积

四、线性变换的逆

五、线性变换的多项式

1.线性变换的和

设T_1,T_2 为线性空间V的两个线性变换，定义它们

的和 T_1+T_2为 :$\left(T_1+T_2\right)\left(x\right)=T_1x+T_2x,\forall x\in V$

则 T_1+T_2 也是V的线性变换.

事实上，$(T_1+T_2)(x+y)\ =T_1(x+y)+T_2(x+y)$

$$=T_1x+T_1y+T_2x+T_2y=(T_1+T_2)x+(T_1+T_2)y$$

$$(T_1+T_2)(kx)=T_1(kx)+T_2(kx)\ =k(T_1x)+k(T_2x)$$

$$=k(T_1x+T_2x)\quad =k(T_1+T_2)x$$

负变换

设 T 为线性空间V的线性变换，定义变换 $-T$ 为：

$$(-T)(x) = -T(x), \qquad \forall x \in V$$

则 $-T$ 也为V的线性变换，称之为 T 的负变换.

线性变换和的基本性质

（1）满足交换律：$T_1 + T_2 = T_2 + T_1$

（2）满足结合律：$\left(T_1 + T_2\right) + T_3 = T_1 + \left(T_2 + T_3\right)$

（3） $T_0 + T_1 = T_1$ ，T_0 为零变换.

（4） $(-T) + T = T_0$

线性变换的数量乘法

设 T 为线性空间V的线性变换，$k \in K$, 定义 k 与 T 的数量乘积 kT 为：

$$(kT)(x) = kT(x), \quad \forall x \in V$$

则 kT 也是V的线性变换.

线性变换数量乘法的基本性质

(1) $k(T_1 + T_2) = kT_1 + kT_2$

(2) $(k + l)T = kT + lT$

(3) $(kl)T = k(lT)$

(4) $1T = T$

注： 线性空间V上的全体线性变换所成集合对于线性变换的加法与数量乘法构成数域K上的一个线性空间，记作 $\mathrm{Hom}(V,V) \triangleq$ {T|T是数域K上线性空间V的线性变换}

线性变换的乘积

设 T_1, T_2 为线性空间V的两个线性变换，定义它们的**乘积** T_1T_2 为： $\left(T_1T_2\right)\left(x\right)=T_1\left(T_2x\right),\quad \forall x\in V$

则 T_1T_2 也是V的线性变换.

事实上, $(T_1T_2)(x+y)=T_1(T_2(x+y))=T_1(T_2(x)+T_2(y))$

$$=T_1(T_2x)+T_1(T_2y)=(T_1T_2)x+(T_1T_2)y$$

$$(T_1T_2)(kx)=T_1\left(T_2(kx)\right)=T_1\left(k(T_2x)\right)$$

$$=k\left(T_1(T_2x)\right)=k\left(T_1T_2\right)x$$

线性变换乘积的基本性质

（1）满足结合律：$(T_1T_2)T_3 = T_1(T_2T_3)$

（2） $T_eT = TT_e = T$ ，T_e 为单位变换

（3）交换律一般不成立，即一般地，

$$T_1T_2 \neq T_2T_1$$

（4）乘法对加法满足左、右分配律：

$$T_1(T_2 + T_3) = T_1T_2 + T_1T_3$$

$$(T_1 + T_2)T_3 = T_1T_3 + T_2T_3$$

例1. 线性空间 $R[x]$中，线性变换

$$D\big(f(x)\big)=f'(x)$$

$$J\big(f(x)\big)=\int_0^x f(t)dt$$

$$(DJ)\big(f(x)\big)=D\left(\int_0^x f(t)dt\right)=f(x),$$ 即 $DJ=T_e$

而，

$$(JD)\big(f(x)\big)=J\big(f'(x)\big)=\int_0^x f'(t)dt=f(x)-f(0)$$

$$\therefore\quad DJ\neq JD.$$

例2. 设A、$B\in R^{n\times n}$ 为两个取定的矩阵，定义变换

$$T_1(X)=AX,$$

$$\forall X\in R^{n\times n}$$

$$T_2(X)=XB,$$

则T_1,T_2皆为 $R^{n\times n}$的线性变换，且对 $\forall X\in R^{n\times n}$,有

$$T_1T_2(X)=T_1\big(T_2(X)\big)=T_1(XB)=A(XB)=AXB,$$

$$T_2T_1(X)=T_2\big(T_1(X)\big)=T_2(AX)=(AX)B=AXB,$$

$$\therefore\quad T_1T_2=T_2T_1$$

线性变换的逆

设 T 为线性空间V的线性变换，若有V的变换 S 使

$$ST = TS = T_e$$

则称 T 为可逆变换，称 S 为 T 的逆变换，记作 T^{-1}.

2．基本性质

(1) 可逆变换 T 的逆变换 T^{-1} 也是V的线性变换.

证：对 $\forall x, y \in V$， $\forall k \in K$,

$$T^{-1}(x+y) = T^{-1}\left(\left(TT^{-1}\right)(x) + \left(TT^{-1}\right)(y)\right)$$

$$= T^{-1}\left(T\left(T^{-1}(x) + T^{-1}(y)\right)\right)$$

$$= \left(T^{-1}T\right)\left(T^{-1}(x) + T^{-1}(y)\right)$$

$$= T^{-1}(x) + T^{-1}(y)$$

$$T^{-1}(kx) = T^{-1}\left(k\left(TT^{-1}\right)(x)\right) = T^{-1}\left(k\left(T\left(T^{-1}(x)\right)\right)\right)$$

$$= T^{-1}\left(T\left(k\left(T^{-1}(x)\right)\right)\right) = k\left(T^{-1}(x)\right) = kT^{-1}(x)$$

$\therefore$ T^{-1} 是V的线性变换.

(2) 线性变换 T 可逆 $\Leftrightarrow$ 线性变换 T 是一一对应.

证："$\Rightarrow$" 设 T 为线性空间V上可逆线性变换.

任取 $x, y \in V$，若 $T(x) = T(y)$，则有

$$x = (T^{-1}T)(x) = T^{-1}(T(x)) = T^{-1}(T(y))$$

$$= (T^{-1}T)(y) = y \qquad \therefore\ T \text{ 为单射.}$$

其次，对 $\forall y \in V$，令 $x = T^{-1}(y)$，则 $x \in V$,且

$T(x) = T(T^{-1}(y)) = TT^{-1}(y) = y$. $\therefore$ T 为满射.

故 T 为一一对应.

"$\Leftarrow$" 若 T 为一一对应，易证 T 的逆映射 S 也为V的线性变换，且 $TS = ST = T_e$ 故 T 可逆，$S = T^{-1}$.

(3) 设 $x_1, x_2, \cdots, x_n$ 是线性空间V的一组基，T 为V的线性变换，则 T 可逆当且仅当 $T(x_1), T(x_2), \cdots, T(x_n)$ 线性无关.

证："$\Rightarrow$" 设 $k_1 T(x_1) + k_2 T(x_2) + \cdots + k_n T(x_n) = 0.$

于是 $T(k_1 x_1 + k_2 x_2 + \cdots + k_n x_n) = 0$

因为 T 可逆，由(2)，T 为单射，又 $T(0) = 0,$

$$\therefore \quad k_1 x_1 + k_2 x_2 + \cdots + k_n x_n = 0$$

而 $x_1, x_2, \cdots, x_n$线性无关，所以 $k_i = 0, \ i = 1,2,\cdots,n.$

故$T(x_1), T(x_2), \cdots, T(x_n)$线性无关.

" $\Leftarrow$ " 若$T(x_1), T(x_2), \cdots, T(x_n)$线性无关，则它也为V的一组基. 因而，对 $\forall y \in V$, 有

$$y = k_1 T(x_1) + k_2 T(x_2) + \cdots + k_n T(x_n),$$

即有 $T(k_1 x_1 + k_2 x_2 + \cdots + k_n x_n) = y$

$\therefore \ T$ 为满射.

其次，任取 $x, y \in V$，设 $x = \sum_{i=1}^{n} a_i x_i$， $y = \sum_{i=1}^{n} b_i x_i$，

若 $T(x) = T(y)$，则有

$$\sum_{i=1}^{n} a_i T(x_i) = \sum_{i=1}^{n} b_i T(x_i),$$

$\because$ $T(x_1), T(x_2), \cdots, T(x_n)$ 线性无关

$\therefore$ $a_i = b_i$， $i = 1, 2, \cdots, n$， 即 $x = y$.

从而，T 为单射. 故 T 为一一对应.

由(2)， T 为可逆变换.

(4) 可逆线性变换把线性无关的向量组变成线性无关的向量组.

证：设 T 为线性空间V的可逆变换，$x_1, x_2, \cdots, x_r \in V$ 线性无关. 若 $k_1 T\left(x_1\right)+k_2 T\left(x_2\right)+\cdots+k_r T\left(x_r\right)=0.$

则有， $T(k_1 x_1+k_2 x_2+\cdots+k_r x_r)=0$

又 T 可逆，于是 T 是一一对应，且 $T(0)=0$

$\therefore \quad k_1 x_1+k_2 x_2+\cdots+k_r x_r=0$

由 $x_1, x_2, \cdots, x_r$ 线性无关，有 $k_1=k_2=\cdots=k_r=0.$

故 $T(x_1), T(x_2), \cdots, T(x_r)$ 线性无关.

五、线性变换的多项式

1．线性变换的幂

设 T 为线性空间V的线性变换，n为自然数，定义

$$T^n = \underbrace{T \cdots T}_{n},$$

称之为 T 的**n次幂**.

当 $n=0$ 时，规定 $T^0 = T_e$（单位变换）．

注：

① 易证 $T^{m+n}=T^mT^n,\ \left(T^m\right)^n=T^{mn},\qquad m,n\geq 0$

② 当 T为可逆变换时，定义T 的负整数幂为

$$T^{-n}=\left(T^{-1}\right)^n$$

③ 一般地，$\left(TS\right)^n\neq T^nS^n.$

2．线性变换的多项式

设 $f(x)=a_m x^m+\cdots+a_1 x+a_0\in P[x]$,

T 为V的一个线性变换，则

$$f(T)=a_m T^m+\cdots+a_1 T+a_0 T_e$$

也是V的一个线性变换，称 $f(T)$为线性变换 T 的多项式.

注：① 在 $P[x]$ 中，若

$$h(x)=f(x)+g(x),\quad p(x)=f(x)g(x)$$

则有，$h(T)=f(T)+g(T)$,

$$p(T)=f(T)g(T)$$

② 对 $\forall f(x),g(x)\in P[x]$, 有

$$f(T)+g(T)=g(T)+f(T)$$

$$f(T)g(T)=g(T)f(T)$$

即线性变换的多项式满足加法和乘法交换律.

练习：设T,S为线性变换，若 $TS-ST=T_e$，

证明：$T^kS-ST^k=kT^{k-1}$，　　$k>1.$

证：对k作数学归纳法.

当k=2时，若 $TS-ST=T_e$　　　　①

对①两端左乘 T，得　$T^2S-TST=T$，

对①两端右乘 T，得　$TST-ST^2=T$

上两式相加，即得　$T^2S-ST^2=2T=2T^{2-1}$.

假设命题对 $k-1$时成立，即

$$T^{k-1}S - ST^{k-1} = (k-1)T^{k-2}. \qquad ②$$

对②两端左乘 T，得

$$T^{k}S - TST^{k-1} = (k-1)T^{k-1}, \qquad ③$$

对①两端右乘 T^{k-1}, 得

$$TST^{k-1} - ST^{k} = T^{k-1}, \qquad ④$$

③＋④，得 $T^{k}S - ST^{k} = kT^{k-1}.$

由归纳原理，命题成立 #

本讲主要内容

- 线性变换在给定基下的矩阵
- 线性变换的矩阵表示
- 象与原象坐标间的关系
- 线性变换在不同基下矩阵之间的关系
- 线性变换的特征值与特征向量

- 设$x_1,x_2,\cdots,x_n$是线性空间V^n的一组基，T为V^n的线性变换. 则对任意 $x\in V^n$ 存在唯一的一组数 $k_1,k_2,\cdots,k_n\in K$, 使 $x=k_1x_1+k_2x_2+\cdots+k_nx_n$

从而，$T(x)=k_1T(x_1)+k_2T(x_2)+\cdots+k_nT(x_n)$.

- 由此知，$T(x)$ 由$T(x_1),T(x_2),\cdots,T(x_n)$完全确定.

设$x_1,x_2,\cdots,x_n$是线性空间V的一组基，T_1,T_2为V的线性变换，若 $T_1(x_i)=T_2(x_i)$，$i=1,2,\cdots,n.$

则 $T_1=T_2$

证：对 $\forall x\in V$， $x=k_1x_1+k_2x_2+\cdots+k_nx_n$

$$T_1(x)=k_1T_1(x_1)+k_2T_1(x_2)+\cdots+k_nT_1(x_n)$$

$$T_2(x)=k_1T_2(x_1)+k_2T_2(x_2)+\cdots+k_nT_2(x_n)$$

由已知，即得 $T_1(x)=T_2(x)$ $\quad\therefore\ T_1=T_2$

- 由此知，一个线性变换完全由它在一组基上的作用所决定.

设 $x_1, x_2, \cdots, x_n$是线性空间V的一组基，对V中

任意n个向量 $\alpha_1, \alpha_2, \cdots, \alpha_n$，都存在线性变换$T$使

$$T(x_i) = \alpha_i, \qquad i = 1, 2, \cdots, n$$

证：$\forall y \in V$， $y = k_1 x_1 + k_2 x_2 + \cdots + k_n x_n$

定义 $T : V \to V$， $T(y) = k_1 \alpha_1 + k_2 \alpha_2 + \cdots + k_n \alpha_n$,

易知 T 为V的一个变换，下证它是线性的.

任取 y, $z \in V$，设 $y = \sum_{i=1}^{n} b_i x_i$， $z = \sum_{i=1}^{n} c_i x_i$

则 $y+z=\sum_{i=1}^{n}(b_i+c_i)\ x_i$, $ky=\sum_{i=1}^{n}(kb_i)x_i$

于是 $T(y+z)=\sum_{i=1}^{n}(b_i+c_i)\ \alpha_i=\sum_{i=1}^{n}b_i\alpha_i+\sum_{i=1}^{n}c_i\alpha_i$

$$=T(y)+T(z)$$

$$T(kx)=\sum_{i=1}^{n}(kb_i)\alpha_i=k\sum_{i=1}^{n}b_i\alpha_i=kT(y)$$

$\therefore$ T 为V的线性变换.

又 $x_i=0x_1+\cdots+0x_{i-1}+x_i+0x_{i+1}+\cdots+0x_n$

$\therefore$ $T(x_i)=\alpha_i$, $i=1,2,\cdots,n$

由此即得

定理 设 $x_1, x_2, \cdots, x_n$ 为线性空间V的一组基，对V中任意n个向量 $\alpha_1, \alpha_2, \cdots, \alpha_n$，存在唯一的线性变换$T$使

$$T\left(x_i\right) = \alpha_i, \qquad i = 1, 2, \cdots, n.$$

1．线性变换的矩阵

设 $x_1, x_2, \cdots, x_n$为数域P上线性空间V的一组基，T为V的线性变换．基向量的象可以被基线性表出,设

$$\begin{cases} T(x_1) = a_{11}x_1 + a_{21}x_2 + \cdots + a_{n1}x_n \\ T(x_2) = a_{12}x_1 + a_{22}x_2 + \cdots + a_{n2}x_n \\ \cdots\cdots \\ T(x_n) = a_{1n}x_1 + a_{2n}x_2 + \cdots + a_{nn}x_n \end{cases}$$

用矩阵表示即为

$$T\left(x_1, x_2, \cdots, x_n\right) = \left(Tx_1, Tx_2, \cdots, Tx_n\right) = \left(x_1, x_2, \cdots, x_n\right)A$$

其中 $$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix},$$

矩阵A称为线性变换T在基 $x_1, x_2, \cdots, x_n$下的矩阵.

注： ① 给定V^n的基 $x_1, x_2, \cdots, x_n$ 和线性变换T,

矩阵A是唯一的.

② 单位变换在任意一组基下的矩阵皆为单位矩阵；

零变换在任意一组基下的矩阵皆为零矩阵；

数乘变换在任意一组基下的矩阵皆为数量矩阵；

例1. 设线性空间 V^3 的线性变换 T 为

$$T(x_1, x_2, x_3) = (x_1, x_2, x_1 + x_2)$$

求 T 在标准基 e_1, e_2, e_3 下的矩阵.

解： $\because$ $T(e_1) = T(1,0,0) = (1,0,1)$

$$T(e_2) = T(0,1,0) = (0,1,1)$$

$$T(e_3) = T(0,0,1) = (0,0,0)$$

$$\therefore \quad T(x_1, x_2, x_3) = (x_1, x_2, x_3)\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{pmatrix}$$

例2. 设线性空间$P_n[t]$的线性变换T为 $Tf(t)=f'(t)$

基 I 为 $f_0=1, f_1=t, f_2=\frac{t^2}{2!},\cdots, f_n=\frac{t^n}{n!}$

基 II 为 $g_0=1, g_1=t, g_2=t^2,\cdots, g_n=t^n$

记T在基 I 下的矩阵为A_1，在基 II 下的矩阵为A_2

$$Tf_0=0, Tf_1=f_0, Tf_2=f_1,\cdots, Tf_n=f_{n-1}$$
$$Tg_0=0, Tg_1=g_0, Tg_2=2g_1,\cdots, Tg_n=ng_{n-1}$$

$$\therefore \quad A_1=\begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ & 0 & 1 & \cdots & 0 \\ & & \ddots & \ddots & \vdots \\ & & & 0 & 1 \\ & & & & 0 \end{bmatrix}, A_2=\begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ & 0 & 2 & \cdots & 0 \\ & & \ddots & \ddots & \vdots \\ & & & 0 & n \\ & & & & 0 \end{bmatrix}$$

线性变换T的值域： $R(T)=\{y \mid y=Tx, x\in V\}$

线性变换T的核： $N(T)=\{x \mid Tx=0, x\in V\}$

定理：线性空间V的线性变换T的值域和核都是V的线性子空间

$$V\neq\varnothing\Rightarrow R(T)\neq\varnothing,$$

$$\forall y_1\in R(T)\Rightarrow\exists x_1\in V,\text{st } y_1\in Tx_1$$

$$\forall y_2\in R(T)\Rightarrow\exists x_2\in V,\text{st } y_2\in Tx_2$$

$$y_1+y_2=Tx_1+Tx_2=T\left(x_1+x_2\right)\in R(T)$$

$$ky_1=k\left(Tx_1\right)=T\left(kx_1\right)\in R(T)$$

所以，$R(T)$ 是V的线性子空间

定义：线性变换T的值域 $R(T)$的维数称为**T的秩**

T的核 $N(T)$ 的维数称为**T的亏（零度）**

例: 在线性空间 $P_n[x]$ 中，令

$$T\left(f(x)\right)=f'(x)$$

则 $$T\left(P_n[x]\right)=P_{n-1}[x]$$

$$N(T)=P$$

所以T的秩为$n-1$，T的零度为1.

定理: 设T是n 维线性空间V的线性变换，$x_1, x_2, \cdots, x_n$

是V的一组基，T在这组基下的矩阵是A，则

1）T 的值域 $R(T)$ 是由基象组生成的子空间，即

$$R(T) = L\left(T(x_1), T(x_2), \cdots, T(x_n)\right)$$

2）T 的秩＝A的秩.

证：1） $\forall y \in V$，设 $y = k_1 x_1 + k_2 x_2 + \cdots + k_n x_n$，

于是 $T(y) = k_1 T(x_1) + k_2 T(x_2) + \cdots + k_n T(x_n)$

$$\in L\left(T(x_1), T(x_2), \cdots, T(x_n)\right)$$

即 $R(T) \subseteq L\left(T(x_1), T(x_2), \cdots, T(x_n)\right)$

又对 $\forall k_1 T(x_1) + k_2 T(x_2) + \cdots + k_n T(x_n)$

有 $k_1 T(x_1) + k_2 T(x_2) + \cdots + k_n T(x_n)$

$$= T(k_1 x_1 + k_2 x_2 + \ldots + k_n x_n) \in R(T)$$

$$\therefore \quad L\left(T(x_1), T(x_2), \cdots, T(x_n)\right) \subseteq R(T).$$

因此，$R(T) = L\left(T(x_1), T(x_2), \cdots, T(x_n)\right)$.

2）由1），T 的秩等于基象组 $T(x_1),T(x_2),\cdots,T(x_n)$ 的秩，又

$$\big(T(x_1),T(x_2),\cdots,T(x_n)\big)=(x_1,x_2,\cdots,x_n)A$$

$T(x_1),T(x_2),\cdots,T(x_n)$ 的秩等于矩阵A的秩

$\therefore$ 秩$(T)=$秩(A).

设T为n维线性空间V的线性变换，则

$$T\text{的秩}+T\text{的零度}=n$$

即 $\dim R(T)+\dim N(T)=n.$

证明：设T的零度等于r，在核$N(T)$中取一组基

$$x_1,x_2,\cdots,x_r$$

并把它扩充为V的一组基：$x_1,x_2,\cdots,x_r,\cdots,x_n$

$R(T)$ 是由基象组 $T(x_1),T(x_2),\cdots,T(x_n)$ 生成的.

但 $T(x_i)=0,\qquad i=1,2,\cdots,r.$

$\therefore\quad R(T)=L\big(T(x_{r+1}),\cdots,T(x_n)\big)$

下证$T(x_{r+1}),\cdots,T(x_n)$为$R(T)$的一组基，即证它们线性无关.

设 $k_{r+1}T(x_{r+1})+\cdots+k_nT(x_n)=0$

则有 $T\big(k_{r+1}x_{r+1}+\cdots+k_nx_n\big)=0$

$\therefore\quad y=k_{r+1}x_{r+1}+\cdots+k_nx_n\in N(T)$

即y可被$x_1,x_2,\cdots,x_r$线性表出.

设 $y = k_1 x_1 + k_2 x_2 + \cdots + k_r x_r$

于是有 $k_1 x_1 + k_2 x_2 + \cdots + k_r x_{r,} - k_{r+1} x_{r+1} - \cdots - k_n x_n = 0$

由于 $x_1, x_2, \cdots, x_n$ 为V的基.

$\therefore k_1 = k_2 = \cdots = k_n = 0$

故$T(x_{r+1}), \cdots, T(x_n)$线性无关，即它为$R(T)$的一组基.

$\therefore$ T 的秩＝n－r .

因此，T 的秩＋ T 的零度＝n.

注意：

虽然 $R(T)$ 与 $N(T)$ 的维数之和等于n，但是

$R(T)+N(T)$ 未必等于V.

例: 在线性空间 $P_n[x]$ 中，令 $T\left(f(x)\right)=f'(x)$

则 $R\left(T\right)=P_{n-1}[x]$ $N(T)=P$

$$R(T)+N(T)=P_{n-1}[x]$$

定理：设$x_1, x_2, \cdots, x_n$为数域K上线性空间V的一组基，在这组基下，V的每一个线性变换都与 $P^{n\times n}$ 中的唯一一个矩阵对应，且具有以下性质：

① 线性变换的和对应于矩阵的和；

②线性变换的数量乘积对应于矩阵的数量乘积；

③线性变换的乘积对应于矩阵的乘积；

④ 可逆线性变换与可逆矩阵对应，且逆变换对应于逆矩阵.

证：设T_1,T_2为两个线性变换，它们在基$x_1,x_2,\cdots,x_n$下的矩阵分别为A、B，即

$$T_1(x_1,x_2,\cdots,x_n)=(x_1,x_2,\cdots,x_n)A$$

$$T_2(x_1,x_2,\cdots,x_n)=(x_1,x_2,\cdots,x_n)B$$

① $\because\ (T_1+T_2)(x_1,x_2,\cdots,x_n)$

$$=T_1(x_1,x_2,\cdots,x_n)+T_2(x_1,x_2,\cdots,x_n)$$

$$=(x_1,x_2,\cdots,x_n)A+(x_1,x_2,\cdots,x_n)B$$

$$=(x_1,x_2,\cdots,x_n)(A+B)$$

$\therefore\ T_1+T_2$在基$x_1,x_2,\cdots,x_n$下的矩阵为A＋B.

② $\because\ (T_1T_2)(x_1,x_2,\cdots,x_n)=T_1(T_2(x_1,x_2,\cdots,x_n))$

$$=T_1((x_1,x_2,\cdots,x_n)B)=T_1(x_1,x_2,\cdots,x_n)B$$
$$=(x_1,x_2,\cdots,x_n)(AB)$$

$\therefore\ T_1T_2$在基 $x_1,x_2,\cdots,x_n$下的矩阵为AB.

③ $\because\ (kT_1)(x_1,x_2,\cdots,x_n)=((kT_1)(x_1),\cdots,(kT_1)(x_n))$

$$=(kT_1(x_1),\cdots,kT_1(x_n))=k(T_1(x_1),\cdots,T_1(x_n))$$
$$=kT_1(x_1,x_2,\cdots,x_n)=k(x_1,x_2,\cdots,x_n)A$$
$$=(x_1,x_2,\cdots,x_n)(kA)$$

$\therefore\ kT_1$ 在基$x_1,x_2,\cdots,x_n$下的矩阵为 kA.

④ 由于单位变换(恒等变换) T_e 对应于单位矩阵E.

所以， $T_1T_2 = E$

与 AB＝BA＝E

相对应.

因此，可逆线性变换 T_1 与可逆矩阵A对应，且逆变换 T_1^{-1} 对应于逆矩阵 A^{-1}.

推论： $f(t)=a_m t^m + a_{m-1}t^{m-1}+\cdots+a_1 t+a_0, t\in K$

T为线性空间V的线性变换，且对V的基 $x_1,x_2,\cdots,x_n$

有 $T\left(x_1,x_2,\cdots,x_n\right)=\left(x_1,x_2,\cdots,x_n\right)A$

则V的线性变换 $f(T)$ 在基 $x_1,x_2,\cdots,x_n$ 下的矩阵是

$$f(A)=a_m A^m + a_{m-1}A^{m-1}+\cdots+a_1 A+a_0 I$$

定理：设线性变换 T 在基 $x_1, x_2, \cdots, x_n$ 下的矩阵为A,

$x \in V$ 在基 $x_1, x_2, \cdots, x_n$ 下的坐标为 $(\xi_1, \xi_2, \cdots, \xi_n)^T$,

$T(x)$ 在基 $x_1, x_2, \cdots, x_n$ 下的坐标为 $(\eta_1, \eta_2, \cdots, \eta_n)^T$,

则有

$$\begin{pmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_n \end{pmatrix} = A \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{pmatrix}.$$

证：由已知有

$$x=\xi_1 x_1+\xi_2 x_2+\cdots+\xi_n x_n=(x_1,x_2,\cdots,x_n)\begin{pmatrix}\xi_1\\ \xi_2\\ \vdots\\ \xi_n\end{pmatrix}$$

$$T(x)=T(x_1,x_2,\cdots,x_n)\begin{pmatrix}\xi_1\\ \xi_2\\ \vdots\\ \xi_n\end{pmatrix}$$

又 $T(x)=(x_1,x_2,\cdots,x_n)(\eta_1,\eta_2,\cdots,\eta_n)^T$

所以 $$\begin{pmatrix}\eta_1\\ \eta_2\\ \vdots\\ \eta_n\end{pmatrix}=A\begin{pmatrix}\xi_1\\ \xi_2\\ \vdots\\ \xi_n\end{pmatrix}$$

定理： 设线性空间V的线性变换T在两组基

$$x_1, x_2, \cdots, x_n \qquad (\text{I})$$

$$y_1, y_2, \cdots, y_n \qquad (\text{II})$$

下的矩阵分别为A、B，且从基(I)到基(II)的过渡矩阵矩阵是C，则

$$B = C^{-1}AC$$

证：由已知，有

$$T(x_1,x_2,\cdots,x_n)=(x_1,x_2,\cdots,x_n)A,$$

$$T(y_1,y_2,\cdots,y_n)=(y_1,y_2,\cdots,y_n)B,$$

$$(y_1,y_2,\cdots,y_n)=(x_1,x_2,\cdots,x_n)C$$

于是，$T(y_1,y_2,\cdots,y_n)=T(x_1,x_2,\cdots,x_n)C$

$$=(x_1,x_2,\cdots,x_n)AC \;=(y_1,y_2,\cdots,y_n)C^{-1}AC$$

由此即得 $\quad B=C^{-1}AC$

设A、B为数域K上的两个n阶矩阵，若存在可逆矩阵 $P \in K^{n\times n}$，使得 $B = P^{-1}AP$

则称矩阵A相似于B，记为 $A \sim B.$

相似是一个等价关系，即满足如下三条性质：

① 反身性：$A \sim A.$ $\quad$ （ $\because \ A = E^{-1}AE.$ ）

② 对称性：$A \sim B \Rightarrow B \sim A.$

（ $\because \ B = P^{-1}AP \Rightarrow A = Q^{-1}BQ,\ Q = P^{-1}.$ ）

③ 传递性：$A \sim B,\ B \sim C \Rightarrow A \sim C.$

（ $\because \ B = P^{-1}AP,\ C = Q^{-1}BQ$

$\Rightarrow C = Q^{-1}BQ = Q^{-1}(P^{-1}AP)Q = (PQ)^{-1}A(PQ)$ ）

定理： 线性变换在不同基下的矩阵是相似的；

反过来，如果两个矩阵相似，那么它们可以看作

同一线性变换在两组基下所对应的矩阵.

证：前一部分显然成立. 下证后一部分.

设 $A \sim B$, 且A是线性变换 T 在基 $x_1, x_2, \cdots, x_n$ 下的矩阵.

$\because B = C^{-1}AC$， 令 $(y_1, y_2, \cdots, y_n) = (x_1, x_2, \cdots, x_n)C$

显然，$y_1, y_2, \cdots, y_n$ 也是一组基，且 T 在这组基下的

矩阵就是B.

相似矩阵的运算性质

① 若 $B_1 = C^{-1}A_1C$，$B_2 = C^{-1}A_2C$， 则

$$B_1 + B_2 = C^{-1}(A_1 + A_2)C,$$

$$B_1B_2 = C^{-1}(A_1A_2)C.$$

即， $A_1 + A_2 \sim B_1 + B_2$，$A_1A_2 \sim B_1B_2$.

② 若 $B = C^{-1}AC$，$f(x) \in P[x]$，则

$$f(B) = C^{-1}f(A)C.$$

特别地， $B^m = C^{-1}A^mC.$

例：设 x_1, x_2为线性空间V一组基， 线性变换T 在这组基下的矩阵为 $A=\begin{pmatrix} 2 & 1 \\ -1 & 0 \end{pmatrix}$,

y_1, y_2 为V的另一组基，且

$$(y_1, y_2)=(x_1, x_2)\begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix},$$

（1）求T在y_1, y_2下的矩阵B.

（2）求 A^k.

解：（1）T在基 y_1, y_2 下的矩阵

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

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

（2）由 $B=C^{-1}AC$， 有 $A=CBC^{-1}$，

于是 $A^k=CB^kC^{-1}$.

$$\therefore\quad A^k=\begin{pmatrix}1&-1\\-1&2\end{pmatrix}\begin{pmatrix}1&1\\0&1\end{pmatrix}^k\begin{pmatrix}1&-1\\-1&2\end{pmatrix}^{-1}$$

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

例：在线性空间P^3中，线性变换 T定义如下：

$$\begin{cases} T(y_1) = (-5, 0, 3) \\ T(y_2) = (0, -1, 6) \\ T(y_3) = (-5, -1, 9) \end{cases}$$

其中，$\begin{cases} y_1 = (-1, 0, 2) \\ y_2 = (0, 1, 1) \\ y_3 = (3, -1, 0) \end{cases}$

（1）求T在标准基 x_1, x_2, x_3 下的矩阵.

（2）求T 在 y_1, y_2, y_3 下的矩阵.

解：（1）由已知，有

$$(y_1,y_2,y_3)=(x_1,x_2,x_3)\begin{pmatrix}-1 & 0 & 3\\ 0 & 1 & -1\\ 2 & 1 & 0\end{pmatrix}\triangleq(x_1,x_2,x_3)C,$$

$$T(y_1,y_2,y_3)=(x_1,x_2,x_3)\begin{pmatrix}-5 & 0 & -5\\ 0 & -1 & -1\\ 3 & 6 & 9\end{pmatrix},$$

设T 在标准基 x_1,x_2,x_3下的矩阵为A，即

$$T(x_1,x_2,x_3)=(x_1,x_2,x_3)A$$

$$\therefore T(y_1,y_2,y_3)=T((x_1,x_2,x_3)C)=T(x_1,x_2,x_3)C$$

$$=(x_1,x_2,x_3)AC$$

因而， $AC=\begin{pmatrix}-5&0&-5\\0&-1&-1\\3&6&9\end{pmatrix}$,

$$\therefore\quad A=\begin{pmatrix}-5&0&-5\\0&-1&-1\\3&6&9\end{pmatrix}C^{-1}=\begin{pmatrix}-5&0&-5\\0&-1&-1\\3&6&9\end{pmatrix}\begin{pmatrix}-1&0&3\\0&1&-1\\2&1&0\end{pmatrix}^{-1}$$

$$=\frac{1}{7}\begin{pmatrix}-5&20&-20\\-4&-5&-2\\27&18&24\end{pmatrix}$$

（2）设 T 在 y_1, y_2, y_3 下的矩阵为B，则A与B相似，且

$$B = C^{-1}AC$$

$$\therefore \quad B = \begin{pmatrix} -1 & 0 & 3 \\ 0 & 1 & -1 \\ 2 & 1 & 0 \end{pmatrix}^{-1} \begin{pmatrix} -5 & 0 & -5 \\ 0 & -1 & -1 \\ 3 & 6 & 9 \end{pmatrix}$$

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

本讲主要内容

- 线性变换在给定基下的矩阵
- 线性变换的矩阵表示
- 象与原象坐标间的关系
- 线性变换在不同基下矩阵之间的关系
- 线性变换的特征值与特征向量

引入

有限维线性空间V中取定一组基后，V的任一线性变换都可以用矩阵来表示．为了研究线性变换性质，希望这个矩阵越简单越好，如对角矩阵．

从本节开始，我们主要讨论，如何选择一组适当的基，使V的某个线性变换在这组基下的矩阵就是一个对角矩阵?

一、特征值与特征向量

定义：设 T 是数域K上线性空间V的一个线性变换，

若对于K中的一个数 λ_0，存在一个V的非零向量 x，

使得

$$T(x)=\lambda_0 x \quad ,$$

则称 λ_0 为 T 的一个特征值，称 x 为 T 的属于特征值

λ_0 的特征向量.

注：① 几何意义：特征向量经线性变换后方向保持相同 $(\lambda_0>0)$ 或相反 $(\lambda_0<0)$. $\lambda_0=0$ 时，$T(x)=0.$

② 若 x 是 T 的属于特征值 λ_0 的特征向量，则 $kx\ \ (k\in K, k\neq 0)$ 也是 T 的属于 λ_0 的特征向量.

$$\left(\ \because\ \ T(kx)=kT(x)=k(\lambda_0 x)=\lambda_0(kx)\ \right)$$

由此知，特征向量不是被特征值所唯一确定的，但是特征值却是被特征向量所唯一确定的，即若 $T(x)=\lambda x$ 且 $T(x)=\mu x$ ，则 $\lambda=\mu$.

二、特征值与特征向量的求法

分析：设 $\dim V = n$，$x_1, x_2, \cdots, x_n$ 是V的一组基，

线性变换 T 在这组基下的矩阵为A.

设 λ_0是 T 的特征值，它的一个特征向量 x在基

$x_1, x_2, \cdots, x_n$下的坐标记为 $\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}$,

则 $T(x)$在基 $x_1, x_2, \cdots, x_n$下的坐标为 $A\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}$,

而$\lambda_0 x$的坐标是 $\lambda_0\begin{pmatrix}\xi_1\\ \vdots\\ \xi_n\end{pmatrix}$, 又 $T(x)=\lambda_0 x$

于是 $A\begin{pmatrix}\xi_1\\ \vdots\\ \xi_n\end{pmatrix}=\lambda_0\begin{pmatrix}\xi_1\\ \vdots\\ \xi_n\end{pmatrix}$, 从而 $(\lambda_0 I-A)\begin{pmatrix}\xi_1\\ \vdots\\ \xi_n\end{pmatrix}=0.$

即 $\begin{pmatrix}\xi_1\\ \vdots\\ \xi_n\end{pmatrix}$ 是线性方程组 $(\lambda_0 I-A)X=0$ 的解,

又 $\because x\neq 0$, $\therefore \begin{pmatrix}\xi_1\\ \vdots\\ \xi_n\end{pmatrix}\neq 0$, $\therefore$ $(\lambda_0 I-A)X=0$ 有非零解.

所以它的系数行列式 $|\lambda_0 I-A|=0.$

以上分析说明：

若λ_0是T的特征值，则 $|\lambda_0 I - A| = 0$.

反之，若$\lambda_0 \in K$满足 $|\lambda_0 I - A| = 0$,

则齐次线性方程组 $(\lambda_0 I - A)X = 0$ 有非零解.

若 $(\xi_1, \xi_2, \cdots, \xi_n)^T$ 是 $(\lambda_0 I - A)X = 0$ 一个非零解，

则向量 $x = \xi_1 x_1 + \cdots + \xi_n x_n$ 就是T的属于λ_0的一个

特征向量.

设 $A \in K^{n\times n}$，λ是一个参数，矩阵 $\lambda I - A$ 称为

A的特征矩阵，它的行列式

$$|\lambda I - A| = \begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\ \cdots & \cdots & & \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{vmatrix}$$

称为A的特征多项式.

（ $|\lambda I - A|$ 是数域K上的一个n次多项式）

注：① 若矩阵A是线性变换 T 关于V的一组基的矩阵，而 λ_0 是 T 的一个特征值，则 λ_0 是特征多项式 $|\lambda I - A|$ 的根，即 $|\lambda_0 I - A| = 0.$

反之，若 λ_0 是A的特征多项式的根，则 λ_0 就是 T 的一个特征值.（所以，特征值也称特征根.）

② 矩阵A的特征多项式的根有时也称为A的特征值，而相应的线性方程组 $(\lambda I - A)X = 0$ 的非零解也就称为A的属于这个特征值的特征向量.

求特征值与特征向量的一般步骤

i) 在V中任取一组基 $x_1, x_2, \cdots, x_n$ 写出 T 在这组基下的矩阵A .

ii) 求A的特征多项式 $|\lambda I - A|$ 在K上的全部根它们就是 T 的全部特征值.

iii) 把所求得的特征值逐个代入方程组

$$(\lambda I - A)X = 0$$

并求出它的一组基础解系.(它们就是属于这个特征值的全部线性无关的特征向量在基 $x_1, x_2, \cdots, x_n$ 下的坐标.)

例: 在线性空间V中，数乘变换K在任意一组基下的矩阵都是数量矩阵kI，它的特征多项式是

$$\left|\lambda I - kI\right| = (\lambda - k)^n.$$

故数乘法变换K的特征值只有数k，且

对 $\forall x \in V \quad (x \neq 0)$，皆有 $K(x) = kx.$

所以，V中任一非零向量皆为数乘变换K的特征向量.

例: 设线性变换 T 在基 x_1, x_2, x_3 下的矩阵是

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

求 T 特征值与特征向量.

解：A的特征多项式

$$|\lambda I - A| = \begin{vmatrix} \lambda - 1 & -2 & -2 \\ -2 & \lambda - 1 & -2 \\ -2 & -2 & \lambda - 1 \end{vmatrix} = (\lambda + 1)^2 (\lambda - 5)$$

故 T 的特征值为： $\lambda_1 = -1$（二重）， $\lambda_2 = 5$

把 $\lambda=-1$ 代入齐次方程组 $(\lambda I-A)X=0$，得

$$\begin{cases}-2x_1-2x_2-2x_3=0\\-2x_1-2x_2-2x_3=0\\-2x_1-2x_2-2x_3=0\end{cases}$$ 即 $x_1+x_2+x_3=0$

它的一个基础解系为：$(1,0,-1)$，$(0,1,-1)$

因此，属于-1的两个线性无关的特征向量为

$$y_1=x_1-x_3,\quad y_2=x_2-x_3$$

而属于-1的全部特征向量为

$$k_1y_1+k_2y_2,\qquad (k_1,k_2\in K\text{不全为零})$$

把 $\lambda=5$ 代入齐次方程组 $(\lambda I-A)X=0$， 得

$$\begin{cases} 4x_1-2x_2-2x_3=0 \\ -2x_1+4x_2-2x_3=0 \\ -2x_1-2x_2+4x_3=0 \end{cases}$$

解得它的一个基础解系为：(1,1,1)

因此，属于5的一个线性无关的特征向量为

$$y_3=x_1+x_2+x_3$$

而属于5的全部特征向量为

$$k_3y_3, \qquad (k_3\in K,\ k_3\neq 0\ )$$

定义： 设T为n维线性空间V的线性变换，λ_0为T的一个特征值，令V_{λ_0}为T的属于λ_0的全部特征向量再添上零向量所成的集合，即 $V_{\lambda_0}=\{x \mid Tx=\lambda_0 x\}$

则V_{λ_0}是V的一个子空间, 称之为T的一个**特征子空间**.

$$\because \quad T(x+y)=T(x)+T(y)=\lambda_0 x+\lambda_0 y=\lambda_0(x+y)$$

$$T(kx)=kT(x)=k(\lambda_0 x)=\lambda_0(kx)$$

$$\therefore \quad x+y\in V_{\lambda_0}, \qquad kx\in V_{\lambda_0}$$

若 T 在n维线性空间V的某组基下的矩阵为A，则

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

即特征子空间 V_{λ_0}的维数等于齐次线性方程组

$$(\lambda_0 I - A)X = 0 \qquad (*)$$

的解空间的维数，且由方程组(*)得到的属于λ_0的全部线性无关的特征向量就是 V_{λ_0}的一组基.

特征多项式的有关性质

1. 设 $A=\left(a_{ij}\right)\in K^{n\times n}$，则A的特征多项式

$$|\lambda I-A|=\begin{vmatrix}\lambda-a_{11} & -a_{12} & \cdots & -a_{1n}\\ -a_{21} & \lambda-a_{22} & \cdots & -a_{2n}\\ \cdots & \cdots & & \\ -a_{n1} & -a_{n2} & \cdots & \lambda-a_{nn}\end{vmatrix}$$

$$=\lambda^{n}-(a_{11}+a_{22}+\cdots+a_{nn})\lambda^{n-1}+\cdots+(-1)^{n}|A|$$

由多项式根与系数的关系还可得

① A的全体特征值的和＝$a_{11}+a_{22}+\cdots+a_{nn}$.

称之为A的迹,记作trA

② A的全体特征值的积＝ $|A|$.

定理：设 $A=\left(a_{ij}\right)_{n\times n}, B=\left(b_{ij}\right)_{n\times n}$ 则 $\mathrm{tr}\left(AB\right)=\mathrm{tr}\left(BA\right)$

证：令 $AB=\left(u_{ij}\right), BA=\left(v_{ij}\right)$，于是有

$$u_{ij}=\sum_{k=1}^{n}a_{ik}b_{kj},\quad v_{ij}=\sum_{l=1}^{n}b_{il}a_{lj}$$

$$\therefore\quad \mathrm{tr}(AB)=\sum_{i=1}^{n}u_{ii}=\sum_{i=1}^{n}\left[\sum_{k=1}^{n}a_{ik}b_{ki}\right]$$

$$=\sum_{k=1}^{n}\left[\sum_{i=1}^{n}b_{ki}a_{ik}\right]=\sum_{k=1}^{n}v_{kk}=\mathrm{tr}(BA)$$

定理：相似矩阵有相似的迹 $\operatorname{tr}(AB)=\operatorname{tr}(BA)$

证：设 $A\sim B$，即 $\exists P\neq 0,\text{st}\quad B=P^{-1}AP$

$$\begin{aligned}\operatorname{tr}(B)&=\operatorname{tr}(P^{-1}AP)\\&=\operatorname{tr}(APP^{-1})\\&=\operatorname{tr}(A)\end{aligned}$$

定理： 相似矩阵具有相同的特征多项式.

证:设 $A \sim B$, 则存在可逆矩阵P，使得

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

于是，
$$\begin{aligned}|\lambda I - B| &= |\lambda I - P^{-1}AP| \\ &= |\lambda P^{-1}IP - P^{-1}AP| \\ &= |P^{-1}(\lambda I - A)P| \\ &= |P^{-1}||\lambda I - A||P| \\ &= |\lambda I - A|\end{aligned}$$

注：① 由定理线性变换T的特征值与基的选择无关.

因此,矩阵A的特征多项式也说成是线性变换 T 的特征多项式；而线性变换 T 的特征值与特征向量有时也说成是矩阵A的特征值与特征向量.

② 有相同特征多项式的矩阵未必相似.

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

它们的特征多项式都是 $(\lambda-1)^2$; 但A、B不相似.

定理：任意n阶矩阵A与三角矩阵相似

证明:对阶数n利用数学归纳法证明. 当n=1时显然成立

设当阶数为n−1时定理成立。

设 $x_1, x_2, \cdots, x_n$ 是n个线性无关的列向量,

其中x_1 为A的特征值λ的特征向量, $Ax_1 = \lambda_1 x_1$

记 $P_1 = (x_1, x_2, \cdots, x_n)$

于是 $AP_1 = (Ax_1, Ax_2, \cdots, Ax_n) = (\lambda x_1, Ax_2, \cdots, Ax_n)$

由于$Ax_i \in C^n$,因此可以由 $x_1, x_2, \cdots, x_n$唯一地线性表示

即有 $Ax_i = b_{1i}x_1 + b_{2i}x_2 + \cdots + b_{ni}x_n$

于是 $AP_1 = (\lambda x_1, Ax_2, \cdots, Ax_n)$

$$= (x_1, x_2, \cdots, x_n)\begin{bmatrix} \lambda & b_{12} & \cdots & b_{1n} \\ 0 & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & b_{nn} \end{bmatrix}$$

即

$$P_1^{-1}AP_1 = \begin{bmatrix} \lambda & b_{12} & \cdots & b_{1n} \\ 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{bmatrix}$$

由归纳假定，对于n–1阶矩阵A_1存在矩阵Q使得

$$Q^{-1}A_1Q = \begin{bmatrix} \lambda_2 & & * \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

记 $P_2 = \begin{bmatrix} 1 & 0^T \\ 0 & Q \end{bmatrix}$, $P = P_1 P_2$

则有 $P^{-1}AP = \left(P_1 P_2\right)^{-1} A \left(P_1 P_2\right) = P_2^{-1}\left(P_1^{-1} A P_1\right) P_2$

$$= P_2^{-1} \begin{bmatrix} \lambda & b_{12} & \cdots & b_{1n} \\ 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{bmatrix} P_2$$

$$= \begin{bmatrix} \lambda_1 & * & \cdots & * \\ & \lambda_2 & \ddots & \vdots \\ & & \ddots & * \\ & & & \lambda_n \end{bmatrix}$$

Hamilton—Caylay定理 设 $A \in K^{n\times n}$ 其特征多项式

$$\varphi(\lambda) = \left|\lambda I - A\right| = \lambda^n + a_1\lambda^{n-1} + \cdots + a_{n-1}\lambda + a_n$$

则 $\varphi(A) = A^n + a_1 A^{n-1} + \cdots + a_{n-1}A + a_n I = 0.$

证明:A的特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_n$ ，即

$$\varphi(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_n)$$

必然存在可逆矩阵 $P_{n\times n}$ ，使得

$$P^{-1}AP = \begin{bmatrix} \lambda_1 & * & \cdots & * \\ & \lambda_2 & \ddots & \vdots \\ & & \ddots & * \\ & & & \lambda_n \end{bmatrix}$$

$$\varphi(P^{-1}AP)=(P^{-1}AP-\lambda_1 I)(P^{-1}AP-\lambda_2 I)\cdots(P^{-1}AP-\lambda_n I)$$

$$=\begin{bmatrix} 0 & * & \cdots & * \\ & \lambda_2-\lambda_1 & \ddots & \vdots \\ & & \ddots & * \\ & & & \lambda_n-\lambda_1 \end{bmatrix}\begin{bmatrix} \lambda_1-\lambda_2 & * & \cdots & * \\ & 0 & \ddots & \vdots \\ & & \ddots & * \\ & & & \lambda_n-\lambda_2 \end{bmatrix}\cdots(P^{-1}AP-\lambda_n I)$$

$$=\begin{bmatrix} 0 & 0 & * & \cdots & * \\ 0 & 0 & * & \cdots & * \\ 0 & 0 & * & \ddots & * \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & & & * \end{bmatrix}\begin{bmatrix} \lambda_1-\lambda_3 & * & * & \cdots & * \\ & \lambda_2-\lambda_3 & * & \cdots & * \\ & & 0 & \ddots & \vdots \\ & & & \ddots & * \\ & & & & \lambda_2-\lambda_3 \end{bmatrix}\cdots(P^{-1}AP-\lambda_n I)$$

$$=0$$

即 $P^{-1}\varphi(A)P=0 \quad \Rightarrow \varphi(A)=0$

例: 设 $A=\begin{pmatrix} 1 & 0 & 2 \\ 0 & -1 & 1 \\ 0 & 1 & 0 \end{pmatrix}$,求 $2A^8-3A^5+A^4+A^2-4I$

解：A的特征多项式 $\varphi(\lambda)=|\lambda I-A|=\lambda^3-2\lambda+1$

用$\varphi(\lambda)$去除 $2\lambda^8-3\lambda^5+\lambda^4+\lambda^2-4\lambda=g(\lambda)$,得

$$g(\lambda)=\varphi(\lambda)(2\lambda^5+4\lambda^3-5\lambda^2+9\lambda-14)+(24\lambda^2-37\lambda+10)$$

$\because \quad \varphi(A)=0,$

$\therefore \quad 2A^8-3A^5+A^4+A^2-4I=24A^2-37A+10I$

$$=\begin{pmatrix} -3 & 48 & -26 \\ 0 & 95 & -61 \\ 0 & -61 & 34 \end{pmatrix}$$

练习1： 已知 $A \in P^{n\times n}$，λ 为A的一个特征值，则

（1） $kA \ (k \in P)$ 必有一个特征值为 $\underline{\quad k\lambda \quad}$；

（2） $A^m \ (m \in Z^+)$ 必有一个特征值为 $\underline{\quad \lambda^m \quad}$；

（3） A可逆时，A^{-1}必有一个特征值为 $\underline{\quad \lambda^{-1} \quad}$；

（4） A可逆时，A^*必有一个特征值为 $\underline{\quad \dfrac{|A|}{\lambda} \quad}$.

（5） $f(x) \in P[x]$，则 $f(A)$必有一个特征值为 $\underline{\ f(\lambda)\ }$.

练习2： 已知3阶方阵A的特征值为：1、－1、2，

则矩阵 $\boldsymbol{B}=\boldsymbol{A}^3-2\boldsymbol{A}^2$ 的特征值为：$\underline{\quad -1,-3,0 \quad}$，

行列式 $|\boldsymbol{B}|=$ $\underline{\quad 0 \quad}$.

作业

- **P77：4，5**
- **P78：7，8**
- **P79：14，16**

本讲主要内容

- 线性变换的特征值与特征向量
- 最小多项式
- 对角矩阵
- 不变子空间

引入

有限维线性空间V中取定一组基后，V的任一线性变换都可以用矩阵来表示. 为了研究线性变换性质，希望这个矩阵越简单越好，如对角矩阵.

从本节开始，我们主要讨论，如何选择一组适当的基，使V的某个线性变换在这组基下的矩阵就是一个对角矩阵?

约束极值与特征值

考虑一个几何约束条件的实对称二次型的极值问题

假定 $x \in R^n, x^T x = 1$ ，求 $x^T Ax$ 的极大值

其中 $A = A^T \in R^{n\times n}$ 。

Lagrauge函数 $L = x^T Ax - \lambda x^T x$

有极值的必要条件 $0 = \nabla L = 2(Ax - \lambda x) = 0$

也就是说，**A**的特征值、特征向量对是极值问题的解

一、特征值与特征向量

定义：设 T 是数域K上线性空间V的一个线性变换，

若对于K中的一个数 λ_0，存在一个V的非零向量 x，

使得

$$T(x) = \lambda_0 x \quad ,$$

则称 λ_0 为 T 的一个特征值，称 x 为 T 的属于特征值

λ_0 的特征向量.

注：① 几何意义：特征向量经线性变换后方向保持相同 $(\lambda_0>0)$ 或相反 $(\lambda_0<0)$. $\lambda_0=0$ 时，$T(x)=0.$

② 若 x 是 T 的属于特征值 λ_0 的特征向量，则 $kx\quad(k\in K, k\neq 0)$ 也是 T 的属于 λ_0 的特征向量.

$$\left(\quad\because\quad T(kx)=kT(x)=k(\lambda_0 x)=\lambda_0(kx)\quad\right)$$

由此知，特征向量不是被特征值所唯一确定的，但是特征值却是被特征向量所唯一确定的，即若 $T(x)=\lambda x$ 且 $T(x)=\mu x$，则 $\lambda=\mu$.

二、特征值与特征向量的求法

分析：设 $\dim V = n$， $x_1, x_2, \cdots, x_n$ 是V的一组基，

线性变换 T 在这组基下的矩阵为A.

设 λ_0是 T 的特征值，它的一个特征向量 x在基

$x_1, x_2, \cdots, x_n$下的坐标记为 $\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}$,

则 $T(x)$在基 $x_1, x_2, \cdots, x_n$下的坐标为 $A\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}$,

而$\lambda_0 x$的坐标是 $\lambda_0\begin{pmatrix}\xi_1\\ \vdots\\ \xi_n\end{pmatrix}$, 又 $T(x)=\lambda_0 x$

于是 $A\begin{pmatrix}\xi_1\\ \vdots\\ \xi_n\end{pmatrix}=\lambda_0\begin{pmatrix}\xi_1\\ \vdots\\ \xi_n\end{pmatrix}$, 从而 $(\lambda_0 I-A)\begin{pmatrix}\xi_1\\ \vdots\\ \xi_n\end{pmatrix}=0.$

即 $\begin{pmatrix}\xi_1\\ \vdots\\ \xi_n\end{pmatrix}$ 是线性方程组 $(\lambda_0 I-A)X=0$ 的解，

又 $\because x\neq 0,\ \therefore \begin{pmatrix}\xi_1\\ \vdots\\ \xi_n\end{pmatrix}\neq 0,\ \therefore\ (\lambda_0 I-A)X=0$ 有非零解.

所以它的系数行列式 $\left|\lambda_0 I-A\right|=0.$

以上分析说明：

若λ_0是T的特征值，则 $|\lambda_0 I - A| = 0.$

反之，若$\lambda_0 \in K$满足 $|\lambda_0 I - A| = 0,$

则齐次线性方程组 $(\lambda_0 I - A)X = 0$ 有非零解.

若 $(\xi_1, \xi_2, \cdots, \xi_n)^T$ 是 $(\lambda_0 I - A)X = 0$ 一个非零解，

则向量 $x = \xi_1 x_1 + \cdots + \xi_n x_n$ 就是T的属于λ_0的一个

特征向量.

设 $A \in K^{n\times n}$，λ是一个参数，矩阵 $\lambda I - A$ 称为

A的特征矩阵，它的行列式

$$|\lambda I - A| = \begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\ \cdots & \cdots & & \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{vmatrix}$$

称为A的特征多项式.

（ $|\lambda I - A|$ 是数域K上的一个n次多项式）

注：① 若矩阵A是线性变换 T 关于V的一组基的矩阵，而 λ_0 是 T 的一个特征值，则 λ_0 是特征多项式 $|\lambda I - A|$ 的根，即 $|\lambda_0 I - A| = 0.$

反之，若 λ_0 是A的特征多项式的根，则 λ_0 就是 T 的一个特征值.（所以，特征值也称特征根.）

② 矩阵A的特征多项式的根有时也称为A的特征值，而相应的线性方程组 $(\lambda I - A)X = 0$ 的非零解也就称为A的属于这个特征值的特征向量.

求特征值与特征向量的一般步骤

i) 在V中任取一组基 $x_1, x_2, \cdots, x_n$ 写出 T 在这组基下的矩阵A .

ii) 求A的特征多项式 $|\lambda I - A|$ 在K上的全部根它们就是 T 的全部特征值.

iii) 把所求得的特征值逐个代入方程组

$$(\lambda I - A)X = 0$$

并求出它的一组基础解系.(它们就是属于这个特征值的全部线性无关的特征向量在基 $x_1, x_2, \cdots, x_n$ 下的坐标.)

例:在线性空间V中，数乘变换K在任意一组基下的矩阵都是数量矩阵kI，它的特征多项式是

$$\left|\lambda I-kI\right|=(\lambda-k)^{n}.$$

故数乘法变换K的特征值只有数k，且

对 $\forall x\in V\quad(x\neq 0)$，皆有 $K(x)=kx.$

所以，V中任一非零向量皆为数乘变换K的特征向量.

例: 设线性变换 T 在基 x_1, x_2, x_3 下的矩阵是

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

求 T 特征值与特征向量.

解：A的特征多项式

$$|\lambda I-A|=\begin{vmatrix}\lambda-1&-2&-2\\-2&\lambda-1&-2\\-2&-2&\lambda-1\end{vmatrix}=(\lambda+1)^2(\lambda-5)$$

故 T 的特征值为： $\lambda_1=-1$（二重），$\lambda_2=5$

把 $\lambda=-1$ 代入齐次方程组 $(\lambda I-A)X=0$，得

$$\begin{cases}-2x_1-2x_2-2x_3=0\\-2x_1-2x_2-2x_3=0\\-2x_1-2x_2-2x_3=0\end{cases}$$ 即 $x_1+x_2+x_3=0$

它的一个基础解系为：$(1,0,-1)$，$(0,1,-1)$

因此，属于-1的两个线性无关的特征向量为

$y_1=x_1-x_3$， $y_2=x_2-x_3$

而属于-1的全部特征向量为

$k_1y_1+k_2y_2$， $(k_1,k_2\in K$ 不全为零 $)$

把 $\lambda=5$ 代入齐次方程组 $(\lambda I-A)X=0$，得

$$\begin{cases} 4x_1-2x_2-2x_3=0 \\ -2x_1+4x_2-2x_3=0 \\ -2x_1-2x_2+4x_3=0 \end{cases}$$

解得它的一个基础解系为：(1,1,1)

因此，属于5的一个线性无关的特征向量为

$$y_3=x_1+x_2+x_3$$

而属于5的全部特征向量为

$$k_3y_3, \quad (k_3\in K, \ k_3\neq 0 \ )$$

定义：设T 为n维线性空间V的线性变换，λ_0为 T 的一个特征值，令 V_{λ_0}为 T 的属于λ_0的全部特征向量再添上零向量所成的集合，即 $V_{\lambda_0}=\{x \mid Tx=\lambda_0 x\}$

则V_{λ_0}是V的一个子空间, 称之为 T 的一个特征子空间.

$$\because \quad T(x+y)=T(x)+T(y)=\lambda_0 x+\lambda_0 y=\lambda_0(x+y)$$

$$T(kx)=kT(x)=k(\lambda_0 x)=\lambda_0(kx)$$

$$\therefore \quad x+y\in V_{\lambda_0}, \qquad kx\in V_{\lambda_0}$$

若 T 在n维线性空间V的某组基下的矩阵为A，则

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

即特征子空间 V_{λ_0}的维数等于齐次线性方程组

$$(\lambda_0 I - A)X = 0 \qquad (*)$$

的解空间的维数，且由方程组(*)得到的属于λ_0的全部线性无关的特征向量就是 V_{λ_0}的一组基.

特征多项式的有关性质

1. 设 $A=\left(a_{ij}\right)\in K^{n\times n}$，则A的特征多项式

$$|\lambda I-A|=\begin{vmatrix}\lambda-a_{11} & -a_{12} & \cdots & -a_{1n}\\ -a_{21} & \lambda-a_{22} & \cdots & -a_{2n}\\ \cdots & \cdots & & \\ -a_{n1} & -a_{n2} & \cdots & \lambda-a_{nn}\end{vmatrix}$$

$$=\lambda^{n}-(a_{11}+a_{22}+\cdots+a_{nn})\lambda^{n-1}+\cdots+(-1)^{n}|A|$$

由多项式根与系数的关系还可得

① A的全体特征值的和＝$a_{11}+a_{22}+\cdots+a_{nn}$.

称之为A的迹,记作trA

② A的全体特征值的积＝$|A|$.

定理：设 $A=(a_{ij})_{n\times n}, B=(b_{ij})_{n\times n}$ 则 $\mathrm{tr}(AB)=\mathrm{tr}(BA)$

证：令 $AB=(u_{ij}), BA=(v_{ij})$，于是有

$$u_{ij}=\sum_{k=1}^{n} a_{ik}b_{kj}, \quad v_{ij}=\sum_{l=1}^{n} b_{il}a_{lj}$$

$$\therefore \quad \mathrm{tr}(AB)=\sum_{i=1}^{n} u_{ii}=\sum_{i=1}^{n}\left[\sum_{k=1}^{n} a_{ik}b_{ki}\right]$$

$$=\sum_{k=1}^{n}\left[\sum_{i=1}^{n} b_{ki}a_{ik}\right]=\sum_{k=1}^{n} v_{kk}=\mathrm{tr}(BA)$$

定理：相似矩阵有相似的迹　$\mathrm{tr}(AB)=\mathrm{tr}(BA)$

证：设 $A\sim B$，即 $\exists P\neq 0,\mathrm{st}\quad B=P^{-1}AP$

$$\begin{aligned}\mathrm{tr}(B)&=\mathrm{tr}(P^{-1}AP)\\&=\mathrm{tr}(APP^{-1})\\&=\mathrm{tr}(A)\end{aligned}$$

定理： 相似矩阵具有相同的特征多项式.

证：设 $A \sim B$，则存在可逆矩阵 P，使得

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

于是，$|\lambda I - B| = |\lambda I - P^{-1}AP|$

$$= |\lambda P^{-1}IP - P^{-1}AP|$$

$$= |P^{-1}(\lambda I - A)P|$$

$$= |P^{-1}||\lambda I - A||P|$$

$$= |\lambda I - A|$$

注：① 由定理线性变换T的特征值与基的选择无关.

因此,矩阵A的特征多项式也说成是线性变换 T 的特征多项式；而线性变换 T 的特征值与特征向量有时也说成是矩阵A的特征值与特征向量.

② 有相同特征多项式的矩阵未必相似.

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

它们的特征多项式都是 $(\lambda-1)^2$; 但A、B不相似.

定理：任意n阶矩阵A与三角矩阵相似

证明:对阶数n利用数学归纳法证明. 当n=1时显然成立

设当阶数为n–1时定理成立。

设 $x_1, x_2, \cdots, x_n$ 是n个线性无关的列向量，

其中 x_1 为A的特征值λ的特征向量，$Ax_1 = \lambda_1 x_1$

记 $P_1 = \left(x_1, x_2, \cdots, x_n\right)$

于是 $AP_1 = \left(Ax_1, Ax_2, \cdots, Ax_n\right) = \left(\lambda x_1, Ax_2, \cdots, Ax_n\right)$

由于 $Ax_i \in C^n$,因此可以由 $x_1, x_2, \cdots, x_n$ 唯一地线性表示

即有 $Ax_i = b_{1i}x_1 + b_{2i}x_2 + \cdots + b_{ni}x_n$

于是 $AP_1 = (\lambda x_1, Ax_2, \cdots, Ax_n)$

$$= (x_1, x_2, \cdots, x_n)\begin{bmatrix} \lambda & b_{12} & \cdots & b_{1n} \\ 0 & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & b_{nn} \end{bmatrix}$$

即

$$P_1^{-1}AP_1 = \begin{bmatrix} \lambda & b_{12} & \cdots & b_{1n} \\ 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{bmatrix}$$

由归纳假定，对于n–1阶矩阵A_1存在矩阵Q使得

$$Q^{-1}A_1Q = \begin{bmatrix} \lambda_2 & & * \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

记 $P_2=\begin{bmatrix}1 & 0^T\\ 0 & Q\end{bmatrix}$, $P=P_1P_2$

则有 $P^{-1}AP=\left(P_1P_2\right)^{-1}A\left(P_1P_2\right)=P_2^{-1}\left(P_1^{-1}AP_1\right)P_2$

$$=P_2^{-1}\begin{bmatrix}\lambda & b_{12} & \cdots & b_{1n}\\ 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{bmatrix}P_2$$

$$=\begin{bmatrix}\lambda_1 & * & \cdots & *\\ & \lambda_2 & \ddots & \vdots\\ & & \ddots & *\\ & & & \lambda_n\end{bmatrix}$$

#

Hamilton—Caylay定理 设 $A \in K^{n\times n}$ 其特征多项式

$$\varphi(\lambda)=|\lambda I - A| = \lambda^n + a_1\lambda^{n-1} + \cdots + a_{n-1}\lambda + a_n$$

则 $\varphi(A) = A^n + a_1 A^{n-1} + \cdots + a_{n-1}A + a_n I = 0.$

证明:A的特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_n$ ，即

$$\varphi(\lambda)=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n)$$

必然存在可逆矩阵 $P_{n\times n}$ ，使得

$$P^{-1}AP=\begin{bmatrix} \lambda_1 & * & \cdots & * \\ & \lambda_2 & \ddots & \vdots \\ & & \ddots & * \\ & & & \lambda_n \end{bmatrix}$$

$$\varphi(P^{-1}AP)=(P^{-1}AP-\lambda_1 I)(P^{-1}AP-\lambda_2 I)\cdots(P^{-1}AP-\lambda_n I)$$

$$=\begin{bmatrix} 0 & * & \cdots & * \\ & \lambda_2-\lambda_1 & \ddots & \vdots \\ & & \ddots & * \\ & & & \lambda_n-\lambda_1 \end{bmatrix}\begin{bmatrix} \lambda_1-\lambda_2 & * & \cdots & * \\ & 0 & \ddots & \vdots \\ & & \ddots & * \\ & & & \lambda_n-\lambda_2 \end{bmatrix}\cdots(P^{-1}AP-\lambda_n I)$$

$$=\begin{bmatrix} 0 & 0 & * & \cdots & * \\ 0 & 0 & * & \cdots & * \\ 0 & 0 & * & \ddots & * \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & & & * \end{bmatrix}\begin{bmatrix} \lambda_1-\lambda_3 & * & * & \cdots & * \\ & \lambda_2-\lambda_3 & * & \cdots & * \\ & & 0 & \ddots & \vdots \\ & & & \ddots & * \\ & & & & \lambda_2-\lambda_3 \end{bmatrix}\cdots(P^{-1}AP-\lambda_n I)$$

$$=0$$

即 $P^{-1}\varphi(A)P=0 \quad \Rightarrow \varphi(A)=0$

例: 设 $A=\begin{pmatrix}1&0&2\\0&-1&1\\0&1&0\end{pmatrix}$,求 $2A^8-3A^5+A^4+A^2-4I$

解：A的特征多项式 $\varphi(\lambda)=|\lambda I-A|=\lambda^3-2\lambda+1$

用$\varphi(\lambda)$去除 $2\lambda^8-3\lambda^5+\lambda^4+\lambda^2-4\lambda=g(\lambda)$,得

$$g(\lambda)=\varphi(\lambda)(2\lambda^5+4\lambda^3-5\lambda^2+9\lambda-14)+(24\lambda^2-37\lambda+10)$$

$\because\quad \varphi(A)=0,$

$\therefore\quad 2A^8-3A^5+A^4+A^2-4I=24A^2-37A+10I$

$$=\begin{pmatrix}-3&48&-26\\0&95&-61\\0&-61&34\end{pmatrix}$$

练习1： 已知 $A \in P^{n\times n}$，λ 为A的一个特征值，则

（1） $kA \ (k \in P)$ 必有一个特征值为 $k\lambda$;

（2） $A^m \ (m \in Z^+)$ 必有一个特征值为 λ^m ;

（3） A可逆时, A^{-1}必有一个特征值为 λ^{-1} ;

（4） A可逆时， A^*必有一个特征值为 $\dfrac{|A|}{\lambda}$.

（5） $f(x) \in P[x]$, 则 $f(A)$必有一个特征值为 $f(\lambda)$.

练习2： 已知3阶方阵A的特征值为：1、－1、2，

则矩阵 $\boldsymbol{B}=\boldsymbol{A}^3-2\boldsymbol{A}^2$ 的特征值为：___$-1,-3,0$___，

行列式 $|\boldsymbol{B}|=$ ___0___.

本讲主要内容

- 线性变换的特征值与特征向量
- 最小多项式
- 对角矩阵
- 不变子空间

由哈密尔顿—凯莱定理，$\forall A\in K^{n\times n},\varphi(\lambda)=|\lambda I-A|$ 是A的特征多项式，则 $\varphi(A)=0.$

因此，对任定一个矩阵 $A\in K^{n\times n}$，总可以找到一个多项式 $\varphi(x)\in P[x]$，使 $\varphi(A)=0.$ 此时，也称多项式$\varphi(x)$以A为根.

本节讨论，以矩阵A为根的多项式的中次数最低的那个与A的对角化之间的关系.

定义：设 $A \in K^{n\times n}$，在数域K上的以A为根的多项式中，次数最低的首项系数为1的那个多项式，称为**A的最小多项式**.常记做 $m(\lambda)$

显然 $m(\lambda)$ 的次数不大于特征多项式 $\varphi(\lambda)$ 的次数

最小多项式的基本性质

定理：矩阵A的最小多项式$m(\lambda)$可以整除以A为根的任意首1多项式 $f(\lambda)$，且 $m(\lambda)$是唯一的.

证：（1）反证法。

假如 $m(\lambda)$ 不能整除 $f(\lambda)$，则有

$$f(\lambda)=m(\lambda)q(\lambda)+r(\lambda)$$

其中 $r(\lambda)$ 的次数小于 $m(\lambda)$ 的次数。于是

$$f(A)=m(A)q(A)+r(A)$$

由 $f(A)=0$ 和 $m(A)=0$ 可得 $r(A)=0$

这与 $m(\lambda)$是A的最小多项式相矛盾

设 $m_1(\lambda)$和$m_2(\lambda)$ 都是A的最小多项式，则

$$m_2(A)=0 \quad \Rightarrow m_1(\lambda)\big|m_2(\lambda)$$

$$m_1(A)=0 \quad \Rightarrow m_2(\lambda)\big|m_1(\lambda)$$

又 $m_1(x), m_2(x)$都是首1多项式,

故 $m_1(\lambda)=m_2(\lambda)$.

定理：矩阵A的最小多项式$m(\lambda)$与其特征多项式 $\varphi(\lambda)$ 的零点相同（不计重数）。

证：显然 $\varphi(A)=0$，且 $m(\lambda)\mid\varphi(\lambda)$

所以 $m(\lambda)$ 的零点是 $\varphi(\lambda)$ 的零点

设 λ_0 是 $\varphi(\lambda)$ 的零点

$$Ax=\lambda_0 x(x\neq 0)\Rightarrow m(A)x=m(\lambda_0)x=0$$

所以 λ_0 也是 $m(\lambda)$ 的零点

推论：$m(\lambda)$ 一定含 $\varphi(\lambda)$ 的全部单因式

但：$m(\lambda)$不一定是$\varphi(\lambda)$的全部单因式的乘积

例：数量矩阵$k\mathbf{I}$的最小多项式是一次多项式 $x-k$;

特别地，单位矩阵的最小多项式是 $\lambda-1$ ；

零矩阵的最小多项式是 λ .

反之，若矩阵A的最小多项式是一次多项式，则

A一定是数量矩阵.

例2、求 $A=\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ 的最小多项式.

解：A的特征多项式为

$$\varphi(\lambda)=|\lambda I - A|=\begin{vmatrix} \lambda-1 & -1 & 0 \\ 0 & \lambda-1 & 0 \\ 0 & 0 & \lambda-1 \end{vmatrix}=(\lambda-1)^3$$

又 $A - I \neq 0,$

$$(A-I)^2 = A^2 - 2A + I$$

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

$\therefore$ A的最小多项式为 $(\lambda-1)^2.$

定理：相似矩阵具有相同的最小多项式.

证：设矩阵A与B相似，$m_A(\lambda), m_B(\lambda)$分别为它们的最小多项式.

由A相似于B，存在可逆矩阵P， 使 $B = P^{-1}AP.$

从而 $m_A(B) = m_A(P^{-1}AP) = P^{-1}m_A(A)P = 0$

$\therefore m_A(\lambda)$也以B为根， 从而 $m_B(\lambda) \mid m_A(\lambda).$

同理可得 $m_A(\lambda) \mid m_B(\lambda).$

又$m_A(\lambda), m_B(\lambda)$ 都是首1多项式，$\therefore m_A(\lambda) = m_B(\lambda).$

注：反之不然，即最小多项式相同的矩阵未必相似.

如：

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

的最小多项式皆为$(x-1)^2(x-2)$，但A与B不相似.

$$\because\quad |\lambda I-A|=(\lambda-1)^3(\lambda-2),$$

$$|\lambda I-B|=(\lambda-1)^2(\lambda-2)^2$$

即 $|\lambda I-A|\neq|\lambda I-B|$.　　所以，A与B不相似.

最小多项式求法

定理：设n阶矩阵A特征多项式φ (λ)，特征矩阵的λI–A的全体n–1阶子式的最大公因式为d(λ)，则A最小多项式为

$$m(\lambda)=\frac{\det(\lambda I-A)}{d(\lambda)}$$

设 $A=\begin{pmatrix} 3 & -3 & 2 \\ -1 & 5 & -2 \\ -1 & 3 & 0 \end{pmatrix}$，求 $m(\lambda)$

解：$\lambda I - A = \begin{pmatrix} \lambda-3 & -3 & 2 \\ -1 & \lambda-5 & -2 \\ -1 & 3 & \lambda \end{pmatrix}$

$$\varphi(\lambda) = |\lambda I - A| = (\lambda-2)^2(\lambda-4)$$

$M_{11} = \lambda^2 - 5\lambda + 6$ $\quad M_{21} = 3\lambda - 6$ $\quad M_{31} = 2\lambda - 4$

$M_{12} = \lambda - 2$ $\quad M_{22} = \lambda^2 - 3\lambda + 2$ $\quad M_{32} = 2\lambda - 4$

$M_{13} = -\lambda + 2$ $\quad M_{23} = -3(\lambda - 2)$ $\quad M_{33} = (\lambda - 2)(\lambda - 6)$

$d(\lambda) = \lambda - 2$

$$\therefore m(\lambda) = \frac{\varphi(\lambda)}{d(\lambda)} = (\lambda-2)(\lambda-4)$$

本讲主要内容

- 线性变换的特征值与特征向量
- 最小多项式
- 对角矩阵
- 不变子空间

一、可对角化的概念

定义1：设T是n维线性空间V的一个线性变换，如果存在V的一个基，使T在这组基下的矩阵为对角矩阵，则称线性变换 T 可对角化.

定义2：矩阵A是数域K上的一个n阶方阵. 如果存在一个K上的n阶可逆矩阵P，使 $P^{-1}AP$ 为对角矩阵，则称矩阵A可对角化.

二、可对角化的条件

定理： 设T为n维线性空间V的一个线性变换，

则T可对角化 $\Leftrightarrow$ T有n个线性无关的特征向量.

证：设T在基$x_1, x_2, \cdots x_n$下的矩阵为对角矩阵

$$diag(\lambda_1, \lambda_2, \cdots, \lambda_n)$$

则有$Tx_i = \lambda_i x_i, \ \ i = 1, 2, \cdots n.$

$\therefore x_1, x_2, \cdots x_n$就是$T$ 的n个线性无关的特征向量.

反之，若T有n个线性无关的特征向量$y_1, y_2, \cdots, y_n$，那么就取$y_1, y_2, \cdots, y_n$为基，则在这组基下T的矩阵是对角矩阵.

定理： 设 T 为n维线性空间V的一个线性变换,

如果$x_1, x_2, \cdots x_k$分别是T 的属于互不相同的特征值

$\lambda_1, \lambda_2, \cdots \lambda_k$ 的特征向量，则$x_1, x_2, \cdots x_k$线性无关.

证：对k作数学归纳法.

当 $k=1$ 时, $\because x_1 \neq 0$, $\therefore x_1$线性无关. 命题成立.

假设对于$k-1$来说，结论成立. 现设 $\lambda_1, \lambda_2, \cdots \lambda_k$ 为 T 的互不相同的特征值，x_i 是属于 λ_i 的特征向量，

即 $Tx_i = \lambda_i x_i, \quad i = 1, 2, \cdots, n.$

设 $a_1 x_1 + a_2 x_2 + \cdots + a_k x_k = 0, \quad a_i \in K$ ①

以 λ_k 乘①式的两端，得

$$a_1\lambda_k x_1 + a_2\lambda_k x_2 + \cdots a_k\lambda_k x_k = 0. \qquad ②$$

又对①式两端施行线性变换 T，得

$$a_1\lambda_1 x_1 + a_2\lambda_2 x_2 + \cdots a_k\lambda_k x_k = 0. \qquad ③$$

③式减②式得

$$a_1(\lambda_1-\lambda_k)x_1 + a_2(\lambda_2-\lambda_k)x_2 + \cdots a_{k-1}(\lambda_{k-1}-\lambda_k)x_{k-1} = 0$$

由归纳假设, $x_1, x_2, \cdots, x_{k-1}$ 线性无关，所以

$$a_i(\lambda_i-\lambda_k) = 0, \quad i = 1,2,\cdots,k-1.$$

但 $\lambda_1,\lambda_2,\cdots,\lambda_k$ 互不相同，所以 $a_1 = a_2 = \cdots = a_{k-1} = 0.$

将之代入①，得 $a_k x_k = 0.$

$\because \quad x_k \neq 0, \qquad \therefore \quad a_k = 0$

故 $x_1, x_2, \cdots, x_k$ 线性无关.

定理25： n阶矩阵与对角矩阵相似的充要条件：A有n个线性无关的特征向量，或A有完备的特征向量系

对角化的一般方法

设 T 为维线性空间V的一个线性变换，$x_1, x_2, \cdots, x_n$ 为V的一组基， T 在这组基下的矩阵为A.

步骤:

1° 求出矩阵A的全部特征值 $\lambda_1, \lambda_2, \cdots, \lambda_k$.

2° 对每一个特征值 λ_i ，求出齐次线性方程组

$$\left(\lambda_i I - A\right) X = 0, \qquad i = 1.2.\cdots k$$

的一个基础解系（此即 T 的属于 λ_i 的全部线性无关的特征向量在基 $x_1, x_2, \cdots, x_n$下的坐标）.

3° 若全部基础解系所合向量个数之和等于n，则

T 有n个线性无关的特征向量 $y_1, y_2, \cdots, y_n$, 从而 T

（或矩阵A）可对角化. 以这些解向量为列，作一个

n阶方阵C，则C可逆，$C^{-1}AC$ 是对角矩阵. 而且

C就是基 $x_1, x_2, \cdots, x_n$到基 $y_1, y_2, \cdots, y_n$的过渡矩阵.

例：设复数域上线性空间V的线性变换 T在某组基

x_1, x_2, x_3 下的矩阵为

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

问T是否可对角化．在可对角化的情况下，写出

基变换的过渡矩阵．

解：A的特征多项式为

$$|\lambda I - A| = \begin{vmatrix} \lambda & 0 & -1 \\ 0 & \lambda - 1 & 0 \\ -1 & 0 & \lambda \end{vmatrix} = (\lambda - 1)^2 (\lambda + 1)$$

得A的特征值是1、1、－1.

解齐次线性方程组 $(1 \cdot I - A) X = 0$, 得 $x_1 = x_3$

故其基础解系为：(1,0,1),(0,1,0)

所以， $y_1 = x_1 + x_3$， $y_2 = x_2$

是T 的属于特征值1的两个线性无关的特征向量.

再解齐次线性方程组 $\left(-1\cdot I-A\right)X=0$，得 $\begin{cases} x_1=-x_3 \\ x_2=0 \end{cases}$

故其基础解系为：$(1,0,-1)$

所以， $y_3=x_1-x_3$

是 T 的属于特征值一1的线性无关的特征向量.

y_1,y_2,y_3 线性无关，故 T 可对角化，且

T 在基 y_1,y_2,y_3 下的矩阵为对角矩阵

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix};$$

$$\left(y_1, y_2, y_3\right) = \left(x_1, x_2, x_3\right)\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & -1 \end{pmatrix}.$$

即基x_1, x_2, x_3到 y_1, y_2, y_3 的过渡矩阵为

$$C = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & -1 \end{pmatrix},$$

$$C^{-1}AC = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

例：问A是否可对角化？若可，求可逆矩阵C，使

$C^{-1}AC$ 为对角矩阵. 这里 $A=\begin{pmatrix} 3 & 2 & -1 \\ -2 & -2 & 2 \\ 3 & 6 & -1 \end{pmatrix}$

解: A的特征多项式为

$$|\lambda I - A| = \begin{vmatrix} \lambda-3 & -2 & 1 \\ 2 & \lambda+2 & -2 \\ -3 & -6 & \lambda+1 \end{vmatrix}$$

$$= \lambda^3 - 12\lambda + 16 = (\lambda-2)^2(\lambda+4)$$

得A的特征值是2、2、-4 .

对于特征值2，求出齐次线性方程组

$$\begin{pmatrix} -1 & -2 & 1 \\ 2 & 4 & -2 \\ -3 & -6 & 3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

的一个基础解系：（－2、1、0），（1、0、1）

对于特征值－4，求出齐次方程组

$$\begin{pmatrix} -7 & -2 & 1 \\ 2 & -2 & -2 \\ -3 & -6 & -3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

的一个基础解系： $(\frac{1}{3}, -\frac{2}{3}, 1)$

所以A可对角化.

令 $$C=\begin{pmatrix} -2 & 1 & \frac{1}{3} \\ 1 & 0 & -\frac{2}{3} \\ 0 & 1 & 1 \end{pmatrix}$$

则 $$C^{-1}AC=\begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -4 \end{pmatrix}$$

例： 在 $P[x]_n(n>1)$ 中，求微分变换D的特征多项式.并证明：D在任何一组基下的矩阵都不可能是对角矩阵（即D不可对角化）.

解：在 $P[x]_n$中取一组基：$1,\ x,\ \dfrac{x^2}{2!},\ \cdots,\ \dfrac{x^{n-1}}{(n-1)!}$

则D在这组基下的矩阵为

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

于是

$$|\lambda I - A| = \begin{pmatrix} \lambda & -1 & 0 & \cdots & 0 \\ 0 & \lambda & -1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & -1 \\ 0 & 0 & 0 & \cdots & \lambda \end{pmatrix} = \lambda^n$$

∴ D的特征值为0（n重）.

又由于对应特征值0的齐次线性方程组 $-AX = 0$

的系数矩阵的秩为n－1，从而方程组的基础解系

只含有一个向量，它小于$P[x]_n$的维数n（＞1）.

故D不可对角化 .

本讲主要内容

- 线性变换的特征值与特征向量
- 最小多项式
- 对角矩阵
- 不变子空间

1、定义

设 T 是数域K上线性空间V的线性变换，V_1是V的子空间，若 $\forall x \in V_1$, 有 $T(x) \in V_1$

则称V_1是 T 的不变子空间

注：

V的平凡子空间（V及零子空间）对于V的任意一个变换 T 来说，都是不变子空间.

不变子空间的简单性质

1）两个不变子空间的交与和仍是不变子空间.

2）设 $V_1 = L(x_1, x_2, \cdots x_s)$，则$V_1$是不变子空间

$\Leftrightarrow T(x_1), T(x_2), \cdots, T(x_s) \in V_1.$

证："$\Rightarrow$" 显然成立.

"$\Leftarrow$" 任取 $x \in V_1$，设 $x = k_1 x_1 + k_2 x_2 + \cdots + k_s x_s$，

则 $T(x) = k_1 T(x_1) + k_2 T(x_2) + \cdots + k_s T(x_s).$

由于 $T(x_1), T(x_2), \cdots, T(x_s) \in V_1$， $\therefore$ $T(x) \in V_1.$

故V_1为T 的不变子空间.

一些重要不变子空间

1） 线性变换T的值域 $R(T)$与核 $N(T)$都是T的不变子空间.

证：$\because$ $R(T)=\{T(x)|x\in V\}\subseteq V,$

$\therefore \forall x\in R(T),$ 有$T(x)\in R(T).$

故 $R(T)$为T的不变子空间.

又任取 $x\in N(T),$ 有 $T(x)=0\in N(T).$

$\therefore$ $N(T)$也为T的不变子空间.

2）任何子空间都是数乘变换K的不变子空间.

$\left(\because \forall x \in V_1,\ \ \mathrm{K}x = kx \in V_1\right)$

3）线性变换T的特征子空间 V_{λ_0} 是T的不变子空间.

$\left(\because \forall x \in V_{\lambda o},\ \text{有} T\left(x\right) = \lambda_o x \in V_{\lambda o}.\right)$

4）由T的特征向量生成的子空间是T的不变子空间.

证：设 $x_1, x_2, \cdots, x_s$ 是 T 的分别属于特征值 $\lambda_1, \lambda_2, \cdots, \lambda_s$ 的特征向量. 任取 $x \in L(x_1, x_2, \cdots, x_s)$,

设 $x = k_1 x_1 + k_2 x_2 + \cdots + k_s x_s$, 则

$$T(x) = k_1 \lambda_1 x_1 + k_2 \lambda_2 x_2 + \cdots + k_s \lambda_s x_s \in L(x_1, x_2, \cdots, x_s)$$

$\therefore$ $L(x_1, x_2, \cdots, x_s)$ 为T的不变子空间.

注：

特别地，由 T 的一个特征向量生成的子空间是一个一维不变子空间. 反过来，一个一维不变子空间必可看成是 T 的一个特征向量生成的子空间.

事实上，若 $V_1 = L(x) = \{kx \mid k \in K, x \neq 0\}$.

则 x 为 $L(x)$ 的一组基.　因为V_1为不变子空间，

$\therefore$ $T(x) \in V_1$，即必存在 $\lambda \in K$，使 $T(x) = \lambda x$.

$\therefore$ x 是T 的特征向量.

定理27：设T是线性空间 V^n 的线性变换，且 V^n 可分解为s个T的不变子空间的直和

$$V^n = V_1 \oplus V_2 \oplus \cdots \oplus V_s$$

在每个不变子空间 V_i 中取基

$$x_{i1}, x_{i2}, \cdots, x_{in_i} (i = 1, 2, \cdots, s) \qquad (*)$$

将其合并作为 V^n 的基，则T在该基下的矩阵为

$$A = \text{diag}(A_1, A_2, \cdots, A_s)$$

其中 A_i $(i = 1, 2, \cdots, s)$ 是T在 V_i 的基(*)下的矩阵

作业

- **P79：16、17、18**

本讲主要内容

- λ－矩阵的概念
- 若当(Jordan)标准形
- 欧式空间

引入

由第五讲知，n维线性空间V的线性变换在某组基下的矩阵为对角形$\Leftrightarrow T$ 有n个线性无关的特征向量 .

$\Leftrightarrow T$ 的所有不同特征子空间的维数之和等于n .

可见，并不是任一线性变换都有一组基，使它在这组基下的矩阵为对角形.

本节介绍，在适当选择基条件下，一般的线性变换的矩阵能化简成什么形状.

一、λ—矩阵的概念

定义：

设K是一个数域，λ 是一个文字，$P[\lambda]$是多项式环，若矩阵A的元素是λ的多项式，即 $P[\lambda]$的元素，则称A为λ —矩阵，并把A写成 $A(\lambda)$.

注：

① $\because K \subset P[\lambda]$，$\therefore$ 数域K上的矩阵—数字矩阵也是 λ —矩阵.

② λ—矩阵也有加法、减法、乘法、数量乘法运算，

其定义与运算规律与数字矩阵相同.

③ 对于 $n\times n$ 的 λ—矩阵，同样有行列式 $|A(\lambda)|$,

它是一个 λ 的多项式，且有

$$|A(\lambda)B(\lambda)|=|A(\lambda)||B(\lambda)|.$$

这里 $A(\lambda),B(\lambda)$ 为同级 λ—矩阵.

④ 与数字矩阵一样，λ—矩阵也有子式的概念.

λ—矩阵的各级子式是 λ 的多项式.

定义：若 λ —矩阵 $A(\lambda)$ 中有一个 $r(r \geq 1)$ 级子式不为零，而所有 $r+1$ 级的子式（若有的话）皆为零，则称 $A(\lambda)$ 的**秩为** r ．

零矩阵的秩规定为 $\mathbf{0}$.

λ—矩阵的初等变换

λ—矩阵的初等变换是指下面三种变换：

① 矩阵两行（列）互换位置；

行变换： $r_i \leftrightarrow r_j$ 列变换： $c_i \leftrightarrow c_j$

② 矩阵的某一行（列）乘以非零常数 k；

行变换： kr_i 列变换： kc_i

③ 矩阵的某一行（列）加另一行（列）的 $p(\lambda)$倍，$p(\lambda)$是一个多项式.

行变换：$r_i + p(\lambda)r_j$ 列变换 $c_i + p(\lambda)c_j$

二、λ一矩阵的行列式因子

行列式因子：$D_k(\lambda)$=最大公因式$\{A(\lambda)$的所有k阶子式$\}$

不 变 因子：$d_k(\lambda)=\dfrac{D_k(\lambda)}{D_{k-1}(\lambda)}\left(D_0(\lambda)=1\right)$

初 等 因子：$d_k(\lambda)$的不可约因式

注：考虑 λ一矩阵$\lambda I-A$，可得A的最小多项式

$$m(\lambda)=d_n(\lambda)=\frac{D_n(\lambda)}{D_{n-1}(\lambda)}$$

例：已知 $A=\begin{bmatrix}-1 & 1 & 0\\ 4 & 3 & 0\\ 1 & 0 & 2\end{bmatrix}$，求 $\lambda I-A$ 的全体初等因子

解： $\lambda I-A=\begin{bmatrix}\lambda+1 & -1 & 0\\ 4 & \lambda-3 & 0\\ -1 & 0 & \lambda-2\end{bmatrix}$ $D_1(\lambda)=1$

因为 $\begin{vmatrix}4 & \lambda-3\\ -1 & 0\end{vmatrix}=\lambda-3$ 与 $\begin{vmatrix}4 & 0\\ -1 & \lambda-2\end{vmatrix}=4(\lambda-2)$ 互质

所以 $D_2(\lambda)=1$ $D_3(\lambda)=\det(\lambda I-A)=(\lambda-2)(\lambda-1)^2$

不变因子为 $d_1(\lambda)=1, d_2(\lambda)=1, d_3(\lambda)=(\lambda-1)^2(\lambda-2)$

全体初等因子为 $(\lambda-1)^2,(\lambda-2)$

例、求$\lambda-$矩阵的不变因子

$$1)\quad A(\lambda)=\begin{pmatrix} \lambda^2+\lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & (\lambda+1)^2 \end{pmatrix}$$

$$2)\quad A(\lambda)=\begin{pmatrix} \lambda-2 & -1 & 0 & 0 \\ 0 & \lambda-2 & -1 & 0 \\ 0 & 0 & \lambda-2 & -1 \\ 0 & 0 & 0 & \lambda-2 \end{pmatrix}$$

解：1） $A(\lambda)$ 的非零1级子式为： $A(\lambda)=\begin{pmatrix}\lambda^2+\lambda & 0 & 0\\ 0 & \lambda & 0\\ 0 & 0 & (\lambda+1)^2\end{pmatrix}$

$$\lambda^2+\lambda,\quad \lambda,\quad (\lambda+1)^2.$$

$$\therefore\quad D_1(\lambda)=1$$

$A(\lambda)$ 的非零二级子式为：

$$\begin{vmatrix}\lambda^2+\lambda & 0\\ 0 & \lambda\end{vmatrix}=\lambda^2(\lambda+1),\quad \begin{vmatrix}\lambda & 0\\ 0 & (\lambda+1)^2\end{vmatrix}=\lambda(\lambda+1)^2,$$

$$\begin{vmatrix}\lambda^2+\lambda & 0\\ 0 & (\lambda+1)^2\end{vmatrix}=\lambda(\lambda+1)^3.$$

$$\therefore D_2(\lambda)=\lambda(\lambda+1).$$

$$A(\lambda)=\begin{pmatrix}\lambda^2+\lambda & 0 & 0\\ 0 & \lambda & 0\\ 0 & 0 & (\lambda+1)^2\end{pmatrix}$$

又 $D_3(\lambda)=|A(\lambda)|=\lambda^2(\lambda+1)^3.$

所以，$A(\lambda)$的不变因子为 ：

$$d_1(\lambda)=D_1(\lambda)=1,\quad d_2(\lambda)=\frac{D_2(\lambda)}{D_1(\lambda)}=\lambda(\lambda+1),$$

$$d_3(\lambda)=\frac{D_3(\lambda)}{D_2(\lambda)}=\lambda(\lambda+1)^2.$$

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

2) $\because \quad \begin{vmatrix} -1 & 0 & 0 \\ \lambda-2 & -1 & 0 \\ 0 & \lambda-2 & -1 \end{vmatrix} = -1,$

$\therefore \quad D_3(\lambda)=1.$

又 $\quad D_1(\lambda) \mid D_2(\lambda), \; D_2(\lambda) \mid D_3(\lambda)$

$\therefore \; D_1(\lambda)=D_2(\lambda)=1.$

而 $\quad D_4(\lambda)=|A(\lambda)|=(\lambda-2)^4.$

$\therefore A(\lambda)$ 的不变因子为

$$d_1(\lambda)=d_2(\lambda)=d_3(\lambda)=1, \; d_4(\lambda)=(\lambda-2)^4.$$

初等变换法求初等因子

$$A(\lambda) \to \begin{bmatrix} f_1(\lambda) & & \\ & \ddots & \\ & & f_n(\lambda) \end{bmatrix}$$ 其中 $f_k(\lambda)$ 是首1多项式

$f_k(\lambda)$ 的不可约因式为 $A(\lambda)$ 的初等因子

例：求上例中 $\lambda I - A$ 的全体初等因子

$$\lambda I - A = \begin{bmatrix} \lambda+1 & -1 & 0 \\ 4 & \lambda-3 & 0 \\ -1 & 0 & \lambda-2 \end{bmatrix} \xrightarrow{c_1 \leftrightarrow c_2} \begin{bmatrix} -1 & \lambda+1 & 0 \\ \lambda-3 & 4 & 0 \\ 0 & -1 & \lambda-2 \end{bmatrix}$$

$$\xrightarrow{r_2 \leftrightarrow (\lambda-3) r_1} \begin{bmatrix} -1 & \lambda+1 & 0 \\ 0 & (\lambda-1)^2 & 0 \\ 0 & -1 & \lambda-2 \end{bmatrix} \xrightarrow[c_2-(\lambda+1)c_1]{(-1)c_1} \begin{bmatrix} 1 & 0 & 0 \\ 0 & (\lambda-1)^2 & 0 \\ 0 & -1 & \lambda-2 \end{bmatrix}$$

$$\xrightarrow[r_3+(\lambda-1)^2 r_2]{r_2 \leftrightarrow r_3} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & \lambda-2 \\ 0 & 0 & (\lambda-1)^2(\lambda-2) \end{bmatrix} \xrightarrow[c_3-(\lambda-2)c_2]{(-1)c_2} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & (\lambda-1)^2(\lambda-2) \end{bmatrix}$$

于是 $f_1(\lambda) = 1,\ f_2(\lambda) = 1,\ f_3(\lambda) = (\lambda-1)^2(\lambda-2)$

全体初等因子为 $(\lambda-1)^2, (\lambda-2)$

设 λ－矩阵$A(\lambda)$的秩为 r，对于正整数 k，$1\le k\le r$，$A(\lambda)$中必有非零的 k级子式，$A(\lambda)$中全部 k级子式的首项系数为1的最大公因式 $D_k(\lambda)$，称为 $A(\lambda)$ 的 k 阶行列式因子.

注：若 秩$\left(A(\lambda)\right)=r$，则 $A(\lambda)$ 有 r个行列式因子.

本讲主要内容

- λ－矩阵的概念
- 若当(Jordan)标准形
- 欧式空间

定理：设T是复数域C上的线性空间V_n的线性变换，任取V_n的一组基， T在该基下的矩阵为A， T的特征多项式 $\varphi(\lambda)=(\lambda-\lambda_1)^{m_1}(\lambda-\lambda_2)^{m_2}\cdots(\lambda-\lambda_s)^{m_s}$ $(m_1+m_2+\cdots+m_s$
则V_n可分解为不变子空间的直和

$$V_n=N_1\oplus N_2\oplus\cdots\oplus N_s$$

其中 $N_i=\{x\,|\,(T-\lambda_i T_e)^{m_i}x=0, x\in V^n\}$是线性变换 的核空间。

若给每个子空间N_i选一组基，它们的并构成V_n 的基，且T在该组基下的矩阵为如下形式的对角块矩阵

$$J=\begin{bmatrix} J_1(\lambda_1) & & & \\ & J_2(\lambda_2) & & \\ & & \ddots & \\ & & & J_s(\lambda_s) \end{bmatrix} \qquad \text{其中}\ J_i(\lambda_i)=\begin{bmatrix} \lambda_i & 1 & \cdots & 0 \\ 0 & \lambda_i & \ddots & 0 \\ \vdots & \vdots & \ddots & 1 \\ 0 & 0 & \cdots & \lambda_i \end{bmatrix}$$

定义:在上面的定义中J称为矩阵A的Jordan 标准形，

$J_i(\lambda_i)$ 为 $(\lambda-\lambda_i I)^{m_i}$ 对应的Jordan 块。

如：$\begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$, $\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$, $\begin{pmatrix} i & 1 \\ 0 & i \end{pmatrix}$ 都是若当块；

而下面的准对角形则是一个若当形矩阵.

$$\begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 & 0 \\ 0 & 0 & 0 & -i & 1 & 0 \\ 0 & 0 & 0 & 0 & -i & 1 \\ 0 & 0 & 0 & 0 & 0 & -i \end{pmatrix} = \begin{pmatrix} J(1,2) & & \\ & J(4,1) & \\ & & J(-i,3) \end{pmatrix}$$

定理：设矩阵A为复数域C的矩阵， 特征多项式的分解

$$\varphi(\lambda)=|\lambda I-A|=(\lambda-\lambda_1)^{m_1}(\lambda-\lambda_2)^{m_2}\cdots(\lambda-\lambda_s)^{m_s}$$

存在，则存在非奇异矩阵P 使得 $P^{-1}AP=J$

例如： $A=\begin{bmatrix}-1 & 1 & 0\\ 4 & 3 & 0\\ 1 & 0 & 2\end{bmatrix}$ 的Jordan标准型为 $J=\begin{bmatrix}1 & 1 & 0\\ 0 & 1 & 0\\ 0 & 0 & 2\end{bmatrix}$

特征向量法求初等因子

设 $\varphi(\lambda)=|\lambda I-A|$ 的一个不可约因式为$(\lambda-\lambda_0)^r$，则

$(\lambda-\lambda_0)^r$ 是A 的k个初等因子的乘积

$\Leftrightarrow (\lambda_0 I-A)x=0$ 的基础解系含k个解向量

$\Leftrightarrow$ 对应特征值 λ_0 有k个线性无关的特征向量

$\Leftrightarrow \quad k=n-\mathrm{rank}(\lambda_0 I-A)$

例：求矩阵 $A=\begin{bmatrix}1 & & & \\ 2 & 1 & & \\ 1 & 3 & 2 & \\ 2 & 3 & 1 & 1\end{bmatrix}$ 的Jordan标准型

解：$\varphi(\lambda)=\det(\lambda I-A)=(\lambda-1)^3(\lambda-2)$

由 $\mathrm{rank}(1I-A)=2$ 知，$(\lambda-1)^3$ 是A的4-2=2个初等因子的乘积，即 $(\lambda-1)^2$ 和 $(\lambda-1)$ 的乘积，

故A的初等因子为 $(\lambda-1)^2,\lambda-1,\lambda-2$

A的Jordan标准型 $J=\begin{bmatrix}1 & 1 & & \\ & 1 & & \\ & & 1 & \\ & & & 2\end{bmatrix}$

例：求矩阵 $A=\begin{pmatrix}1 & -1 & 2\\ 3 & -3 & 6\\ 2 & -2 & 4\end{pmatrix}$ 的若当标准形.

解： $\lambda I-A=\begin{pmatrix}\lambda-1 & 1 & -2\\ -3 & \lambda+3 & -6\\ -2 & 2 & \lambda-4\end{pmatrix}$

$$\to\begin{pmatrix}\lambda-1 & 1 & -2\\ -\lambda^2-2\lambda & 0 & 2\lambda\\ -2\lambda & 0 & \lambda\end{pmatrix}\to\begin{pmatrix}1 & -2 & \lambda-1\\ 0 & 2\lambda & -\lambda^2-2\lambda\\ 0 & \lambda & -2\lambda\end{pmatrix}$$

$$\to\begin{pmatrix}1 & 0 & 0\\ 0 & 2\lambda & -\lambda^2-2\lambda\\ 0 & \lambda & -2\lambda\end{pmatrix}\to\begin{pmatrix}1 & 0 & 0\\ 0 & 0 & -\lambda^2+2\lambda\\ 0 & \lambda & -2\lambda\end{pmatrix}$$

$$\to\begin{pmatrix}1&0&0\\0&0&-\lambda^2+2\lambda\\0&\lambda&0\end{pmatrix}\to\begin{pmatrix}1&0&0\\0&\lambda&0\\0&0&\lambda(\lambda-2)\end{pmatrix}$$

$\therefore A$ 的初等因子为 $\lambda,\ \lambda,\ \lambda-2.$

故 A的若当标准形为 $\begin{pmatrix}0&0&0\\0&0&0\\0&0&2\end{pmatrix}.$

例：已知12级矩阵A的不变因子为

$$\underbrace{1,1,\cdots,1}_{9个},(\lambda-1)^2,(\lambda-1)^2(\lambda+1),(\lambda-1)^2(\lambda+1)(\lambda^2+1)^2$$

求A的若当标准形.

解：依题意，A的初等因子为

$$(\lambda-1)^2,\ (\lambda-1)^2,\ (\lambda-1)^2,\ (\lambda+1),$$

$$(\lambda+1),\ (\lambda-i)^2,\ (\lambda+i)^2$$

$\therefore A$的若当标准形为

$$\begin{pmatrix}
1 & 1 & & & & & & & & & & \\
0 & 1 & & & & & & & & & & \\
 & & 1 & 1 & & & & & & & & \\
 & & 0 & 1 & & & & & & & & \\
 & & & & 1 & 1 & & & & & & \\
 & & & & 0 & 1 & & & & & & \\
 & & & & & & -1 & & & & & \\
 & & & & & & & -1 & & & & \\
 & & & & & & & & i & 1 & & \\
 & & & & & & & & & i & & \\
 & & & & & & & & & & -i & 1 \\
 & & & & & & & & & & & -i
\end{pmatrix}$$

矩阵的Jordan标准型的求法

求矩阵的Jordan标准型时，需要计算矩阵的初等因子组，矩阵的全部初等因子，其方法有两类：

（1）行列式法：计算n 阶矩阵A的特征矩阵 $\lambda I - A$ 的各阶行列式因子 $D_k(\lambda)$ ，及其不变因子

$$d_k(\lambda)=\frac{D_k(\lambda)}{D_{k-1}(\lambda)} \quad \left(k=1,2,\cdots,n; D_0(\lambda)=1\right)$$

那么 $d_1(\lambda),d_2(\lambda),\cdots,d_n(\lambda)$ 的全体不可约因式（单因式连同连同其幂指数）为A的初等因子组

（2）初等变换法：

$$\lambda \boldsymbol{I} - \boldsymbol{A} \to \begin{bmatrix} f_1(\lambda) & & \\ & \ddots & \\ & & f_n(\lambda) \end{bmatrix}$$

其中 $f_k(\lambda)$ 是首1多项式

本讲主要内容

- λ－矩阵的概念
- 若当(Jordan)标准形
- 欧式空间

问题的引入：

1、线性空间中，向量之间的基本运算为线性运算，其具体模型为几何空间 R^2、R^3，但几何空间的度量性质(如长度、夹角)等在一般线性空间中没有涉及.

2、在解析几何中，向量的长度，夹角等度量性质都可以通过内积反映出来：

长度： $|x| = \sqrt{x \cdot x}$

夹角 $< x, y >$ ： $\cos < x, y > = \dfrac{x \cdot y}{|x||y|}$

3、几何空间中向量的内积具有比较明显的代数性质.

定义：设V是实数域 R上的线性空间，对V中任意两个向量x y,定义一个二元实函数，记 (x,y)，(x,y)满足性质：$\forall x,y,z\in V$，$\forall k\in R$

(1) $(x,y)=(y,x)$ （交换率）

(2) $(kx,y)=k(x,y)$ （齐次性）

(3) $(x+y,z)=\left(x,z\right)+(y,z)$ （分配率）

(4) $(x,x)\geq 0$,当且仅当 $x=0$ 时$(x,x)=0$. （正定性）

则称 (x,y)为 x和 y 的内积，并称这种定义了内积的实数域 R上的线性空间V为欧氏空间.

注：欧氏空间 V是特殊的线性空间

① V为实数域 R上的线性空间;

② V除向量的线性运算外，还有“内积”运算;

③ $(x, y) \in R.$

例1. 在 R^n 中，对于向量

$$x=(a_1,a_2,\cdots,a_n),\quad y=(b_1,b_2,\cdots,b_n)$$

1）定义 $(x,y)=a_1b_1+a_2b_2+\cdots+a_nb_n$ （1）

易证 (x,y) 满足定义中的性质 (1)～(4).

所以，(x,y) 为内积.

这样 R^n 对于内积 (x,y) 就成为一个欧氏空间.

（ 当 $n=3$ 时，1）即为几何空间 R^3 中内积在直角坐标系下的表达式．(x,y) 即 $x\cdot y$．）

2）定义

$$(x,y)' = a_1b_1 + 2a_2b_2 + \cdots + ka_kb_k + \cdots + na_nb_n$$

易证$(x,y)'$满足定义中的性质(1)～(4).

所以$(x,y)'$也为内积.

从而R^n对于内积$(x,y)'$也构成一个欧氏空间.

注意：由于对 $\forall x,y\in V$，未必有 $(x,y)=(x,y)'$

所以1），2）是两种不同的内积.

从而 R^n 对于这两种内积就构成了不同的欧氏空间.

例2. 在 $R^{m\times n}$ 中: $A=\left(a_{ij}\right)_{m\times n}$, $B=\left(b_{ij}\right)_{m\times n}$

2）定义 $(A,B)=\sum_{i=1}^{m}\sum_{j=1}^{n}a_{ij}b_{ij}=tr(AB^T)$ （2）

易证(A,B)满足定义中的性质 (1)～(4).

所以, (A,B)为内积.

这样 $R^{m\times n}$ 对于内积 (A,B) 就成为一个欧氏空间.

例3. $C[a,b]$ 为闭区间 $[a,b]$ 上的所有实连续函数所成线性空间，对于函数 $f(x), g(x)$ ，定义

$$(f,g)=\int_a^b f(x)g(x)\,dx \qquad (3)$$

则 $C[a,b]$ 对于（3）作成一个欧氏空间.

证：$\forall f(x), g(x), h(x)\in C(a,b), \quad \forall k\in R$

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

(2) $(kf,g)=\int_a^b kf(x)g(x)\,dx=k\int_a^b f(x)g(x)\,dx$

$=k(f,g)$

(3) $$(f+g,h)=\int_a^b\big(f(x)+g(x)\big)h(x)\,dx$$

$$=\int_a^b f(x)h(x)\,dx+\int_a^b g(x)h(x)\,dx$$

$$=(f,h)+(g,h)$$

(4) $$(f,f)=\int_a^b f^2(x)\,dx$$

$\because\ f^2(x)\geq 0,\qquad \therefore\ (f,f)\geq 0.$

且若 $f(x)\neq 0,$ 则 $f^2(x)>0,$ 从而 $(f,f)>0.$

故 $(f,f)=0\Leftrightarrow f(x)=0.$

因此, (f,g) 为内积, $C[a,b]$为欧氏空间.

2. 内积的简单性质

V为欧氏空间， $\forall x, y, z \in V, \forall k \in R$

1) $(x, ky) = k(x, y), \quad (kx, ky) = k^2(x, y)$

2) $(0, y) = (x, 0) = 0$

3) $(x, y+z) = (x, y) + (x, z)$

推广： $\left(\sum_{i=1}^{n} \xi_i x_i, \sum_{i=1}^{n} \eta_j y_j\right) = \sum_{i,j=1}^{n} \xi_i \eta_j (x_i, y_j)$

3. n 维欧氏空间中内积的矩阵表示

设V为欧氏空间, $x_1, x_2, \cdots, x_n$为V的一组基，对V中任意两个向量

$$x = \xi_1 x_1 + \xi_2 x_2 + \cdots + \xi_n x_n$$

$$y = \eta_1 x_1 + \eta_2 x_2 + \cdots + \eta_n x_n$$

$$(x, y) = (\sum_{i=1}^{n} \xi_i x_i, \sum_{j=1}^{n} \eta_j x_j) = \sum_{i=1}^{n} \sum_{j=1}^{n} \varepsilon_i \eta_j (x_i, x_j) \qquad (4)$$

令 $a_{ij} = (x_i, x_j), \quad i, j = 1, 2, \cdots n.$

$$A=\left(a_{ij}\right)_{n\times n},\quad X=\begin{pmatrix}\xi_1\\ \xi_2\\ \vdots\\ \xi_n\end{pmatrix},\quad Y=\begin{pmatrix}\eta_1\\ \eta_2\\ \vdots\\ \eta_n\end{pmatrix} \tag{5}$$

则 $$(x,y)=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}\xi_i\eta_j=X'AY \tag{6}$$

定义：矩阵 $A=\begin{pmatrix}(x_1,x_1) & (x_1,x_2) & \cdots & (x_1,x_n)\\ (x_2,x_1) & (x_2,x_2) & \cdots & (x_2,x_n)\\ \cdots & \cdots & \cdots & \cdots\\ (x_n,x_1) & (x_n,x_2) & \cdots & (x_n,x_n)\end{pmatrix}$

称为基 $x_1,x_2,\cdots,x_n$的度量矩阵.

注：

① 度量矩阵A是实对称矩阵.

② 由内积的正定性，度量矩阵A还是正定矩阵.

事实上，对 $\forall x \in V, x \neq 0$，即 $X = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{pmatrix} \neq 0$

有 $(x, x) = X'AX > 0$

$\therefore$ A为正定矩阵.

③对同一内积而言，不同基的度量矩阵是合同的.

$$B = C^T AC$$

证：设$x_1, x_2, \cdots, x_n$；$y_1, y_2, \cdots, y_n$为欧氏空间V的两组基，它们的度量矩阵分别为A、B，且

$$(y_1, y_2, \cdots, y_n) = (x_1, x_2, \cdots, x_n)C$$

设 $C = \left(c_{ij}\right)_{n\times n} = \left(C_1, C_2, \cdots, C_n\right),$

则 $$y_i = \sum_{k=1}^{n} c_{ki} x_k, i = 1, 2, \cdots, n$$

于是

$$(y_i, y_j) = \left(\sum_{k=1}^{n} c_{ki} x_k , \sum_{l=1}^{n} c_{lj} x_l \right) = \sum_{k=1}^{n} \sum_{l=1}^{n} (x_k, x_l) c_{ki} c_{lj}$$

$$= \sum_{k=1}^{n} \sum_{l=1}^{n} a_{kl} c_{ki} c_{lj} = C_i' A C_j$$

$$\therefore \quad B = \left((y_i, y_j) \right) = \left(C_i' A C_j \right)$$

$$= \begin{pmatrix} C_1' \\ C_2' \\ \vdots \\ C_n' \end{pmatrix} A \left(C_1, C_2, \cdots, C_n \right) = C' A C$$

④向量的内积由度量矩阵A完全确定,与基的选择无关

$$\forall x, y \in V^n$$

$$x = \xi_1 x_1 + \xi_2 x_2 + \cdots + \xi_n x_n = \left(x_1, x_2, \cdots, x_n\right) X_1$$

$$x = \eta_1 y_1 + \eta_2 y_2 + \cdots + \eta_n y_n = \left(y_1, y_2, \cdots, y_n\right) X_2$$

$$= \left(x_1, x_2, \cdots, x_n\right) C X_2 \qquad \Rightarrow X_1 = C X_2$$

$$y = \left(x_1, x_2, \cdots, x_n\right) Y_1$$

$$\Rightarrow Y_1 = C Y_2$$

$$y = \left(y_1, y_2, \cdots, y_n\right) Y_2 = \left(x_1, x_2, \cdots, x_n\right) C Y_2$$

基（I）下：$\left(x, y\right) = X_1^T A Y_1$

基（II）下：$\left(x, y\right) = X_2^T B Y_2 = X_2^T C^T A C Y_2 = X_1^T A Y_1$

3．欧氏空间中向量的长度

(1) 引入长度概念的可能性

1）在 R^3 向量 x 的长度（模）$|x| = \sqrt{x \cdot x}$.

2) 欧氏空间V中，$\forall x \in V$， $(x, x) \geq 0$

使得 $\sqrt{x \cdot x}$ 有意义.

2．向量长度的定义

$\forall x \in V$， $|x| = \sqrt{(x, x)}$ 称为向量 x 的长度（模）.

特别地，当 $|x| = 1$ 时，称 x 为单位向量.

3. 向量长度的简单性质

1) $|x| \geq 0;\quad |x| = 0 \Leftrightarrow x = 0$

2) $|kx| = |k||x|$

3）非零向量 x 的单位化： $\dfrac{1}{|x|}x$.

4.欧氏空间中向量的夹角

（1） 引入夹角概念的可能性与困难

1）在 R^3 中向量 x 与 y 的夹角

$$< x, y > = \arccos \frac{x \cdot y}{|x||y|} \qquad (*)$$

2）在一般欧氏空间中推广（*）的形式，首先

应证明不等式：$$\left| \frac{(x, y)}{|x||y|} \right| \le 1$$

此即,

2．柯西－布涅柯夫斯基不等式

对欧氏空间V中任意两个向量 x、y，有

$$|(x,y)| \le |x||y| \qquad (**)$$

当且仅当 x、y 线性相关时等号成立.

证：当 $y=0$ 时， $(x,0)=0,\quad |y|=0$

$\therefore\quad (x,y)=|x||y|=0.$ 结论成立.

当 $y\neq 0$ 时，作向量 $z=x+ty,\qquad t\in R$

由内积的正定性，对 $\forall t \in R$，皆有

$$(z,z)=(x+ty,x+ty)$$

$$=(x,x)+2(x,y)t+(y,y)t^2 \geq 0 \qquad (1)$$

取 $t=-\dfrac{(x,y)}{(y,y)}$ 代入（1）式，得

$$(x,x)-2(x,y)\frac{(x,y)}{(y,y)}+(y,y)\frac{(x,y)^2}{(y,y)^2} \geq 0$$

即 $\quad (x,y)^2 \leq (x,x)(y,y)$

两边开方，即得 $\quad \left|(x,y)\right| \leq |x||y|.$

当 x、y 线性相关时，不妨设 $x = ky$

于是，$\left|(x,y)\right| = \left|(ky,y)\right| = \left|k(y,y)\right| = |k||y|^2$.

$$|x||y| = |ky||y| = |k||y|^2$$

$\therefore\ \left|(x,y)\right| = |x||y|$. (**)式等号成立.

反之，若（**）式等号成立，由以上证明过程知

或者 $y = 0$，或者 $x - \dfrac{(x,y)}{(y,y)}y = 0$

也即 x、y 线性相关.

3．柯西—布涅柯夫斯基不等式的应用

柯西
不等式

1）
$$\left|a_1b_1+a_2b_2+\cdots+a_nb_n\right|$$

$$\le\sqrt{a_1^2+a_2^2+\cdots+a_n^2}\sqrt{b_1^2+b_2^2+\cdots+b_n^2}$$

施瓦兹
不等式

$$a_i,\, b_i\in R,\quad i=1,2,\cdots,n.$$

2）
$$\left|\int_a^b f(x)g(x)dx\right|\le\sqrt{\int_a^b f^2(x)dx}\sqrt{\int_a^b g^2(x)dx}$$

证：在 $C(a,b)$ 中， $f(x)$ 与 $g(x)$ 的内积定义为

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

由柯西—布涅柯夫斯基不等式有

$$\left|(f(x),g(x))\right|\le\left|f(x)\right|\left|g(x)\right|$$

从而得证.

3)

三角
不等式

对欧氏空间中的任意两个向量 x、y，有

$$|x+y| \le |x| + |y| \qquad (*)$$

证：$|x+y|^2 = (x+y, x+y)$

$$= (x,x) + 2(x,y) + (y,y)$$

$$\le |x|^2 + 2|x||y| + |y|^2 = (|x|+|y|)^2$$

两边开方，即得（*）成立.

欧氏空间中两非零向量的夹角

定义1：设V为欧氏空间，x、y 为V中任意两非零向量，x、y 的夹角定义为

$$\langle x, y\rangle = arc\cos\frac{(x, y)}{|x||y|}$$

$$\left(0 \le \langle x, y\rangle \le \pi \right)$$

定义2：设 x、y 为欧氏空间中两个向量，若内积

$$(x, y) = 0$$

则称 x 与 y **正交**或**互相垂直**，记作 $x \perp y$.

注：

① 零向量与任意向量正交.

② $x \perp y \Longleftrightarrow \langle x, y\rangle = \dfrac{\pi}{2}$，即 $\cos\langle x, y\rangle = 0$.

勾股定理

设V为欧氏空间， $\forall x, y \in V$

$$x \perp y \Longleftrightarrow |x+y|^2 = |x|^2 + |y|^2$$

证： $\because \quad |x+y|^2 = (x+y, x+y)$

$$= (x,x) + 2(x,y) + (y,y)$$

$$\therefore \quad |x+y|^2 = |x|^2 + |y|^2 \Longleftrightarrow (x,y) = 0$$

$$\Longleftrightarrow \quad x \perp y.$$

定理：若欧氏空间V中向量 $x_1, x_2, \cdots, x_m$ 两两正交，

即 $(x_i, x_j) = 0, \quad i \neq j, \quad i, j = 1, 2, \cdots, m$

则 $|x_1 + x_2 + \cdots + x_m|^2 = |x_1|^2 + |x_2|^2 + \cdots + |x_m|^2.$

证：若 $(x_i, x_j) = 0, \quad i \neq j$

则 $|x_1 + x_2 + \cdots + x_m|^2 = (\sum_{i=1}^{m} x_i, \sum_{j=1}^{m} x_j)$

$$= \sum_{i=1}^{m} (x_i, x_i) + \sum_{i \neq j}^{m} (x_i, x_j)$$

$$= \sum_{i=1}^{m} (x_i, x_i) = |x_1|^2 + |x_2|^2 + \cdots + |x_m|^2$$

例：已知 $x=(2,1,3,2)$， $y=(1,2,-2,1)$

在通常的内积定义下，求 $|x|,(x,y),\langle x,y\rangle,|x-y|$.

解：$|x|=\sqrt{(x,x)}=\sqrt{2^2+1^2+3^2+2^2}=\sqrt{18}=3\sqrt{2}$

$(x,y)=2\times1+1\times2+3\times(-2)+2\times1=0 \quad \therefore \ \langle x,y\rangle=\dfrac{\pi}{2}$

又 $x-y=(1,-1,5,1)$

$\therefore \ |x-y|=\sqrt{1^2+(-1)^2+5^2+1^2}=\sqrt{28}=2\sqrt{7}$

通常称 $|x-y|$为 x 与 y 的距离，记作 $d(x,y)$.

定理：欧式空间 $V^n(n>1)$ 存在标准正交基

证：对 V^n 的基 $x_1,x_2,\cdots,x_n$ 进行正交化，可得正交基

$$y_1,y_2,\cdots,y_n \qquad\qquad (y_i,y_j)=0,\quad i\neq j$$

再进行单位化，可得单位正交基

$$z_1,z_2,\cdots,z_n \qquad\qquad z_j=\frac{y_j}{|y_j|}$$

例：标准正交基的特征：欧式空间V^n的标准正交基

$$x_1, x_2, \cdots, x_n$$

且 $x = \xi_1 x_1 + \xi_2 x_2 + \cdots + \xi_n x_n, y = \eta_1 x_1 + \eta_2 x_2 + \cdots + \eta_n x_n$

（1） 基 $x_1, x_2, \cdots, x_n$ 的度量矩阵$A=I$

（2） $\xi_i = (x, x_i), \eta_j = (y, y_j)$

（3） $(x, y) = \sum_{i=1}^{n} \xi_i \eta_i$

五：欧氏空间的子空间

欧氏空间V的子空间在所V中定义的内积之下也是

一个欧氏空间，称之为V的欧氏子空间.

欧氏空间V的子空间V_1,给定 $y \in V$，若 $\forall x \in V_1$

都有 $y \perp x$ ，则称y 正交于V_1 ，记做 $y \perp V_1$

欧氏空间 V^n，子空间 V_1，则 $V_1^{\perp}=\{y \mid y\in V, y\perp V_1\}$

是V^n的子空间

证明：$\theta\in V_1^{\perp}\Rightarrow V_1^{\perp}$非空。$\forall y,z\in V_1^{\perp},\forall x\in V_1,\forall k\in R$

$$\left(y+z,x\right)=\left(y,x\right)+\left(z,x\right)=0\Rightarrow\left(y+z\right)\perp V_1^{\perp}:\left(y+z\right)\in V_1^{\perp}$$

$$\left(ky,x\right)=k\left(y,x\right)=0\Rightarrow\left(ky\right)\perp V_1\ :\left(ky\right)\in V_1^{\perp}$$

所以 $V_1^{\perp}$是 V^n的子空间。（称$V_1^{\perp}$为是V_1 的正交补)

定理：设欧氏空间 V^n，子空间 V_1，则

$$V^n=V_1\oplus V_1^{\perp}$$

定理：设 $A=\left(a_{ij}\right)_{m\times n}\in R^{m\times n}$ ，则

(1) $\left[R(A)\right]^{\perp}=N(A^T)$, 且 $R(A)\oplus N(A^T)=R^m$

(2) $\left[R(A^T)\right]^{\perp}=N(A)$, 且 $R(A^T)\oplus N(A)=R^n$

三、欧氏空间中的正交变换

1. 定义

欧氏空间V的线性变换 T 如果保持向量的内积不变，

即，$$\big(T(x),T(y)\big)=(x,y),\quad \forall x,y\in V$$

则称 T 为正交变换.

注：欧氏空间中的正交变换是几何空间中保持长度不变的正交变换的推广.

2.欧氏空间中的正交变换

定理：设 T 是欧氏空间V的一个线性变换.

下述命题是等价的：

1） T 是正交变换；

2） T 保持向量长度不变，即

$$|T(x)| = |x|, \qquad \forall x \in V;$$

3） T 保持向量间的距离不变，即

$$d\big(T(x), T(y)\big) = d\big(x, y\big), \qquad \forall x, y \in V$$

证明：首先证明1)与2)等价.

1) $\Rightarrow$ 2)：若 T 是正交变换，则

$$\left(T(x),T(x)\right)=(x,x),\quad \forall x\in V$$

即，$$\left|T(x)\right|^2=\left|x\right|^2$$

两边开方得，$$\left|T(x)\right|=\left|x\right|,\quad \forall x\in V,$$

2) $\Rightarrow$ 1)：若 T 保持向量长度不变，则对 $\forall x,y\in V$

有，$$\left(T(x),T(x)\right)=(x,x),\qquad (1)$$

$$\left(T(y),T(y)\right)=(y,y),\qquad (2)$$

$$\left(T(x+y),T(x+y)\right)=(x+y,x+y), \qquad (3)$$

把(3)展开得，

$$\left(T(x),T(x)\right)+2\left(T(x),T(y)\right)+\left(T(y),T(y)\right)$$

$$=(x,x)+2(x,y)+(y,y)$$

再由(1)(2)即得，

$$\left(T(x),T(y)\right)=(x,y)$$

$\therefore$ T 是正交变换.

再证明2)与3)等价.

$$2) \Rightarrow 3): \quad \because \quad T(x)-T(y)=T(x-y),$$

$$\therefore \quad d\big(T(x),T(y)\big)=\left|T(x)-T(y)\right|$$

$$=\left|T(x-y)\right|=\left|x-y\right| \qquad (\text{根据2)})$$

$$=d(x,y)$$

故 3）成立.

$$3) \Rightarrow 2): \text{若} \quad d\big(T(x),T(y)\big)=d\big(x,y\big), \quad \forall x,y \in V$$

$$\text{则有，} \quad d\big(T(x),T(0)\big)=d\big(x,0\big), \quad \forall x \in V$$

$$\text{即，} \quad \left|T(x)\right|=\left|x\right|, \quad \forall x \in V.$$

故 2）成立.

1． n维欧氏空间中的正交变换是保持标准正交基不变的线性变换.

1). 若 T 是 n 维欧氏空间V的正交变换，$x_1, x_2, \cdots, x_n$ 是V的标准正交基，则 $T(x_1), T(x_2), \cdots, T(x_n)$ 也是V的标准正交基.

事实上，由正交变换的定义及标准正交基的性质即有，

$$\left(T(x_i), T(x_j)\right) = (x_i, x_j) = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

2). 若线性变换 T 使V的标准正交基 $x_1, x_2, \cdots, x_n$变成标准正交基$T(x_1), T(x_2), \cdots, T(x_n)$，则 T 为V的正交变换.

证明：任取 $x, y \in V$，设

$$x = \xi_1 x_1 + \xi_2 x_2 + \cdots \xi_n x_n$$

$$y = \eta_1 x_1 + \eta_2 x_2 + \cdots \eta_n x_n,$$

由 $x_1, x_2, \cdots, x_n$为标准正交基，有

$$(x, y) = \sum_{i=1}^{n} \xi_i \eta_i$$

又 $T(x)=\sum_{i=1}^{n}\xi_i T(x_i)$, $T(y)=\sum_{j=1}^{n}\eta_j T(x_j)$

由于$T(x_1),T(x_2),\cdots,T(x_n)$为标准正交基，得

$$\left(T(x),T(y)\right)=\sum_{i=1}^{n}\xi_i\eta_i$$

$\therefore \quad \left(T(x),T(y)\right)=(x,y)$

故 T 是正交变换.

2． n维欧氏空间V中的线性变换 T是正交变换

$\Longleftrightarrow T$ 在任一组标准正交基下的矩阵是正交矩阵.

证明："$\Rightarrow$" 设$x_1, x_2, \cdots, x_n$为V的标准正交基，且

$$T(x_1, x_2, \cdots, x_n) = (Tx_1, Tx_2, \cdots, Tx_n)$$
$$= (x_1, x_2, \cdots, x_n) A$$

当 T 是正交变换时，由1知，$Tx_1, Tx_2, \cdots, Tx_n$ 也是V的标准正交基， 而由标准正交基 $x_1, x_2, \cdots, x_n$ 到标准正交基$Tx_1, Tx_2, \cdots, Tx_n$的过渡矩阵是正交矩阵.

所以，A是正交矩阵.

$"\Leftarrow"$ 设$x_1, x_2, \cdots, x_n$为V的标准正交基，且

$$T(x_1, x_2, \cdots, x_n) = (x_1, x_2, \cdots, x_n) A$$

即，$(Tx_1, Tx_2, \cdots, Tx_n) = (x_1, x_2, \cdots, x_n) A$

由于当A是正交矩阵时，$Tx_1, Tx_2, \cdots, Tx_n$也是V的标准正交基，再由1即得T为正交变换.

3. n维欧氏空间中正交变换的分类：

设 n 维欧氏空间V中的线性变换 T 在标准正交基 $x_1, x_2, \cdots, x_n$下的矩阵是正交矩阵A，则 $|A| = \pm 1.$

1）如果 $|A| = 1$，则称 T 为第一类的（旋转）；

2）如果 $|A| = -1$，则称 T 为第二类的.

例、在欧氏空间中任取一组标准正交基 $x_1, x_2, \cdots, x_n$，

定义线性变换 T 为：

$$Tx_1 = -x_1$$

$$Tx_i = x_i, \qquad i = 2, 3, \cdots n.$$

则 T 为第二类的正交变换，也称之为镜面反射.

四、实对称矩阵的一些性质

定理：设A是实对称矩阵，则A的特征值皆为实数.

定理：设A是实对称矩阵，在n维欧氏空间R^n上定义一个线性变换T如下：

$$T(x) = Ax, \qquad \forall x \in R^n$$

则对任意 $x, y \in R^n$，有

$$\big(T(x), y\big) = \big(x, T(y)\big),$$

或

$$y'(Ax) = x'(Ay).$$

证：取R^n的一组标准正交基，

$$e_1=\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}, e_2=\begin{pmatrix}0\\1\\\vdots\\0\end{pmatrix}, \dots, e_n=\begin{pmatrix}0\\\vdots\\0\\1\end{pmatrix}$$

则T在基 $e_1,e_2,\dots,e_n$下的矩阵为A，即

$$T(e_1,e_2,\dots,e_n)=(e_1,e_2,\dots,e_n)A$$

任取 $x=\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix},\quad y=\begin{pmatrix}y_1\\y_2\\\vdots\\y_n\end{pmatrix}\in R^n,$

即 $x = x_1e_1 + x_2e_2 + \dots + x_ne_n \overset{\triangle}{=} (e_1, e_2, \dots, e_n)X,$

$$y = y_1e_1 + y_2e_2 + \dots + y_ne_n \overset{\triangle}{=} (e_1, e_2, \dots, e_n)Y,$$

于是

$$T(x) = T(e_1, e_2, \dots, e_n)X = (e_1, e_2, \dots, e_n)AX,$$

$$T(y) = T(e_1, e_2, \dots, e_n)Y = (e_1, e_2, \dots, e_n)AY,$$

又 $e_1, e_2, \dots, e_n$是标准正交基,

$$\therefore \quad \left(T(x), y\right) = (AX)'Y = (X'A')Y = X'AY$$

$$= X'(AY) = \left(x, T(y)\right)$$

又注意到在R^n中 $x = X$， $y = Y$,

即有 $$y(Ax) = \left(y, T(x)\right) = \left(T(x), y\right)$$

$$= \left(x, T(y)\right) = x'(Ay).$$

二、对称变换

1. 定义

设 T 为欧氏空间V中的线性变换，如果满足

$$\big(T(x),y\big)=\big(x,T(y)\big),\qquad \forall x,y\in V,$$

则称 T 为对称变换.

1）n维欧氏空间V的对称变换与n级实对称矩阵在标准正交基下是相互确定的：

① 实对称矩阵可确定一个对称变换.

事实上，设 $A \in R^{n\times n}, A' = A,$ $x_1, x_2, \ldots, x_n$为V的一组标准正交基. 定义V的线性变换 T ：

$$T(x_1, \ldots x_n) = (x_1, \ldots x_n)A$$

则 T 即为V的对称变换.

② 对称变换在标准正交基下的矩阵是实对称矩阵.

事实上，设 T 为n维欧氏空间V上的对称变换，

$x_1, x_2, \cdots, x_n$ 为V的一组标准正交基， $A = (a_{ij}) \in R^{n\times n}$

为 T 在这组基下的矩阵，即

$$T(x_1, x_2, \cdots, x_n) = (x_1, x_2, \cdots, x_n)A$$

或

$$T(x_i) = a_{1i}x_1 + a_{2i}x_2 + \cdots + a_{ni}x_n$$

$$= \sum_{k=1}^{n} a_{ki}x_k, \qquad i = 1, 2, \cdots, n$$

于是 $\left(T(x_i), x_j\right) = \left(\sum_{k=1}^{n} a_{ki} x_k,\ x_j\right) = \sum_{k=1}^{n} a_{ki}(x_k, x_j)$

$$= a_{ji}(x_j, x_j) = a_{ji}$$

$$\left(x_i, T(x_j)\right) = \left(x_i, \sum_{k=1}^{n} a_{kj} x_k \right) = \sum_{k=1}^{n} a_{kj}(x_i, x_k)$$

$$= a_{ij}(x_i, x_i) = a_{ij}$$

由 T 是对称变换，有 $\left(T(x_i), x_j\right) = \left(x_i, T(x_j)\right)$

即 $\alpha_{ij} = \alpha_{ji}, \qquad i, j = 1, 2, \cdots\cdots n,$

所以A为对称矩阵.

定理：

设实对称矩阵A的特征值 $\lambda_1 \neq \lambda_2$，对应的特征向量为 x_1 和 x_2，则 $\left(x_1, x_2\right) = 0$

证明：$\left.\begin{array}{l} Ax_1 = \lambda_1 x_1 \\ Ax_2 = \lambda_2 x_2 \end{array}\right\} \Rightarrow x_1^T A x_2 = \begin{cases} x_1^T\left(Ax_2\right) = \lambda_2\left(x_1^T x_2\right) \\ \left(Ax_1\right)^T x_2 = \lambda_1\left(x_1^T x_2\right) \end{cases}$

$$\Rightarrow \left(\lambda_2 - \lambda_1\right)\left(x_1^T x_2\right) = 0 \qquad \Rightarrow x_1^T x_2 = 0$$

即：$\left(x_1, x_2\right) = 0$

作业

- P79：19
- P106：1
- P107：6、9、10

本讲主要内容

- 向量范数及l_p范数

定义：如果V是数域K上的线性空间，且对于V的任一向量x，对应一个实数值 $\|x\|$，满足以下三个条件

1) 非负性： $\|x\| \geq 0$， 且 $\|x\| = 0 \Leftrightarrow x = 0$

2) 齐次性： $\|kx\| = |k| \cdot \|x\|, \quad \forall k \in K$

3) 三角不等式：$\|x + y\| \leq \|x\| + \|y\|$

则称 $\|x\|$ 为V上向量x的范数，简称为向量范数。

注意：2)中 $|k|$ 当K为实数时为绝对值，

当K为复数域时为复数的模。

向量的范数具有下列简单性质：

(1) 当 $\|x\| \neq 0$ 时，$\left\| \frac{1}{\|x\|} x \right\| = 1$ $\quad \because \left\| \frac{1}{\|x\|} x \right\| = \frac{1}{\|x\|} \|x\| = 1$

(2) $\forall x \in V$ ，$\|-x\| = \|x\|$ $\quad \because \|-x\| = |-1| \|x\| = \|x\|$

(3) $\forall x, y \in V$ ，$\|x\| - \|y\| \leq \|x - y\|$

$$\because \|x\| = \|(x - y) + y\| \leq \|x - y\| + \|y\| \Rightarrow \|x\| - \|y\| \leq \|x - y\|$$

(4) $\forall x, y \in V$ ，$\big| \|x\| - \|y\| \big| \leq \|x - y\|$

$$\|\boldsymbol{x}\| = \|(\boldsymbol{x} - \boldsymbol{y}) + \boldsymbol{y}\| \leq \|\boldsymbol{x} - \boldsymbol{y}\| + \|\boldsymbol{y}\| \Rightarrow \|\boldsymbol{x}\| - \|\boldsymbol{y}\| \leq \|\boldsymbol{x} - \boldsymbol{y}\|$$

同样 $\|\boldsymbol{y}\| - \|\boldsymbol{x}\| \leq \|\boldsymbol{y} - \boldsymbol{x}\| = \|\boldsymbol{x} - \boldsymbol{y}\|$

例1：线性空间C^n，设 $x=(x_1,x_2,\cdots,x_n)^T\in C^n$

1： $\|x\|_1=\sum|\xi_i|$ 是一种向量范数，记为1-范数

2： $\|x\|=\sqrt{(x,x)}$ 是一种向量范数，记2-范数

3: $\|x\|=\max\limits_i|x_i|$ 是一种向量范数,记为 ∞ -范数

4： $\|x\|_p=\left(\sum\limits_{i=1}^{n}|x_i|^p\right)^{\frac{1}{p}}$ $(1\le p<\infty)$

是一种向量范数，记为p-范数或 l_p 范数

证明：向量p -范数 $\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}} \quad (1 \le p < \infty)$

证：性质（1）、（2）显然是满足的

设 $x = (\xi_1, \xi_2, \cdots, \xi_n), y = (\eta_1, \eta_2, \cdots, \eta_n)$ ，则

$$p = 1: \|x + y\|_1 = \sum |\xi_i + \eta_i| \le \sum |\xi_i| + |\eta_i| = \|x\|_1 + \|y\|_1$$

$p > 1: x + y = \theta$ 时,结论成立; $x + y \ne \theta$ 时,应用Holder不等式

$$\sum |a_i b_i| \le \left(\sum |a_i|^p\right)^{\frac{1}{p}} \left(\sum |b_i|^q\right)^{\frac{1}{q}} (p > 1, q > 1, \frac{1}{p} + \frac{1}{q} = 1)$$

(利用 $(p-1)q = p$)

$$\left(\|x+y\|_p\right)^p=\sum_{i=1}^{n}|\xi_i+\eta_i|^p=\sum_{i=1}^{n}\left(|\xi_i+\eta_i|\cdot|\xi_i+\eta_i|^{p-1}\right)$$

$$\le\sum_{i=1}^{n}|\xi_i||\xi_i+\eta_i|^{p-1}+\sum_{i=1}^{n}|\eta_i||\xi_i+\eta_i|^{p-1}\qquad \sum|a_ib_i|\le\left(\sum|a_i|^p\right)^{\frac{1}{p}}\left(\sum|b_i|^q\right)^{\frac{1}{q}}$$

$$\le\left(\sum_{i=1}^{n}|\xi_i|^P\right)^{1/P}\left(\sum\left(|\xi_i+\eta_i|^{p-1}\right)^q\right)^{1/q}+\left(\sum_{i=1}^{n}|\eta_i|^P\right)^{1/P}\left(\sum\left(|\xi_i+\eta_i|^{p-1}\right)^q\right)^{1/q}$$

$$\le\|x\|_p\left(\sum|\xi_i+\eta_i|^p\right)^{1/q}+\|y\|_p\left(\sum|\xi_i+\eta_i|^p\right)^{1/q}\qquad \frac{1}{q}=\frac{p-1}{p}$$

$$=\left(\|x\|_p+\|y\|_p\right)\left(\|x+y\|_p\right)^{p-1}$$

因此： $\left(\|x\|_p+\|y\|_p\right)\ge\|x+y\|_p$

所以 $\|x\|_p=\left(\sum_{i=1}^{n}|x_i|^p\right)^{\frac{1}{p}}$ $(1\le p<\infty)$ 是向量 x 的范数

例2：线性空间V^n中，任取它的一组基 $x_1,\cdots,x_n$

则对于任意向量x,它可以表示为

$$x=\xi_1 x_1+\cdots+\xi_n x_n$$

与 $\alpha=(\xi_1,\cdots,\xi_n)^T\in C^n$ 是同构的

所以 $\|x\|_p=\|\alpha\|_p$ 是V^n中元素x的p一范数

例3: $C[a,b]$为闭区间 $[a,b]$上的所有实连续函数所成线性空间，可以验证以下定义式均满足范数条件

$$\|f(x)\|_1=\int_a^b f(x)dt \qquad \|f(x)\|_\infty=\max_{t\in[a,b]}|f(x)|$$

$$\|f(x)\|_p=\left(\int_a^b |f(x)|^p dt\right)^{\frac{1}{p}},1<p<\infty$$

例4：设A为n阶实对称正定矩阵，对$x\in Rn$,

定义$\|x\|_A=\left(x^TAx\right)^{1/2}$称为加权范数或椭圆范数

由正定矩阵定义可知$\|x\|_A=0\Leftrightarrow x=0$; $\|x\|_A\neq 0\Leftrightarrow x\neq 0$

对任意数 $\alpha\in R$ ，有

$$\|\alpha x\|_A=\sqrt{(\alpha x)^T A\alpha x}=\sqrt{\alpha^2x^TAx}=|\alpha|\sqrt{x^TAx}=|\alpha|\|x\|_A$$

由A正定且实对称 $\Rightarrow\exists$ 正交矩阵Q，使得

$$Q^TAQ=diag(\lambda_1,\cdots,\lambda_n),\quad \lambda_i>0,i=1,\cdots,n$$

定义 $B=diag(\sqrt{\lambda_1},\cdots,\sqrt{\lambda_n})Q^T$ 可得 $A=B^TB$

$$\because \|x\|_A=\left(x^TB^TBx\right)^{1/2}=\left[(Bx)^T(Bx)\right]^{1/2}=\|Bx\|_2$$

$$\therefore \|x+y\|_A=\|B(x+y)\|_2\leq\|Bx\|_2+\|By\|_2=\|x\|_A+\|y\|_A$$

例5：设$\|y\|_\alpha$ 是 C^m中的一个向量范数，给定矩阵 $A\in C^{m\times n}$，它的n个列向量线性无关。对于C^m中的一个向量$x=(x_1,x_2,\cdots,x_n)^T$，规定$\|x\|_\beta=\|Ax\|_\alpha$ 则 $\|x\|_\beta$ 也是 C^m 中的一个向量范数。

证：1）设$A=(a_1,a_2,\cdots,a_n)$，由假设知$a_1,a_2,\cdots,a_n$ 线性无关。

当 $x\neq 0$ $\quad Ax=(a_1,\cdots,a_n)\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}=a_1x_1+\cdots+a_nx_n\neq 0$

又因为$\|y\|_\alpha$是C^m中的一个向量范数，有 $\|Ax\|_\alpha>0$

即 $\|x\|_\beta>0$

当 $x=0$ 时，$Ax=0$ ，所以 $\|x\|_\beta=\|Ax\|_\alpha=0$

2） $\forall k\in C,\ \|kx\|_\beta=\|A(kx)\|_\alpha=\|kAx\|_\alpha=|k|\|Ax\|_\alpha=|k|\|x\|_\beta$

3） $\forall x=(x_1,x_2,\cdots,x_n)^T, y=(y_1,y_2,\cdots,y_n)^T\in C^n$ 有

$$\|x+y\|_\beta=\|A(x+y)\|_\alpha=\|Ax+Ay\|_\alpha\le\|Ax\|_\alpha+\|Ay\|_\alpha=\|x\|_\beta+\|y\|_\beta$$

所以，$\|x\|_\beta$ 是 C^n 中的一个向量范数。

由此可知，当给定 $A\in C^{m\times n}$ 时，可以由 C^m 中的一个向量范数确定 C^n 中的一个向量范数。

三、范数等价

定义：有限维线性空间 V^n 中任意两个向量范数 $\|x\|_\alpha$ 和 $\|x\|_\beta$，如果存在着正常数 c_1 和 c_2，使得 $c_1\|x\|_\beta \le \|x\|_\alpha \le c_2\|x\|_\beta \qquad (\forall x \in V^n)$

则称范数 $\|x\|_\alpha$ 与 $\|x\|_\beta$ 等价

（1）自反性：$1\cdot\|x\|_\alpha \le \|x\|_\alpha \le 1\cdot\|x\|_\alpha, \forall x \in V^n$

（2）对称性：$\frac{1}{c_2}\|x\|_\alpha \le \|x\|_\beta \le \frac{1}{c_1}\|x\|_\alpha, \forall x \in V^n$

（3）传递性：$\left.\begin{aligned} c_1\|x\|_\beta \le \|x\|_\alpha \le c_2\|x\|_\beta \\ c_3\|x\|_\gamma \le \|x\|_\beta \le c_4\|x\|_\gamma \end{aligned}\right\} \quad \forall x \in V^n$

$$\Rightarrow c_5\|x\|_\gamma \le \|x\|_\alpha \le c_6\|x\|_\gamma$$

例6：向量空间 V^n 中，对 $\forall x=(\xi_1,\xi_2,\cdots,\xi_n)^T$ ，有

（1） $\|x\|_1=\sum|\xi_i|\le n\bullet\max_i|\xi_i|=n\|x\|_\infty \quad \|x\|_1\ge\sum|\xi_i|=1\bullet\|x\|_\infty$

$\therefore 1\bullet\|x\|_\infty\le\|x\|_1\le n\bullet\|x\|_\infty$

（2） $\|x\|_2=\left(\sum|\xi_i|^2\right)^{\frac{1}{2}}\le\left(n\bullet\max_i|\xi_i|^2\right)^{\frac{1}{2}}=\sqrt{n}\|x\|_\infty$

$\|x\|_2\ge\left(\max_i|\xi_i|^2\right)^{\frac{1}{2}}=1\bullet\|x\|_\infty \qquad \therefore 1\bullet\|x\|_\infty\le\|x\|_2\le n\bullet\|x\|_\infty$

（3） $\frac{1}{\sqrt{n}}\bullet\|x\|_2\le\|x\|_1\le n\bullet\|x\|_2$

定理：有限维线性空间中任意两个向量范数都等价。

证明思路　1)范数等价为等价关系,满足传递性;

2)任意范数为坐标函数的连续函数；

3)在单位超球面上有大于零的极大极小值,

与2-范数等价。

定义：若$\left\{x^{(k)}\right\}(k=1,2,\cdots)$ 是线性空间V^n 中的向量序列，如果存在 $\forall x\in V^n$ ，使得 $\lim\limits_{k\to+\infty}\left\|x^{(k)}-x\right\|_\alpha=0$

则称序列 $\left\{x^{(k)}\right\}$ 按 $\alpha-$ 范数收敛于 x

定理：向量空间C^n 中，

$$\lim_{k\to+\infty}x^{(k)}=x\Leftrightarrow\forall\|x\|,\lim_{k\to+\infty}\left\|x^{(k)}-x\right\|=0$$

定理：向量空间C^n 中，

$$\lim_{k\to+\infty} x^{(k)} = x \Leftrightarrow \forall\|x\|, \lim_{k\to+\infty}\left\|x^{(k)} - x\right\| = 0$$

证明：只需对 $\|x\| = \|x\|_1$ 证明即可。

$$\begin{aligned}
x^k \to x \ &\Leftrightarrow \xi_i^{(k)} \to \xi_i (i = 1, 2, \cdots, n) \\
&\Leftrightarrow \left|\xi_i^{(k)} - \xi_i\right| \to 0 (i = 1, 2, \cdots, n) \\
&\Leftrightarrow \sum_{i=1}^{n}\left|\xi_i^{(k)} - \xi_i\right| \to 0 \\
&\Leftrightarrow \left\|x^{(k)} - x\right\|_1 \to 0
\end{aligned}$$

作业

- P121：4
- P122：5

本讲主要内容

- 矩阵范数
- 从属范数
- 范数的应用

定义：矩阵空间 $C^{m\times n}$ 中， $\forall A\in C^{m\times n}$ ，

定义实数值$\|A\|$，且满足以下条件

1)正定条件 ： $\|A\|\geq 0$， 且 $\|A\|=0\Leftrightarrow A=0_{m\times n}$

2)齐次条件： $\|kA\|=|k|\cdot\|A\|$， $\forall k\in K$

3)三角不等式 ：$\|A+B\|\leq\|A\|+\|B\|$， $B\in C^{m\times n}$

则称$\|A\|$为A的广义范数。

若对于$C^{m\times n}$,$C^{n\times l}$ 及 $C^{m\times n}$上的同类广义矩阵范数有

4)相容条件 ： $\|AB\|\leq\|A\|\|B\|$， $B\in C^{n\times l}$

则称$\|A\|$为A的范数。

定义：设 $C^{m\times n}$ 的矩阵函数 $\|A\|_M$, C^m 与 C^n 中的同类范数 $\|x\|_V$ ，若 $\|Ax\|_V \le \|A\|_M \|x\|_V$

则称矩阵范数 $\|A\|_M$ 与向量范数 $\|x\|_V$ 相容

例1、设 $A=\left(a_{ij}\right)_{m\times n}\in C^{m\times n},x=\left(\xi_1,\cdots,\xi_n\right)^T$ ，证明

$\|A\|_{m1}=\sum_{i,j=1}^{n}|a_{ij}|$ 是矩阵函数，且与 $\|x\|_1$ 相容

证明：（1）～（3）显然成立，

$$\|Ax\|_1=\sum_{i=1}^{m}\left|a_{i1}\xi_1+\cdots+a_{in}\xi_n\right|\le\sum_{i=1}^{m}\left(|a_{i1}||\xi_1|+\cdots+|a_{in}||\xi_n|\right)$$

$$\le\sum_{i=1}^{m}\left[\left(|a_{i1}|+\cdots+|a_{in}|\right)\left(|\xi_1|+\cdots+|\xi_n|\right)\right]$$

$$=\left[\sum_{i=1}^{m}\left(|a_{i1}|+\cdots+|a_{in}|\right)\right]\left(|\xi_1|+\cdots+|\xi_n|\right)=\|A\|_{m1}\|x\|_1$$

因此，$\|A\|_{m1}$ 与 $\|x\|_1$ 相容。

（4）证明相容性

划分 $B_{n\times l}=\left(b_1,\cdots b_l\right)$ ，则 $AB=\left(Ab_1,\cdots,Ab_l\right)$，且有

$$\begin{aligned}\|AB\|_{m1}&=\|Ab_1\|_1+\cdots+\|Ab_l\|_1\\&\le\|A\|_{m1}\|b_1\|_1+\cdots+\|A\|_{m1}\|b_l\|_1\\&=\|A\|_{m1}\left(\|b_1\|_1+\cdots+\|b_l\|_1\right)\\&=\|A\|_{m1}\|B\|_{m1}\end{aligned}$$

因此，$\|A\|_{m1}=\sum_{i,j=1}^{n}\left|a_{ij}\right|$ 是矩阵函数，且与 $\|x\|_1$ 相容

例2、设 $A=\left(a_{ij}\right)_{m\times n}\in C^{m\times n}, x=\left(\xi_1,\cdots,\xi_n\right)^T$ ，证明

$$\|A\|_{m\infty}=n\bullet\max_{i,j}|a_{ij}|$$ 是矩阵函数，且与 $\|x\|_\infty$ 相容

证明：（1）～（3）成立，设 $B=\left(b_{ij}\right)_{n\times l}$ ，则

$$\|AB\|_{m\infty}=l\bullet\max_{i,j}\left|\sum_{k=1}^n a_{ik}b_{kj}\right|\le l\bullet\max_{i,j}\left(\sum_{k=1}^n|a_{ik}||b_{kj}|\right)$$

$$\le l\bullet n\bullet\max_{i,j}\left(|a_{ik}||b_{kj}|\right)\le\left(n\bullet\max_{i,j}|a_{ij}|\right)\bullet\left(l\bullet\max_{i,j}|b_{ij}|\right)=\|A\|_{m\infty}\|B\|_{m\infty}$$

$$\|Ax\|_\infty=\max_i\left|\sum_{k=1}^n a_{ik}\xi_k\right|\le\max_i\sum_{k=1}^n|a_{ik}||\xi_k|$$

$$\le\max_i\sum_{k=1}^n|a_{ik}|\bullet\max_i|\xi_i|\le\left(n\bullet\max_i|a_{ik}|\right)\bullet\max_i|\xi_i|=\|A\|_{m\infty}\|x\|_\infty$$

例3、设 $A=\left(a_{ij}\right)_{m\times n}\in C^{m\times n}, x=\left(\xi_1,\cdots,\xi_n\right)^T$ ，证明

$$\|A\|_{m2}=\left(\sum_{i=1}^{n}\sum_{j=1}^{n}\left|a_{ij}\right|^2\right)^{\frac{1}{2}}$$ 是矩阵函数，且与 $\|x\|_2$ 相容

证明：（1）～（2）成立， $B=\left(b_{ij}\right)_{n\times l}$

设 $B_{m\times n}$,划分 $A=\left(a_1,\cdots,a_n\right), B=\left(b_1,\cdots,b_n\right)$ ，则有

$$\|A+B\|_{m2}^2=\|a_1+b_1\|_2^2+\cdots+\|a_n+b_n\|_2^2$$

$$\leq\left(\|a_1\|_2+\|b_1\|_2\right)^2+\cdots+\left(\|a_n\|_2+\|b_n\|_2\right)^2$$

$$\leq\|A\|_{m2}^2+2\left(\|a_1\|_2\|b_1\|_2+\cdots+\|a_n\|_2\|b_n\|_2\right)+\|B\|_{m2}^2$$

$$\leq\|A\|_{m2}^2+2\left(\sum\|a_i\|_2^2\right)^{\frac{1}{2}}\left(\sum\|b_i\|_2^2\right)^{\frac{1}{2}}+\|B\|_{m2}^2=\left(\|A\|_{m2}+\|B\|_{m2}\right)$$

设 $B_{n\times l}$， $AB=\left(\sum_{k=1}^{n}a_{ik}b_{kj}\right)_{m\times l}$ ，则有

$$\|AB\|_{m2}=\sum_{i,j=1}^{n}\left|\sum_{k=1}^{n}a_{ik}b_{kj}\right|^2 \le \sum_{i,j=1}^{n}\left(\sum_{k=1}^{n}|a_{ik}||b_{kj}|\right)^2$$

$$\le \sum_{i,j=1}^{n}\left[\left(\sum_{k=1}^{n}|a_{ik}|^2\right)\cdot\left(\sum_{k=1}^{n}|b_{kj}|^2\right)\right] \le \sum_{i=1}^{n}\left[\left(\sum_{k=1}^{n}|a_{ik}|^2\right)\cdot\sum_{j=1}^{n}\left(\sum_{k=1}^{n}|b_{kj}|^2\right)\right]$$

$$\le \left(\sum_{i,k=1}^{n}|a_{ik}|^2\right)\cdot\left(\sum_{k,j=1}^{n}|b_{kj}|^2\right) = \|A\|_{m2}\cdot\|B\|_{m2}$$

特别的，取 $B=x\in C^{n\times 1}$ ，则有

$$\|Ax\|_2=\|AB\|_{m2} \le \|A\|_{m2}\cdot\|B\|_{m2} = \|A\|_{m2}\cdot\|x\|_2$$

注：

1. $\|A\|_{m2}=\left[\mathrm{tr}\left(A^{\mathrm{H}}A\right)\right]^{\frac{1}{2}}=\left[\mathrm{tr}\left(AA^{\mathrm{H}}\right)\right]^{\frac{1}{2}}$ 称为矩阵的Frobenius范数，记做 $\|A\|_{\mathrm{F}}$

2. $C^{m\times n}$ 中的矩阵范数等价：对于任意的两种矩阵范数 $\|A\|_{\alpha}$ 和 $\|A\|_{\beta}$，存在 $0\le c_1\le c_2$，使得

$$c_1\|A\|_{\beta}\le\|A\|_{\alpha}\le c_2\|A\|_{\beta}\quad\forall A_{m\times n}$$

3. $C^{m\times n}$ 中，$\lim\limits_{k\to\infty}A^{(k)}=A\Leftrightarrow\forall\|A\|,\ \lim\limits_{k\to\infty}\left\|A^{(k)}-A\right\|=0$

定理3：设 $A_{m\times n}$ 及酉矩阵 $P_{m\times m}$ 和 $Q_{n\times n}$，有

$$\|PA\|_F = \|A\|_F = \|AQ\|_F$$

证明：

$$\|PA\|_F^2 = \mathrm{tr}\left((PA)^H(PA)\right) = \mathrm{tr}\left(A^H P^H PA\right)$$
$$= \mathrm{tr}\left(A^H A\right)$$

$$\|AQ\|_F^2 = \mathrm{tr}\left((AQ)^H(AQ)\right) = \mathrm{tr}\left(Q^H A^H AQ\right)$$

$$= \mathrm{tr}\left(AQQ^H A^H\right) = \mathrm{tr}\left(A^H A\right)$$

推论：酉（正交）相似的矩阵的F-范数相等

引理：对 $C^{m\times n}$ 的矩阵范数 $\|A\|$，存在向量范数 $\|x\|_V$

使得 $\|Ax\|_V \le \|A\|\cdot\|x\|_V$

例7：对$C^{m\times n}$的矩阵范数 $\|A\|$,任取非零列向量 $y\in C^n$

（1） $\|x\|_V = \|xy^H\|$ 是 C^n 的向量范数；

（2） $\|A\|$ 与 $\|x\|_V$ 相容。

证明：（1）略；（2）

$$\|Ax\|_V = \|(Ax)y^H\| = \|A(xy^H)\| \le \|A\|\cdot\|xy^H\| = \|A\|\cdot\|x\|_V$$

三、从属范数

定理：对 C^m 与 C^n 上的同类向量范数 $\|x\|_V$ ，定义

$$\|A\| = \max_{\|x\|_V=1}\|Ax\|_V \qquad \left(\forall A_{m\times n}, x\in C^n\right)$$

则 $\|A\|$ 是 $C^{m\times n}$ 中矩阵A 的范数，且 $\|A\|$ 与 $\|x\|_V$ 相容

$\|A\|$ 称为由 $\|x\|_V$ 导出的矩阵范数（或称为从属范数）

等价定义：

$$\max_{\|x\|_V=1}\|Ax\|_V = \max_{x\neq\theta}\frac{\|Ax\|_V}{\|x\|_V}$$

证明:(1) $A \neq 0: \exists x_0$ 满足 $\|x_0\|_V = 1, st\ Ax_0 \neq \theta$ 从而

$$\|A\| \geq \|Ax_0\|_V > 0$$

$$A = 0: \|A\| = \max_{\|x\|_V=1} \|Ax\|_V = \max_{\|x\|_V=1} \|0\|_V = 0$$

(2) 略

(3) 对 $A+B: \exists x_1, \|x_1\|_V = 1, st \max_{\|x\|_V=1} \|(A+B)x\|_V = \|(A+B)x_1\|_V$

$$\|A+B\| = \|Ax_1 + Bx_1\|_V \leq \|Ax_1\|_V + \|Bx_1\|_V \leq \|A\| + \|B\|$$

(4) 先证 $\|Ay\|_V \leq \|A\| \|y\|_V \quad (y \in C^n)$

$y = \theta$ ：显然成立。

$y \neq \theta$: 定义 $y_0 = \dfrac{y}{\|y\|_V}$ 满足

$$\|y_0\|_V = 1 \Rightarrow \|Ay_0\|_V \le \max_{\|y\|_V=1} \|Ay\|_V = \|A\|$$

故 $\|Ay\|_V = \left\|A\left(\|y\|_V\, y_0\right)\right\|_V = \|Ay_0\|_V \|y\|_V \le \|A\| \cdot \|y\|_V$

对 AB: $\exists x_2, \|x_2\|_V = 1, \mathrm{st} \max\limits_{\|x\|_V=1} \|(A+B)x\|_V = \|(A+B)x_2\|_V$

$$\|AB\| = \|(AB)x_2\|_V = \|A(Bx_2)\|_V \le \|A\| \cdot \|Bx_2\|_V \le \|A\| \cdot \|B\|$$

故定理成立。

注：

（1）一般的矩阵范数：$\because I = I \cdot I$

$$\|I\| \le \|I\| \cdot \|I\| \qquad \therefore \|I\| \ge 1$$

例如：$\|I\|_{m1} = n, \quad \|I\|_F = \sqrt{n}$

（2）矩阵的从属范数：$\|I\| = \max\limits_{\|x\|_V = 1} \|Ix\|_V = 1$

（3）常用的从属范数：

$$\|x\|_V \qquad \|x\|_1 \qquad \|x\|_2 \qquad \|x\|_\infty$$

$$\|A\|_M \qquad \|A\|_1 \qquad \|A\|_2 \qquad \|A\|_\infty$$

定理：设 $A=\left(a_{ij}\right)_{m\times n}$ ，则

（1）列和范数： $\|A\|_1=\max\limits_{j}\left\{\sum\limits_{i=1}^{m}|a_{ij}|\right\}$

（2）谱范数： $\|A\|_2=\sqrt{\lambda_1},\lambda_1=\max\left\{\lambda\left(A^HA\right)\right\}$

（3）行和范数： $\|A\|_\infty=\max\limits_{i}\left\{\sum\limits_{j=1}^{m}|a_{ij}|\right\}$

先证列和范数：$\|A\|_1 = \max_j \left\{ \sum_{i=1}^{m} |a_{ij}| \right\}$

证明：（1）记 $t = \max_j \left\{ \sum_{i=1}^{m} |a_{ij}| \right\}$，

如果 $x \in C^n$ 满足 $\|x\|_1 = 1$ ，则

$$\|Ax\|_1 = \sum_{i=1}^{m} \left| \sum_{j=1}^{n} a_{ij} \xi_j \right| \le \sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}| \cdot |\xi_j| = \sum_{j=1}^{n} \left[|\xi_j| \cdot \left(\sum_{i=1}^{m} |a_{ij}| \right) \right]$$

$$\le \sum_{j=1}^{n} \left\lfloor |\xi_j| t \right\rfloor \le t \cdot \|x\|_1 = t \qquad \therefore \|A\|_1 \le t$$

选k 使得 $t=\max_j\left\{\sum_{i=1}^m |a_{ij}|\right\}=\sum_{i=1}^m |a_{ik}|$ ，令

$e_k=(0,\cdots,0,1,0,\cdots,0)^T$ 则 $\|e_k\|_1=1$ ，而且

$$\|A\|_1\geq\|Ae_k\|_1=\sum_{i=1}^m |a_{ik}| =\max_j\left\{\sum_{i=1}^m |a_{ij}|\right\}=t$$

所以 $\|A\|_1=\max_j\left\{\sum_{i=1}^m |a_{ij}|\right\}$

范数的应用

定理6：设 $A\in C^{n\times n}$,且对$C^{n\times n}$的矩阵范数 $\|\bullet\|$ ，

满足 $\|A\|<1$,则矩阵$I-A$非奇异，且有

$$\left\|(I-A)^{-1}\right\|\le\frac{\|I\|}{1-\|A\|}$$

证明：选取向量范数 $\|x\|_v$，使得$\|A\|$与$\|x\|_v$相容。

如果 $\det(I-A)=0$,则 $(I-A)x=0$ 有非零解 x_0

$$x_0=Ax_0\Rightarrow\|x_0\|_V=\|Ax_0\|_V\le\|A\|\bullet\|x_0\|_V<\|x_0\|_V$$

产生了矛盾，故 $\det(I-A)\ne0$ ， $I-A$可逆。

证明：选取向量范数 $\|x\|_v$，使得 $\|A\|$ 与 $\|x\|_v$ 相容。

如果 $\det(I-A)=0$,则 $(I-A)x=0$ 有非零解 x_0

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

$$\Rightarrow (I-A)^{-1}=I+(I-A)^{-1}A$$

$$\Rightarrow \left\|(I-A)^{-1}\right\| \le \|I\|+\left\|(I-A)^{-1}A\right\|$$

$$\Rightarrow \left\|(I-A)^{-1}\right\| \le \|I\|+\left\|(I-A)^{-1}\right\|\|A\|$$

$$\Rightarrow \left\|(I-A)^{-1}\right\| \le \frac{\|I\|}{1-\|A\|}$$

定理7：设 $A \in C^{n\times n}$,且对$C^{n\times n}$的矩阵范数 $\|\bullet\|$ ，满足 $\|A\|<1$,则矩阵$I-A$非奇异，且有

$$\left\|I-(I-A)^{-1}\right\| \le \frac{\|I\|}{1-\|A\|}$$

证明：已证 $\|A\|<1$，则 $I-A$可逆。

恒等式： $(I-A)-I=-A$

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

左乘 A ： $A-A(I-A)^{-1}=-A^2(I-A)^{-1}$

左乘 A ：$A - A(I-A)^{-1} = -A^2(I-A)^{-1}$

$$\Rightarrow \left\| A(I-A)^{-1} \right\| \le \|A\| + \|A\| \cdot \left\| A(I-A)^{-1} \right\|$$

$$\Rightarrow \left\| A(I-A)^{-1} \right\| \le \frac{\|A\|}{1-\|A\|}$$

$$\Rightarrow \left\| I-(I-A)^{-1} \right\| = \left\| -A(I-A)^{-1} \right\| \le \frac{\|A\|}{1-\|A\|}$$

定理8：设$A,B\in C^{n\times n}$，A可逆,且满足 $\left\|A^{-1}B\right\|<1$

(1) $A+B$ 可逆

(2) $F=I-\left(I-A^{-1}B\right)^{-1}$：$\|F\|\le\dfrac{\left\|A^{-1}B\right\|}{1-\left\|A^{-1}B\right\|}$

(3) $$\frac{\left\|A^{-1}-\left(A+B\right)^{-1}\right\|}{\left\|A^{-1}\right\|}\le\frac{\left\|A^{-1}B\right\|}{1-\left\|A^{-1}B\right\|}$$

证明：利用定理6和定理7可证明

矩阵条件数: $\mathrm{cond}(A)=\|A\|\cdot\|A^{-1}\|$

谱半径:对 $\forall A\in C^{n\times n}$, 谱半径为 $\rho(A)\in\max_i|\lambda_i|$

定理9: $\forall A\in C^{n\times n},\forall\|\cdot\|_M$ ， 有 $\rho(A)\le\|A\|_M$

证明：对矩阵范数 $\|\cdot\|_M$ ， 存在向量范数 $\|\cdot\|_V$ ，

使得 $\|Ax\|_V\le\|A\|_M\cdot\|x\|_V$

设 $Ax_i=\lambda_i x_i \quad (x_i\ne\theta)$ ， 则有 $\|A\|_M$

$$|\lambda_i|\cdot\|x_i\|_V=\|\lambda_i x_i\|_V=\|Ax_i\|_V\le\|A\|_M\cdot\|x_i\|_V$$

$$|\lambda_i|\le\|A\|_M \quad\Rightarrow\rho(A)\le\|A\|_M$$

定理10: $\forall \boldsymbol{A} \in \boldsymbol{C}^{n\times n}$，对$\forall \varepsilon > \mathbf{0}, \exists$ 矩阵范数$\|\bullet\|_M$

使得 $\|\boldsymbol{A}\|_M \le \rho(\boldsymbol{A}) + \varepsilon$

证明：根据Jordan标准型理论：存在可逆矩阵$\boldsymbol{P}_{n\times n}$

使得 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{J} = \Lambda + \tilde{\boldsymbol{I}}$

$$\Lambda = \mathbf{diag}(\lambda_1, \cdots, \lambda_n), \tilde{\boldsymbol{I}} = \mathbf{diag}([\delta_1, \cdots, \delta_{n-1}], \mathbf{1})$$

其中 δ_i 等于0或1，于是有

$$\boldsymbol{D} = \mathbf{diag}(\mathbf{1}, \varepsilon, \cdots, \varepsilon^{n-1})$$

$$(\boldsymbol{PD})^{-1}\boldsymbol{APD} = \boldsymbol{D}^{-1}\boldsymbol{JD} \quad = \Lambda + \varepsilon\tilde{\boldsymbol{I}}$$

令$S=PD$可逆，那么

$$\left\|S^{-1}AS\right\|_1=\left\|\Lambda+\varepsilon\tilde{I}\right\|_1\le\rho(A)+\varepsilon$$

可证 $\|B\|_M=\left\|S^{-1}AS\right\|_1 \quad \left(B\in C^{n\times n}\right)$ 是$C^{n\times n}$中的矩阵范数

于是有

$$\|A\|_M=\left\|S^{-1}AS\right\|_1\le\rho(A)+\varepsilon$$

讨论:

(1) $\|A\|_{M1},\|A\|_{M2}$ 可与同一种 $\|x\|_V$ 相容？

(2) $\|A\|_M$ 可与不同的 $\|x\|_{V1},\|x\|_{V2}$ 相容？

(3) $\forall\|A\|_M$ 与 $\forall\|x\|_V$ 不一定相容？

分析: (1) $\|A\|_{M1},\|A\|_1$ 与 $\|x\|_1$ 相容

(2) $\|A\|_{M1}$ 与 $\|x\|_p\,(p\geq 1)$ 相容

$$x=(\xi_1,\cdots,\xi_n)^T, E_{ij}x=(0,\cdots,0,\xi_n,0,\cdots,)^T \Rightarrow \|E_{ij}x\|_p\leq\|x\|_p$$

$$Ax = \sum_{i,j} a_{ij} E_{ij} x$$

$$\|Ax\|_p = \sum_{i,j} |a_{ij}| \|E_{ij} x\|_p \quad \le \left(\sum_{i,j} |a_{ij}| \right) \cdot \|x\|_p \quad = \|A\|_{M1} \cdot \|x\|_p$$

(3) $\|A\|_1 = \max_j \left(\sum_i |a_{ij}| \right)$ 与 $\|x\|_\infty = \max_i (|\xi_i|)$ 不相容

$$n > 1 : A_0 = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 0 & 0 & & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}_{m\times n} \quad x_0 = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}_{n\times 1}, A_0 x_0 = \begin{bmatrix} n \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$\|A_0\|_1 = 1, \qquad \|x_0\|_\infty = 1, \qquad \|A_0 x_0\|_\infty = n$$

$$\|A_0 x_0\|_\infty = n > 1 = \|A_0\|_1 \cdot \|x_0\|_\infty$$

构造方法:

(1) 由向量范数构造新的向量范数

$S_{m\times n}$ 列满秩, $\|x\|=\|Sx\|_V$ 是 C^n 中的向量范数

(2) 由矩阵范数构造向量范数

非零列向量 $y_0\in C^n, \|x\|=\|xy_0^T\|_M$ 是 C^m 中的向量范数

(3) 由向量范数构造矩阵范数

$\|A\|=\max\limits_{\|x\|_V=1}\|Ax\|_V$ 是 $C^{m\times n}$ 中的矩阵范数

(4) 由矩阵范数构造新的矩阵范数

$S_{n\times n}$ 可逆, $\|A\|=\|S^{-1}AS\|_M$ 是 $C^{n\times n}$ 中的矩阵范数

作业

- P132：1、4

本讲主要内容

- 矩阵序列
- 矩阵级数
- 矩阵函数

引言：

- 一元多项式 $f(t)=c_0+c_1t+\cdots+c_mt^m$

- 矩阵多项式 $f(A)=c_0+c_1A+\cdots+c_mA^m, \quad (\forall A\in C^{n\times n})$

- $f(A)$以矩阵为自变量且取值为矩阵的一类函数

本章研究一般的以矩阵为自变量且取值为矩阵的函数

——矩阵函数

一、敛散性

定义：将矩阵序列 $A^{(k)}=\left(a_{ij}^{(k)}\right)_{m\times n}$,记作$\left\{A^{(k)}\right\}$

当 $\lim\limits_{k\to\infty} a_{ij}^{(k)}=a_{ij}(\forall i,j)$ 时，称矩阵序列$\{A^{(k)}\}$收敛于

矩阵$A=(a_{ij})$。记作

$$\lim_{k\to\infty} A^{(k)}=A \quad \text{，或者} \quad A^{(k)}\to A\left(k\to\infty\right)$$

若数列 $\left(a_{ij}^{(k)}\right)$ 之一发散，称 $\left\{A^{(k)}\right\}$ 发散

性质：

(1) 若 $\lim_{k\to\infty} A^{(k)} = A_{m\times n}, \lim_{k\to\infty} B^{(k)} = B_{m\times n}$ 则

$$\lim_{k\to\infty}(aA^{(k)} + bB^{(k)}) = aA + bB, \quad \forall a, b$$

(2) 若 $\lim_{k\to\infty} A^{(k)} = A_{m\times n}, \lim_{k\to\infty} B^{(k)} = B_{n\times l}$ 则

$$\lim_{k\to\infty}(A^{(k)}B^{(k)}) = AB$$

(3) 若 $A^{(k)}$ 与 A 是可逆矩阵，且 $\lim_{k\to\infty} A^{(k)} = A$ ，则

$$\lim_{k\to\infty}(A^{(k)})^{-1} = A^{-1}$$

定理1：设 $A^{(k)}, A \in C^{m\times n}$，则

(1) $\lim\limits_{k\to\infty} A^{(k)} = 0 \iff \forall \|\bullet\|, \lim\limits_{k\to\infty} \left\|A^{(k)}\right\| = 0$

(2) $\lim\limits_{k\to\infty} A^{(k)} = A \iff \forall \|\bullet\|, \lim\limits_{k\to\infty} \| A^{(k)} - A \| = 0$

证明：(1)考虑F -矩阵范数

$$\lim_{k\to\infty} A^{(k)} = 0 \iff \lim_{k\to\infty} a_{ij}^{(k)} = 0 \quad (all\ i, j)$$

$$\iff \lim_{k\to\infty} \sum_{i=1}^{m}\sum_{j=1}^{n} \left|a_{ij}^{(k)}\right|^2 = 0$$

$$\iff \lim_{k\to\infty} \left\|A^{(k)}\right\|_F = 0$$

(2)由 $\lim\limits_{k\to\infty} A^{(k)} = A \Leftrightarrow \lim\limits_{k\to\infty}(A^{(k)} - A) = 0$ 可直接的得到

敛散性的另一定义：矩阵序列 $\{\boldsymbol{A}^{(k)}\}$收敛的充要条件

为对任给$\varepsilon>0$ 存在$N(\varepsilon)$, 当 $k, l \geq N(\varepsilon)$ 时有

$$\left\|\boldsymbol{A}^{(k)}-\boldsymbol{A}^{(l)}\right\|<\varepsilon$$

其中 $\|\cdot\|$ 为任意的广义矩阵范数。

例1： $\mathbf{A}^{(n)}=\begin{pmatrix} 1-\frac{1}{n} & \sin(\frac{1}{n}) \\ e^{-n} & \sum_{k=1}^{n}\frac{\sin(k)}{k^2} \end{pmatrix}$ ，证明其敛散性

因为求不出 $\mathbf{A}^{(n)}$ 的极限从而很难应用定义证明收敛。

相反，由于 $\left|\sum_{k=m+1}^{n}\frac{\sin(k)}{k^2}\right| \le \left|\sum_{k=m+1}^{n}\frac{1}{k^2}\right| \le \left|\sum_{k=m+1}^{n}\frac{1}{k(k-1)}\right| \le \frac{1}{m}$

从而只要取l充分大，则当$m, n > l$ 时就有

$$\left|\sum_{k=m+1}^{n}\frac{\sin(k)}{k^2}\right| \le \varepsilon$$

这样 $\mathbf{A}^{(n)}$ 收敛

定义：若 $A_{n\times n}$ 满足 $\lim\limits_{k\to\infty} A^k = 0_{n\times n}$，称$A$为收敛矩阵

定理2：A为收敛矩阵 $\Longleftrightarrow$ $\rho(A)<1$

证明：充分性。已知 $\rho(A)<1$，对 $\varepsilon=\frac{1}{2}\left[1-\rho(A)\right]>0$

存在矩阵范数 $\|\cdot\|_M$，使得

$$\|A\|_M \le \rho(A)+\varepsilon = \frac{1}{2}\left[1+\rho(A)\right]<1$$

于是有 $\left\|A^k\right\|_M \le \|A\|_M^k \to 0$，故由定理1可得 $A^k \to 0$

必要性：已知 $A^k \to 0$，设$Ax=\lambda x(x\ne 0)$，则有

$$\lambda^k x = A^k x \to 0 \quad \Rightarrow \lambda^k \to 0 \quad \Rightarrow |\lambda|<1$$

故 $\rho(A)<1$

定理3：若矩阵范数 $\|\bullet\|_M$ 使 $\|A\|_M < 1$ ，则 $A^k \to 0$

证明： $\rho(A) \le \|A\|_M < 1 \quad \Rightarrow A^k \to 0$

例： $A = \begin{bmatrix} 0.1 & 0.3 \\ 0.7 & 0.6 \end{bmatrix}$

$$\|A\|_1 = 0.9 < 1 \Rightarrow A^k \to 0$$

定义：设有矩阵序列$\{A^{(k)}\}$，其中$A^{(k)}=(a_{ij}^{(k)})\in C^{m\times n}$

称$A^{(0)}+A^{(1)}+\cdots+A^{(k)}+\cdots$为矩阵级数。记为$\sum_{k=0}^{\infty}A^{(k)}$

部分和 $S^{(N)}=\sum_{k=0}^{N}A^{(k)}$ 构成矩阵序列 $\{S^{(N)}\}$

敛散性：若$\lim_{N\to\infty}S^{(N)}=S$，称$\sum A^{(k)}$ 收敛于S，记做

$$\sum A^{(k)}=S$$

若 $\{S^{(N)}\}$ 发散，称 $\sum A^{(k)}$ 发散

性质1：$\sum A^{(k)} = S \Longleftrightarrow \sum_{k=0}^{\infty} a_{ij}^{(k)} = s_{ij}\ (all\quad i, j)$

证明：左 $\Longleftrightarrow \lim\limits_{N\to\infty} S^{(N)} = S$

$\Longleftrightarrow \lim\limits_{N\to\infty} s_{ij}^{(N)} = s_{ij}\ (all\quad i, j)$

$\Longleftrightarrow \lim\limits_{N\to\infty} \sum_{k=0}^{N} a_{ij}^{(k)} = s_{ij}\ (all\quad i, j)$

$\Longleftrightarrow$ 右

性质2：若 $\sum\left|a_{ij}^{(k)}\right|$ 收敛 $(all \quad i,j)$，称 $\sum A^{(k)}$ 绝对收敛。

（1） $\sum A^{(k)}$ 绝对收敛 $\Rightarrow \sum A^{(k)}$ 收敛

（2） 若 $\sum A^{(k)}$ 绝对收敛于S，对 $\sum A^{(k)}$ 任意重组重排得 $\sum B^{(k)}$，则 $\sum B^{(k)}$ 绝对收敛于S。

性质3： $\sum A^{(k)} = S$ 绝对收敛 $\Leftrightarrow \forall \|\cdot\|, \sum \|A^{(k)}\|$ 收敛

证明：只需考虑矩阵范数 $\|\cdot\|_{m1}$

证明：（必要性）由于矩阵级数 $\sum_{k=0}^{\infty} A^{(k)}$ 绝对收敛，

因此mn个数项级数 $\sum_{k=0}^{\infty} a_{ij}^{(k)}$ 绝对收敛。因此存在

$M > 0$，使得对任意N，都有 $\sum_{k=0}^{N} \left| a_{ij}^{(k)} \right| < M, (\forall i, j)$

故 $$\sum_{k=0}^{N} \left\| A^{(k)} \right\|_{m1} = \sum_{k=0}^{N} \left(\sum_{i,j} | a_{ij}^{(k)} | \right) \le mnM$$

因此 $\sum_{k=0}^{\infty} \| A^{(k)} \|_{m1}$ 收敛

(充分性) $\sum_{k=0}^{\infty} \left\| A^{(k)} \right\|_{m1}$ 收敛。因此 $\sum_{k=0}^{\infty} | a_{ij}^{(k)} | \le \sum_{k=0}^{\infty} \left\| A^{(k)} \right\|_{m1}$

故矩阵级数 $\sum_{k=0}^{\infty} A^{(k)}$ 绝对收敛

性质4：$\sum A^{(k)}$ 收敛于S $\Rightarrow \sum_{k=0}^{\infty} PA^{(k)}Q$ 收敛于 PSQ

$$\sum A^{(k)} \text{ 绝对收敛} \Rightarrow \sum_{k=0}^{\infty} PA^{(k)}Q \text{ 绝对收敛}$$

证明：只需考虑矩阵范数 $\|\bullet\|_{m1}$

证明：(1) $S^{(N)} = \sum_{k=0}^{N} A^{(k)} \to S \Rightarrow \sum_{k=0}^{N} PA^{(k)}Q = PS^{(N)}Q \to PSQ$

$$\Rightarrow \sum_{k=0}^{\infty} PA^{(k)}Q \to PSQ$$

性质4：

$$\sum A^{(k)} \text{ 绝对收敛} \Rightarrow \sum_{k=0}^{\infty} PA^{(k)}Q \text{ 绝对收敛}$$

(2)矩阵范数 $\|\cdot\|$，由性质3知 $\sum \left\|A^{(k)}\right\|$ 收敛

因为 $\left\|PA^{(k)}Q\right\| \leq \|P\| \left\|A^{(k)}\right\| \|Q\| = M\left\|A^{(k)}\right\| \qquad \left(M = \|P\|\|Q\|\right)$

所以 $\sum_{k=0}^{N}\left\|PA^{(k)}Q\right\| \leq \sum_{k=0}^{N}\left(M\left\|A^{(k)}\right\|\right) = M\sum_{k=0}^{N}\left\|A^{(k)}\right\|$ 有界

故 $\sum A^{(k)}$ 绝对收敛 $\Rightarrow \sum_{k=0}^{\infty} PA^{(k)}Q$ 绝对收敛

性质5：$\sum_{k=1}^{\infty} A^{(k)}$ 绝对收敛于 $S_{m\times n}$，$\sum_{k=1}^{\infty} B^{(k)}$ 绝对收敛于 $T_{n\times l}$

则Cauchy积

$$A^{(1)}B^{(1)} + \left[A^{(1)}B^{(2)} + A^{(2)}B^{(1)}\right] + \left[A^{(1)}B^{(3)} + A^{(2)}B^{(2)} + A^{(3)}B^{(1)}\right]$$
$$+\cdots+\left[A^{(1)}B^{(k)} + A^{(2)}B^{(k-1)} + \cdots + A^{(k)}B^{(1)}\right] + \cdots$$

绝对收敛于ST，记作 $\sum A^{(k)} \bullet \sum B^{(k)} = ST$

证明：只需利用性质3即可

Neumann级数： $A_{n\times n}, \sum_{k=1}^{\infty} A^k, \left(A^0 = I\right)$

定理4： $A_{n\times n}, \sum A^k$ 收敛 $\Leftrightarrow A^k \to 0$

$\sum A^k$ 收敛时，其和为 $(I-A)^{-1}$

证明：必要性。$\sum A^k$ 收敛时 $\sum\left(A^k\right)_{ij}, \forall i,j$ 收敛

即 $\left(A^k\right)_{ij} \to 0$ ，也就是 $A^k \to 0$

充分性。$A^k \to 0$ 由定理2可知 $\rho(A) < 1 \Leftrightarrow (I-A)$ 可逆

$$\left(I+A+A^2+\cdots+A^N\right)(I-A) = I - A^{N+1}$$

$$\left(I+A+A^2+\cdots+A^N\right) = (I-A)^{-1} - A^{N+1}(I-A)^{-1} \to (I-A)^{-1}, N \to \infty$$

即 $\sum A^k = (I-A)^{-1}$

定理5：$A_{n\times n},\|A\|<1\Rightarrow\left\|(I-A)^{-1}-\sum_{k=0}^{N}A^k\right\|\le\frac{\|A\|^{N+1}}{1-\|A\|},N=0,1,2$

证明：$\|A\|<1\Rightarrow\rho(A)<1\quad\Rightarrow(I-A)$ 可逆

$$\left(I+A+A^2+\cdots+A^N\right)(I-A)=I-A^{N+1}$$

右乘 $(I-A)^{-1}$，移项可得

$$(I-A)^{-1}-\left(I+A+A^2+\cdots+A^N\right)=A^{N+1}(I-A)^{-1}$$

恒等式 $A^{N+1}=A^{N+1}(I-A)^{-1}(I-A)=A^{N+1}(I-A)^{-1}-A^{N+1}(I-A)^{-1}A$

$$A^{N+1}(I-A)^{-1}=A^{N+1}+A^{N+1}(I-A)^{-1}A$$

$$\left\|A^{N+1}(I-A)^{-1}\right\|\le\left\|A^{N+1}\right\|+\left\|A^{N+1}(I-A)^{-1}\right\|\bullet\|A\|$$

故 $\left\|A^{N+1}(I-A)^{-1}\right\|\le\frac{\left\|A^{N+1}\right\|}{1-\|A\|}\Rightarrow\left\|(I-A)^{-1}-\sum_{k=0}^{N}A^k\right\|\le\frac{\|A\|^{N+1}}{1-\|A\|}$

幂级数： 对函数 $f(z)=\sum c_k z^k, \left(|z|<r\right)$ 方阵 $A_{n\times n}$,

构造矩阵幂级数 $f(A)=\sum c_k A^k$

定理6： (1) $\rho(A)<r \quad \Rightarrow \sum c_k A^k$ 绝对收敛

(2) $\rho(A)>r \quad \Rightarrow \sum c_k A^k$ 发散

证明:对A，取 $\varepsilon=\frac{1}{2}[r-\rho(A)]>0, \exists\|\bullet\|_{\varepsilon}$ ，使得

$$\|A\|_{\varepsilon}\leq\rho(A)+\varepsilon=\frac{1}{2}[r+\rho(A)]<r$$

$$\left\|c_k A^k\right\|_{\varepsilon}\leq|c_k|\|A\|_{\varepsilon}^k\leq|c_k|[\rho(A)+\varepsilon]^k$$

当 $|z|<r$ 时，$\sum|c_k||z|^k$ 收敛，于是

$\sum|c_k|[\rho(A)+\varepsilon]^k$ 收敛 $\Rightarrow\sum\left\|c_k A^k\right\|_{\varepsilon}$ 收敛 $\Rightarrow\sum c_k A^k$ 收敛

设A的特征值λ满足|λ|=ρ(A)，x为λ相应的特征向量

$$\sum_{k=0}^{n}c_k(A^k x)=\sum_{k=0}^{n}c_k(\lambda^k x)=\left(\sum_{k=0}^{n}c_k\lambda^k\right)x$$

由于ρ(A) > r ,那么 $\left(\sum_{k=0}^{n}c_k\lambda^k\right)x$ 发散(注意x为非零向量)

从而 $\sum_{k=0}^{n}c_k(A^k x)$ 发散，这样 $\sum_{k=0}^{\infty}c_k A^k$ 发散

定义：设一元函数 $f(z)$ 能展开为z的幂级数

$$f(z)=\sum_{k=0}^{\infty} c_k z^k \quad \left(|z|<r, r>0\right)$$

其中r>0表示该幂级数的收敛半径。当n阶矩阵A的谱半径$\rho(A)<r$ 时，把收敛的矩阵幂级数$\sum_{k=0}^{\infty} c_k A^k$ 的和为f(A), 即 $f(A)=\sum_{k=0}^{\infty} c_k A^k$

性质(代入规则)：若$f(z)=g(z)$, 则$f(A)=g(A)$.

例1：

$$e^{z}=1+\frac{1}{1!}z+\cdots+\frac{1}{k!}z^{k}+\cdots\quad\left(r=+\infty\right)$$

$$e^{A}=I+\frac{1}{1!}A+\cdots+\frac{1}{k!}A^{k}+\cdots\quad\left(\forall A_{n\times n}\right)$$

$$\sin z=z-\frac{1}{3!}z^{3}+\cdots+\left(-1\right)^{k}\frac{1}{\left(2k+1\right)!}z^{\left(2k+1\right)}+\cdots\quad\left(r=+\infty\right)$$

$$\sin A=A-\frac{1}{3!}A^{3}+\cdots+\left(-1\right)^{k}\frac{1}{\left(2k+1\right)!}A^{\left(2k+1\right)}+\cdots\quad\left(\forall A_{n\times n}\right)$$

例2：$f(z)=\dfrac{1}{1-z}=\sum_{k=0}^{\infty} z^k \quad (|z|<1)$，$\quad f(A)=\dfrac{1}{1-A}=\sum_{k=0}^{\infty} A^k \quad (\rho(A)<1)$

例3： $\forall A_{n\times n}, e^{jA}=\cos A+j\sin A \quad \left(j=\sqrt{-1}\right)$

$$\cos A=\frac{1}{2}\left(e^{jA}+e^{-jA}\right), \qquad \cos(-A)=\cos A$$

$$\sin A=\frac{1}{2j}\left(e^{jA}-e^{-jA}\right), \qquad \sin(-A)=\sin A$$

证明：在 e^{jA} 中,视“jA”为整体,并按奇偶次幂分开

$$e^{jA}=\left[I+\frac{1}{2!}(jA)^2+\frac{1}{4!}(jA)^4+\cdots\right]+\left[\frac{1}{1!}(jA)+\frac{1}{3!}(jA)^3+\cdots\right]$$

$$=\cos A+j\sin A \quad \left(j=\sqrt{-1}\right)$$

例4： $A=\begin{bmatrix}1 & 1\\ 0 & 0\end{bmatrix}, B=\begin{bmatrix}1 & -1\\ 0 & 0\end{bmatrix}$, 求 e^A, e^B 及 e^{A+B}

解：$A^2 = A$： $e^A = I + \left(\frac{1}{1!}+\cdots+\frac{1}{k!}+\cdots\right)A = I+(e-1)A$

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

$B^2 = B$： $e^B = I + \left(\frac{1}{1!}+\cdots+\frac{1}{k!}+\cdots\right)B = I+(e-1)B$

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

$$A+B=\begin{bmatrix}2 & 0\\ 0 & 0\end{bmatrix}=2\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix} \qquad (A+B)^k = 2^k\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}$$

$$e^{A+B} = I+\left(\frac{2}{1!}+\frac{2^2}{2!}+\cdots\right)\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix} = I+(e^2-1)\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix} = \begin{bmatrix}e^2 & 0\\ 0 & 1\end{bmatrix}$$

注意：

$$e^{A+B}=\begin{bmatrix} e^2 & 0 \\ 0 & 1 \end{bmatrix}$$

$$e^{A}e^{B}=\begin{bmatrix} e^2 & -(e-1)^2 \\ 0 & 1 \end{bmatrix},\quad e^{B}e^{A}=\begin{bmatrix} e^2 & (e-1)^2 \\ 0 & 1 \end{bmatrix}$$

$$e^{A+B}\neq e^{A}e^{B}\neq e^{B}e^{A}$$

定理7： $A_{n\times n}, B_{n\times n}, AB=BA \Rightarrow e^{A+B}=e^{AB}=e^{BA}$

证明：$e^{A}e^{B}=\left[I+\frac{1}{1!}A+\frac{1}{2!}A^{2}+\cdots\right]\left[I+\frac{1}{1!}B+\frac{1}{2!}B^{2}+\cdots\right]$

$$=I+\left(A+B\right)+\frac{1}{2!}\left(A^{2}+2AB+B^{2}\right)+\frac{1}{3!}\left(A^{3}+3AB^{2}+3A^{2}B^{2}+B^{3}\right)+\cdots$$

$$=I+\left(A+B\right)+\frac{1}{2!}\left(A+B\right)^{2}+\frac{1}{3!}\left(A+B\right)^{3}+\cdots$$

$$=e^{A+B}$$

同理： $e^{B}e^{A}=e^{B+A}=e^{A+B}$

注： (1) $e^{A}e^{-A}=e^{o}=I\Rightarrow\left(e^{A}\right)^{-1}=e^{-A}\quad\forall A$

(2) $\left(e^{A}\right)^{m}=e^{mA}\quad m=2,3,\cdots$

例5： $A_{n\times n}, B_{n\times n}, AB=BA$

$$\cos(A+B)=\cos A\cos B-\sin A\sin B$$

$$\sin(A+B)=\sin A\cos B+\cos A\sin B$$

证明： $\cos A\cos B-\sin A\sin B=$

$$=\frac{1}{2}\left[e^{jA}+e^{-jA}\right]\bullet\frac{1}{2}\left[e^{jB}+e^{-jB}\right]-\frac{1}{2j}\left[e^{jA}-e^{-jA}\right]\bullet\frac{1}{2j}\left[e^{jB}-e^{-jB}\right]$$

$$=\frac{1}{4}\left[e^{j(A+B)}+\cdots+\cdots+e^{-j(A+B)}\right]+\frac{1}{4}\left[e^{j(A+B)}-\cdots-\cdots+e^{-j(A+B)}\right]$$

$$=\frac{1}{2}\left[e^{j(A+B)}+e^{-j(A+B)}\right]$$

$$=\cos(A+B)$$

矩阵函数值的求法

1.待定系数法:设n阶矩阵A 的特征多项式$\varphi(\lambda)=\det(\lambda I-A)$

如果首1多项式 $\psi(\lambda)=\lambda^m+b_1\lambda^{m-1}+\cdots+b_{m-1}\lambda+b_m\ (1\le m\le n)$

满足 $\psi(\lambda)|\varphi(\lambda)$ ，分解

$$\psi(\lambda)=(\lambda-\lambda_1)^{m_1}\cdots(\lambda-\lambda_s)^{m_s}\quad \lambda_i\ne\lambda_j,\sum m_i=m$$

因为 λ_i 是A的特征值，所以 $|\lambda_i|\le\rho(A)<r$ ，从而

$f(\lambda_i)=\sum c_k\lambda_i^k$ 绝对收敛。

设 $f(z)==\sum c_k z^k=\psi(\lambda)g(z)+r(z)$

$$r(z)=b_0+b_1z+\cdots+b_{m-1}z^{m-1}$$

由 $\psi(\lambda_i)=0,\psi^{(1)}(\lambda_i)=0,\cdots,\psi^{(m_i-1)}(\lambda_i)=0$ 可得

$$i=1,2,\cdots,s$$

$$r(\lambda_i)=f(\lambda_i)$$

$$r'(\lambda_i)=f'(\lambda_i)$$

$$\cdots$$

$$r^{(m_i-1)}(\lambda_i)=f^{(m_i-1)}(\lambda_i)$$

解此方程组得出 $b_0,b_1,\cdots,b_{m-1}$。因为 $\psi(A)=O$ 所以

$$f(A)=\sum c_k A^k=\psi(A)g(A)+r(A)\quad =r(A)$$

即 $\quad f(A)=b_0 I+b_1 A+\cdots+b_{m-1}A^{m-1}$

例6：$A=\begin{bmatrix}2&0&0\\1&1&1\\1&-1&3\end{bmatrix}$，　求　e^{A}, e^{tA}　$(t\in R)$

解：$\varphi(\lambda)=\det(\lambda I-A)=(\lambda-2)^{3}$

$$(A-2I)=\begin{bmatrix}0&0&0\\1&-1&1\\1&-1&1\end{bmatrix},(A-2I)^{2}=O$$

取　$\psi(\lambda)=m(\lambda)=(\lambda-2)^{2}$

(1)　$f(\lambda)=e^{\lambda}=\psi(\lambda)g(\lambda)+(a+b\lambda)$

$$f'(\lambda)=e^{\lambda}=[\psi(\lambda)g(\lambda)]'+b$$

$$\left.\begin{aligned}f(2)=e^{2}&:(a+2b)=e^{2}\\ f'(2)=e^{2}&:b=e^{2}\end{aligned}\right\}\Rightarrow\left\{\begin{aligned}a&=-e^{2}\\ b&=e^{2}\end{aligned}\right.$$

$$e^{A}=e^{2}(A-I)=e^{2}\begin{bmatrix}1&0&0\\1&0&1\\1&-1&2\end{bmatrix}$$

$$(2)\quad f(\lambda)=e^{t\lambda}=\psi(\lambda)g(\lambda)+(a+b\lambda)$$

$$f'(\lambda)=te^{t\lambda}=\left[\psi(\lambda)g(\lambda)\right]'+b$$

$$\left.\begin{array}{l} f(2)=e^{2t}:(a+2b)=e^{2t} \\ f'(2)=te^{2t}:b=te^{2t} \end{array}\right\} \Rightarrow \left\{\begin{array}{l} a=(1-2t)e^{2t} \\ b=te^{2t} \end{array}\right.$$

$$e^{tA}=e^{2t}\left[(1-2t)I+tA\right]=e^{2t}\begin{bmatrix} 1 & 0 & 0 \\ t & 1-t & t \\ t & -t & 1+t \end{bmatrix}$$

本讲主要内容

- 矩阵函数的值的计算（续）
- 矩阵函数的一般定义
- 矩阵函数的性质
- 矩阵的微分和积分

2. 数项级数求和法。

利用首一多项式 $\psi(\lambda)$ ，且满足$\psi(A)=0$，即

$$A^m+b_1A^{m-1}+\cdots+b_{m-1}A+b_mI=0$$

或者 $A^m=k_0^{(0)}I+k_1^{(0)}A+\cdots+k_{m-1}^{(0)}A^{m-1} \quad \left(k_i^{(0)}=-b_{m-i}\right)$

可以求出 $A^{m+1}=A^mA=k_0^{(1)}I+k_1^{(1)}A+\cdots+k_{m-1}^{(1)}A^{m-1}$

$$\vdots$$

$$A^{m+l}=k_0^{(l)}I+k_1^{(l)}A+\cdots+k_{m-1}^{(l)}A^{m-1}$$

$$\vdots$$

于是 $f(A)=\sum_{k=0}^{\infty}c_kA^k=\left(c_0I+c_1A+\cdots+c_{m-1}A^{m-1}\right)+$

$$c_m\left(k_0^{(0)}I+k_1^{(0)}A+\cdots+k_{m-1}^{(0)}A^{m-1}\right)+\cdots$$

$$=\left(c_0+\sum_{l=0}^{\infty}c_{m+l}k_0^{(l)}\right)I+\left(c_1+\sum_{l=0}^{\infty}c_{m+l}k_1^{(l)}\right)A+\cdots+\left(c_{m-1}+\sum_{l=0}^{\infty}c_{m+l}k_{m-1}^{(l)}\right)A^{m-1}$$

例7： $A=\begin{bmatrix} \pi & 0 & 0 & 0 \\ & -\pi & 0 & 0 \\ & & 0 & 1 \\ & & & 0 \end{bmatrix}$ ，求 $\sin A$

解：$\varphi(\lambda)=|\lambda I-A|=\lambda^4-\pi^2\lambda^2$，取 $\psi(\lambda)=\varphi(\lambda)$

$$\psi(A)=0 \Rightarrow A^4=\pi^2A^2,\ A^5=\pi^2A^3, A^7=\pi^4A^3,\cdots$$

$$\sin A=A-\frac{1}{3!}A^3+\frac{1}{5!}A^5-\frac{1}{7!}A^7+\cdots$$

$$=A+\left[-\frac{1}{3!}+\frac{\pi^2}{5!}-\frac{\pi^4}{7!}+\cdots\right]A^3$$

$$=A+\frac{1}{\pi^3}[\sin\pi-\pi]A^3=A-\frac{1}{\pi^2}A^3=\begin{bmatrix} 0 & 0 & 0 & 0 \\ & 0 & 0 & 0 \\ & & 0 & 1 \\ & & & 0 \end{bmatrix}$$

$$\because A^3=\text{diag}\left(\pi^3,-\pi^3,0,0\right)$$

3. 对角阵法

设 $P^{-1}AP=\mathrm{diag}(\lambda_1,\cdots,\lambda_n)=\Lambda$，则 $A^k=P\Lambda^kP^{-1}$,

且有 $\sum_{k=0}^{N}c_kA^k=P\sum_{k=0}^{N}c_k\Lambda^kP^{-1}$

$$=P\,\mathrm{diag}\left(\sum_{k=0}^{N}c_k\lambda_1^k,\cdots,\sum_{k=0}^{N}c_k\lambda_n^k\right)P^{-1}$$

于是

$$f(A)=\sum_{k=0}^{\infty}c_kA^k=P\bullet\mathrm{diag}\left(f(\lambda_1),\cdots,f(\lambda_n)\right)\bullet P^{-1}$$

例8： $P^{-1}AP=\Lambda$ ：

$$e^{A}=P\bullet\mathrm{diag}\left(e^{\lambda_1},\cdots,e^{\lambda_n}\right)\bullet P^{-1}$$

$$e^{tA}=P\bullet\mathrm{diag}\left(e^{\lambda_1 t},\cdots,e^{\lambda_n t}\right)\bullet P^{-1}$$

$$\sin A=P\bullet\mathrm{diag}\left(\sin\lambda_1,\cdots,\sin\lambda_n\right)\bullet P^{-1}$$

4.Jordan标准型法

设 $P^{-1}AP = J = \mathrm{diag}\left(J_1, \cdots, J_s\right),\ J_i = \lambda_1 I + I^{(1)}$

易证 $I^{(k)} I^{(1)} = I^{(1)} I^{(k)} = I^{(k+1)}, I^{(m_i)} = O$

$$k \le m_i - 1: J_i^k = \lambda_i^k I + C_k^1 \lambda_i^{k-1} I^{(1)} + \cdots + C_k^{k-1} \lambda_i I^{(k-1)} + I^{(k)}$$

$$k \ge m_i : J_i^k = \lambda_i^k I + C_k^1 \lambda_i^{k-1} I^{(1)} + \cdots + C_k^{m_i-1} \lambda_i^{k-m_i+1} I^{(m_i-1)}$$

$$f\left(J_i\right) = \sum_{k=0}^{\infty} c_k J_i^k = f\left(\lambda_i\right) I + \frac{f'\left(\lambda_i\right)}{1!} I^{(1)} + \cdots + \frac{f^{(m_i-1)}\left(\lambda_i\right)}{(m_i-1)!} I^{(m_i-1)}$$

$$f\left(A\right) = \sum_{k=0}^{\infty} c_k A^k = P \bullet \sum_{k=0}^{\infty} c_k J^k \bullet P^{-1} = P \bullet diag\left(f(J_1), \cdots, f(J_s)\right) \bullet P^{-}$$

三、矩阵函数的一般定义

展开式 $f(z)=\sum c_k z^k$， $\left(|z|<r, r>0\right)$，要求

（1） $f^{(k)}(0)$ 存在 $\left(k=0,1,2,\cdots\right)$

（2） $\lim\limits_{k\to\infty}\dfrac{f^{(k+1)}(\xi)}{(k+1)!}z^{k+1}=0 \quad \left(|z|<r\right)$

对于一元函数 $f(z)=\dfrac{1}{z}$ 等，还不能定义矩阵函数。

基于矩阵函数值的Jordan标准形算法，拓宽定义

矩阵函数的一般定义

设 $P^{-1}AP = J = \mathrm{diag}\left(J_1, \cdots, J_s\right)$, $J_i = \lambda_1 I + I^{(1)}$

如果 $f(z)$ 在 λ_i 处有 $m_i - 1$ 阶导数，令

$$f\left(J_i\right) = \sum_{k=0}^{\infty} c_k J_i^k = f\left(\lambda_i\right) I + \frac{f'\left(\lambda_i\right)}{1!} I^{(1)} + \cdots + \frac{f^{(m_i-1)}\left(\lambda_i\right)}{(m_i - 1)!} I^{(m_i-1)}$$

$$f(A) = \sum_{k=0}^{\infty} c_k A^k = P \bullet \sum_{k=0}^{\infty} c_k J^k \bullet P^{-1} = P \bullet diag\left(f(J_1), \cdots, f(J_s)\right) \bullet P^{-}$$

称 $f(A)$ 为对应于 $f(z)$ 的矩阵函数

[注] ① 拓宽定义不要求$f(z)$能展为“z”的幂级数，但要求在A 的特征值 λ_i （重数为 m_i）处有 m_i-1 阶导数,后者较前者弱！

② 当能够展为“z”的幂级数时, 矩阵函数的拓宽定义与级数原始定义是一致的.

例9：$A=\begin{bmatrix} 2 & 1 & & \\ & 2 & 1 & \\ & & 2 & 1 \\ & & & 2 \end{bmatrix}, f(z)=\dfrac{1}{z}$，求 $f(A)$

解：$f(z)=\dfrac{1}{z}, f'(z)=-z^{-2}, f''(z)=2z^{-3}, f'''(z)=-6z^{-4}$

$$f(A)=f(J)$$

$$=f(2)\bullet I+f'(2)\bullet I^{(1)}+\frac{f''(2)}{2!}\bullet I^{(2)}+\frac{f'''(2)}{3!}\bullet I^{(3)}$$

$$=\begin{bmatrix} 0.5 & -0.25 & 0.125 & -0.0625 \\ & 0.5 & -0.25 & 0.125 \\ & & 0.5 & -0.25 \\ & & & 0.5 \end{bmatrix}$$

例10：$A=\begin{bmatrix}1&1&0\\0&1&0\\0&0&2\end{bmatrix}$，$f(z)=\sqrt{z}$ ，求 $f(A)$

解： $f(z)=\sqrt{z}, f'(z)=\dfrac{1}{2\sqrt{z}}$

$$J_1=\begin{bmatrix}1&1\\&1\end{bmatrix}:\quad f(J_1)=f(1)\cdot I+f'(J_1)\cdot I^{(1)}=\begin{bmatrix}1&1/2\\&1\end{bmatrix}$$

$$J_2=[2]:\quad f(J_2)=f(2)\cdot I=\left[\sqrt{2}\right]$$

$$f(A)=f(J)=\begin{bmatrix}f(J_1)&\\&f(J_2)\end{bmatrix}=\begin{bmatrix}1&1/2&0\\&1&0\\&&\sqrt{2}\end{bmatrix}$$

四、矩阵函数的性质

级数定义或拓宽定义给出的矩阵函数具有下列性质：

(1) $f(z) = f_1(z) + f_2(z) \Rightarrow f(A) = f_1(A) + f_2(A)$

$$f^{(l)}(\lambda_i) = f_1^{(l)}(\lambda_i) + f_2^{(l)}(\lambda_i)$$

$$\Rightarrow f^{(l)}(J_i) = f_1^{(l)}(J_i) + f_2^{(l)}(J_i)$$

$$f(A) = P \bullet \left\{ \begin{bmatrix} f_1(J_1) & & \\ & \ddots & \\ & & f_1(J_s) \end{bmatrix} + \begin{bmatrix} f_2(J_1) & & \\ & \ddots & \\ & & f_2(J_s) \end{bmatrix} \right\} \bullet P^{-1}$$

$$= f_1(A) + f_2(A)$$

(2) $f(z) = f_1(z) \cdot f_2(z)$

$$\Rightarrow f(A) = f_1(A) \cdot f_2(A) = f_2(A) \cdot f_1(A)$$

$$f_1(J_i) \cdot f_2(J_i) = \left[f_1 \cdot I + f_1' \cdot I^{(1)} + \frac{f_1''}{2!} \cdot I^{(2)} + \cdots + \frac{f_1^{(m_i-1)}}{(m_i-1)!} \cdot I^{(m_i-1)} \right] \cdot$$

$$\left[f_2 \cdot I + \frac{f'}{1!} \cdot I^{(1)} + \frac{f_2''}{2!} \cdot I^{(2)} + \cdots + \frac{f_2^{(m_i-1)}}{(m_i-1)!} \cdot I^{(m_i-1)} \right]$$

$$= (f_1 f_2) \cdot I + \frac{f_1' f_2 + f_1 f_2'}{1!} \cdot I^{(1)} + \frac{f_1'' f_2 + 2 f_1' f' + f_1 f_2''}{2!} \cdot I^{(2)} + \cdots$$

$$= (f_1 f_2) \cdot I + \frac{(f_1 f_2)'}{1!} \cdot I^{(1)} + \frac{(f_1 f_2)''}{2!} \cdot I^{(2)} + \cdots + \frac{(f_1 f_2)^{(m_i-1)}}{(m_i-1)!} \cdot I^{(m_i-1)}$$

$$= f(J_i)$$

$$f(A) = P\begin{bmatrix} f(J_1) & & \\ & \ddots & \\ & & f(J_s) \end{bmatrix}P^{-1}$$

$$= P\begin{bmatrix} f_1(J_1) & & \\ & \ddots & \\ & & f_1(J_s) \end{bmatrix}P^{-1} \bullet P\begin{bmatrix} f_2(J_1) & & \\ & \ddots & \\ & & f_2(J_s) \end{bmatrix}P^{-1}$$

$$= f_1(A) \bullet f_2(A)$$

4、矩阵的微分和积分

定义: 如果矩阵 $A(t)=\left(a_{ij}(t)\right)_{m\times n}$,的每一个元素 $a_{ij}(t)$

是变量t 的可微函数，则$A(t)$关于t的导数(微商)定义为

$$\frac{dA(t)}{dt}=\left(a'_{ij}(t)\right)_{m\times n}, \quad \text{或者} \quad A'(t)=\left(a'_{ij}(t)\right)_{m\times n}$$

定理8：设 $A(t),B(t)$ 可导，则有

(1) $$\frac{d}{dt}\left[A(t)+B(t)\right]=\frac{d}{dt}A(t)+\frac{d}{dt}B(t)$$

(2) $A_{m\times n}, f(t)$ 可导 $$\frac{d}{dt}\left[f(t)A(t)\right]=f'(t)A(t)+f(t)A'(t)$$

(3) $A_{m\times n}, A_{n\times l}:$ $$\frac{d}{dt}\left[A(t)B(t)\right]=A'(t)B(t)+A(t)B'(t)$$

证明：（3）左 $$=\frac{d}{dt}\left(\sum_k a_{ik}(t)b_{kj}(t)\right)_{m\times l}$$

$$=\left(\sum_k a'_{ik}(t)b_{kj}(t)+\sum_k a_{ik}(t)b'_{kj}(t)\right)_{m\times l}$$

$$=\left(\sum_k a'_{ik}(t)b_{kj}(t)\right)_{m\times l}+\left(\sum_k a_{ik}(t)b'_{kj}(t)\right)_{m\times l}$$ =右

定理9：设 $A_{n\times n}$ 为数量矩阵，则有

(1) $$\frac{d}{dt}e^{tA}=Ae^{tA}=e^{tA}A$$

(2) $$\frac{d}{dt}\cos(tA)=-A\bullet\sin(tA)=-\sin(tA)\bullet A$$

(3) $$\frac{d}{dt}\sin(tA)=A\bullet\cos(tA)=\cos(tA)\bullet A$$

证明：（1）$e^{tA}=\sum_{n=0}^{\infty}\frac{1}{n!}(tA)^n$ 绝对收敛

$$\left(e^{tA}\right)_{ij}=\delta_{ij}+\frac{t}{1!}(A)_{ij}+\frac{t^2}{2!}\left(A^2\right)_{ij}+\cdots+\frac{t^k}{k!}\left(A^k\right)_{ij}+\cdots$$ 绝对收敛

$$\frac{d}{dt}\left(e^{tA}\right)_{ij} = 0 + \left(A\right)_{ij} + \frac{t}{1!}\left(A^2\right)_{ij} + \cdots + \frac{t^{k-1}}{(k-1)!}\left(A^k\right)_{ij} + \cdots$$

绝对收敛

$$\frac{d}{dt}e^{tA} = A + \frac{t}{1!}A^2 + \cdots + \frac{t^{k-1}}{(k-1)!}A^k + \cdots$$

绝对收敛

$$= \begin{cases} A\left[I + \frac{t}{1!}A + \cdots + \frac{t^{k-1}}{(k-1)!}A^{k-1} + \cdots\right] & = Ae^{tA} \\ \left[I + \frac{t}{1!}A + \cdots + \frac{t^{k-1}}{(k-1)!}A^{k-1} + \cdots\right]A & = e^{tA}A \end{cases}$$

定义: 如果矩阵 $A(t)=\left(a_{ij}(t)\right)_{m\times n}$ 的每一个元素 $a_{ij}(t)$ 在 $[t_0,t]$ 上可积，称$A(t)$可积，记为

$$\int_{t_0}^{t} A(\tau)d\tau=\left(\int_{t_0}^{t} a_{ij}(\tau)d\tau\right)_{m\times n}$$

(1) $\int_{t_0}^{t}\left[A(\tau)+B(\tau)\right]d\tau=\int_{t_0}^{t} A(\tau)d\tau+\int_{t_0}^{t} B(\tau)d\tau$

(2) A为常数矩阵：$\int_{t_0}^{t}\left[A\bullet B(\tau)\right]d\tau=A\bullet\left[\int_{t_0}^{t} B(\tau)d\tau\right]$

B为常数矩阵：$\int_{t_0}^{t}\left[A(\tau)\bullet B\right]d\tau=\left[\int_{t_0}^{t} A(\tau)d\tau\right]\bullet B$

(3) 设 $a_{ij}(t)\in C[t_0,t_1], a\in[t_0,t_1]$ 则：$\frac{d}{dt}\int_{a}^{t} A(\tau)d\tau=A(t)$

(4) 设 $a_{ij}'(t)\in C[t_0,t_1]$， 则：$\int_{t_0}^{t_1} A'(\tau)d\tau=A(t_1)-A(t_0)$

其它微分概念

函数对矩阵的导数(包括向量)

定义:设$X=(\xi_{ij})_{m\times n}$,mn元函数

$$f(X)=f(\xi_{11},\xi_{12},\cdots,\xi_{1n},\cdots,\xi_{m\times n})$$

定义$f(X)$对矩阵X的导数为

$$\frac{df}{dX}=\left(\frac{\partial f}{\partial \xi_{ij}}\right)_{m\times n}=\begin{bmatrix}\frac{\partial f}{\partial \xi_{11}} & \cdots & \frac{\partial f}{\partial \xi_{1n}}\\ \vdots & & \vdots\\ \frac{\partial f}{\partial \xi_{m1}} & \cdots & \frac{\partial f}{\partial \xi_{mn}}\end{bmatrix}$$

例11: $x=\begin{bmatrix}\xi_1\\ \vdots\\ \xi_n\end{bmatrix}$: $f(x)=f(\xi_1,\xi_2,\cdots,\xi_n)$ $\quad \dfrac{df}{dx}=\begin{bmatrix}\dfrac{\partial f}{\partial \xi_1}\\ \vdots\\ \dfrac{\partial f}{\partial \xi_n}\end{bmatrix}$

例12: $A=(a_{ij})_{m\times n}$, $x=\begin{bmatrix}\xi_1\\ \vdots\\ \xi_n\end{bmatrix}$: $f(x)=x^T Ax$, 求 $\dfrac{df}{dx}$

$$f(x)=\xi_1\sum_{j=1}^{n}a_{1j}\xi_j+\cdots+\xi_k\sum_{j=1}^{n}a_{kj}\xi_j+\cdots+\xi_n\sum_{j=1}^{n}a_{nj}\xi_j$$

$$\frac{\partial f}{\partial \xi_k}=\xi_1 a_{1k}+\cdots+\xi_{k-1}a_{k-1,k}+\left(\sum_{j=1}^{n}a_{kj}\xi_j+\xi_k a_{kk}\right)$$

$$+\xi_{k+1}a_{k+1,k}+\cdots+\xi_n a_{nk}\quad=\sum_{j=1}^{n}a_{kj}\xi_j+\sum_{i=1}^{n}a_{ik}\xi_i$$

$$\therefore \frac{df}{dx} = \left(A + A^T \right) x$$

如果 $A = A^T$ ，有 $\dfrac{df}{dx} = 2Ax$

例13: $X=\left(\xi_{ij}\right)_{m\times n}$: $f(X)=\left[\operatorname{tr}(X)\right]^2$ 求 $\left.\dfrac{df}{dX}\right|_{X=I_n}$

解：

$$f(X)=\left(\xi_{11}+\xi_{22}+\cdots+\xi_{nn}\right)^2$$

$$\frac{df}{dX}=2\left(\xi_{11}+\xi_{22}+\cdots+\xi_{nn}\right)I_n$$

$$\left.\frac{df}{dX}\right|_{X=I_n}=2nI_n$$

例14: $A\in R^{m\times n}, b\in R^m$，若 $x\in R^n$ 使得 $\|Ax-b\|_2=\min$,

则 $A^TAx=A^Tb$

解： $f(x)=\|Ax-b\|_2^2=(Ax-b)^T(Ax-b)$

$$=x^TA^TAx-2b^TAx+b^Tb$$

$$g(x)=b^TAx=b_1\sum_{j=1}^{n}a_{1j}\xi_j+\cdots+b_m\sum_{j=1}^{n}a_{mj}\xi_j$$

$$\frac{dg}{dx}=\begin{bmatrix}\frac{\partial g}{\partial \xi_1}\\ \vdots\\ \frac{\partial g}{\partial \xi_n}\end{bmatrix}=\begin{bmatrix}b_1a_{11}+\cdots+b_ma_{m1}\\ \vdots\\ b_1a_{1n}+\cdots+b_ma_{mn}\end{bmatrix}=A^Tb$$

$$\frac{df}{dx}=2A^TAx-2A^Tb=0\quad\Rightarrow A^TAx=A^Tb$$

【注】$r(A^TA)=r(A)\Rightarrow r(A^TA|A^Tb)=r(A^TA)\Rightarrow A^TAx=A^Tb$ 有解

5、函数矩阵对矩阵的导数

定义:设 $X=\left(\xi_{ij}\right)_{m\times n}$, $f_{kl}(X)=f_{kl}\left(\xi_{11},\xi_{12},\cdots,\xi_{1n},\cdots,\xi_{m\times n}\right)$

$$F=\begin{bmatrix} f_{11} & \cdots & f_{1s} \\ \vdots & & \vdots \\ f_{r1} & \cdots & f_{rs} \end{bmatrix},\qquad \frac{\partial F}{\partial \xi_{ij}}=\begin{bmatrix} \frac{\partial f_{11}}{\partial \xi_{ij}} & \cdots & \frac{\partial f_{1s}}{\partial \xi_{ij}} \\ \vdots & & \vdots \\ \frac{\partial f_{r1}}{\partial \xi_{ij}} & \cdots & \frac{\partial f_{rs}}{\partial \xi_{ij}} \end{bmatrix},$$

定义 $$\frac{dF}{dX}=\begin{bmatrix} \frac{\partial F}{\partial \xi_{11}} & \cdots & \frac{\partial F}{\partial \xi_{1n}} \\ \vdots & & \vdots \\ \frac{\partial F}{\partial \xi_{m1}} & \cdots & \frac{\partial F}{\partial \xi_{mn}} \end{bmatrix}$$

- 可表示为

$$\frac{dF}{dX}=\left(\frac{1}{dX}\right)\otimes dF$$

例15: $x=\begin{bmatrix}\xi_1\\ \vdots\\ \xi_n\end{bmatrix},\ F(x)=\left[f_1(x), f_2(x),\cdots, f_l(x)\right]$

$$\frac{dF}{dx}=\begin{bmatrix}\frac{\partial f_1}{\partial \xi_1} & \cdots & \frac{\partial f_l}{\partial \xi_1}\\ \vdots & & \vdots\\ \frac{\partial f_1}{\partial \xi_n} & \cdots & \frac{\partial f_l}{\partial \xi_n}\end{bmatrix}$$

例16: $A=\left(a_{ij}\right)_{n\times n},\quad x=\begin{bmatrix}\xi_1\\ \vdots\\ \xi_n\end{bmatrix}\quad Ax=\begin{bmatrix}\sum_{j=1}^{n}a_{1j}\xi_j\\ \vdots\\ \sum_{j=1}^{n}a_{nj}\xi_j\end{bmatrix}$

$$\frac{d(Ax)}{dx^T}=\begin{bmatrix}a_{11} & \cdots & a_{1n}\\ \vdots & & \vdots\\ a_{n1} & \cdots & a_{nn}\end{bmatrix}=A$$

作业

- **P163：1、2、5、6**
- **P170：4、5、6**

本讲主要内容

- 矩阵分析的应用
- 三角分解
- Givens变换

矩阵分析的应用

$$\begin{cases} \xi_1'(t) = a_{11}\xi_1(t) + a_{12}\xi_2(t) + \cdots + a_{1n}\xi_n(t) + b_1(t) \\ \xi_2'(t) = a_{21}\xi_1(t) + a_{22}\xi_2(t) + \cdots + a_{2n}\xi_n(t) + b_2(t) \\ \cdots \\ \xi_n'(t) = a_{n1}\xi_1(t) + a_{n2}\xi_2(t) + \cdots + a_{nn}\xi_n(t) + b_n(t) \end{cases}$$

$$A = \begin{bmatrix} a_{11} & \cdots & a_{12} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}, x(t) = \begin{bmatrix} \xi_1(t) \\ \vdots \\ \xi_n(t) \end{bmatrix}, b(t) = \begin{bmatrix} b_1(t) \\ \vdots \\ b_n(t) \end{bmatrix}, c(t) = \begin{bmatrix} c_1(t) \\ \vdots \\ c_n(t) \end{bmatrix}$$

齐次微分方程： $x'(t) = A \bullet x(t)$

非齐次微分方程：$x'(t) = A \bullet x(t) + b(t)$

齐次微分方程的解法

定理10：齐次方程 $x'(t)=A\bullet x(t)$ 满足 $x(t_0)=x_0$ 的解存在并且唯一

证：存在性　设 $x(t)=e^{(t-t_0)A}x_0$ ，则

$x'(t)=Ae^{(t-t_0)A}x_0=A\bullet x(t)$　　$x(t_0)=e^{O}x_0=x_0$

唯一性　设 $x(t)$ 满足 $x'(t)=A\bullet x(t), x(t_0)=x_0$

$$x'(t)-Ax(t)=0 \Rightarrow e^{-tA}x'(t)+e^{-tA}(-A)x(t)=0$$

$$\Rightarrow \left[e^{-tA}x(t)\right]'=0 \quad \Rightarrow e^{-tA}x(t)=c \quad \Rightarrow x(t)=ce^{tA}$$

因为 $x(t_0)=x_0$, 所以 $x_0=ce^{t_0A} \Rightarrow c=e^{-t_0A}x_0$

因此 $x(t)=e^{tA}e^{-t_0A}x_0=e^{(t-t_0)A}x_0$

例1：设 $A=\begin{bmatrix} 1 & 1 & 0 \\ & 1 & 0 \\ & & 2 \end{bmatrix}$ 求 $x'(t)=A\bullet x(t)$ 的通解

解： $e^{tA}=\begin{bmatrix} e^t & te^t & 0 \\ & e^t & 0 \\ & & e^{2t} \end{bmatrix}$

$$x(t)=ce^{tA}=\begin{bmatrix} c_1e^t+c_2te^t \\ c_2e^t \\ c_3e^{2t} \end{bmatrix}$$

例2：矩阵函数 e^{tA} 的列向量 $x_1(t),\cdots,x_n(t)$ 构成

齐次方程 $x'(t)=A\bullet x(t)$ 的基础解系

解： e^{tA} 可逆 $\Rightarrow x_1(t),\cdots,x_n(t)$ 线性无关

取 $c=e_j \Rightarrow x_j(t)=e^{tA}c$ 是 $x'(t)=A\bullet x(t)$ 的一个解

通解 $x(t)=ce^{tA}=c_1\bullet x_1(t)+\cdots+c_n\bullet x_n(t)$

非齐次微分方程的解法

方程(1)： $x'(t) = A \bullet x(t)$

方程(2)： $x'(t) = A \bullet x(t) + b(t)$

$$\left.\begin{array}{l}\tilde{x}(t)\text{是(2)的特解}\\ x(t)\text{是(2)的通解}\end{array}\right\} \Rightarrow \left\{\begin{array}{l}\tilde{x}'(t) = A \bullet \tilde{x}(t) + b(t)\\ x'(t) = A \bullet x(t) + b(t)\end{array}\right.$$

$$\Rightarrow \left[\tilde{x}(t) - x(t)\right]' = A\left[\tilde{x}(t) - x(t)\right] \Rightarrow \tilde{x}(t) - x(t)\ \text{是(1)的解}$$

$$\Rightarrow \tilde{x}(t) - x(t) = c_1 \bullet x_1(t) + \cdots + c_n \bullet x_n(t)$$

$$\Rightarrow x(t) = e^{tA} + \tilde{x}(t)$$

非齐次微分方程的解法

采用常向量变易法求 $\tilde{x}(t)$.设 $\tilde{x}(t)=e^{tA}c(t)$ 满足(2),有

$$Ae^{tA}c(t)+e^{tA}c'(t)=Ae^{tA}c(t)+b(t)$$

$$c'(t)=e^{-tA}b(t)\Rightarrow c(t)=\int_{t_0}^{t}e^{-\tau A}b(\tau)d\tau \quad \text{(原函数之一)}$$

故(2)的通解为 $x(t)=e^{tA}\left[c+\int_{t_0}^{t}e^{-\tau A}b(\tau)d\tau\right]$

特解为 $x(t)\Big|_{x(t_0)=x_0}=e^{tA}\left[e^{-t_0A}x_0+\int_{t_0}^{t}e^{-\tau A}b(\tau)d\tau\right]$

[注]当 $t_0=0$时，特解 $x(t)\Big|_{x(0)=x_0}=e^{tA}\left[x_0+\int_{0}^{t}e^{-\tau A}b(\tau)d\tau\right]$

例3：设 $A=\begin{bmatrix}1&1&0\\&1&0\\&&2\end{bmatrix}, b(t)=\begin{bmatrix}1\\0\\e^{2t}\end{bmatrix}, x(0)=\begin{bmatrix}-1\\1\\0\end{bmatrix}$

求 $x'(t)=A\bullet x(t)+b(t)$ 满足初始条件 $x(0)$ 的特解

解： $e^{tA}=\begin{bmatrix}e^t&te^t&0\\&e^t&0\\&&e^{2t}\end{bmatrix}$

$$e^{-\tau A}b(\tau)=\begin{bmatrix}e^{-\tau}\\0\\1\end{bmatrix},\quad \int_{t_0}^{t}e^{-\tau A}b(\tau)d\tau=\begin{bmatrix}1-e^{-t}\\0\\t\end{bmatrix}$$

$$x(t)=e^{tA}\bullet\left[\begin{pmatrix}-1\\0\\1\end{pmatrix}+\begin{pmatrix}1-e^{-t}\\0\\t\end{pmatrix}\right]=\begin{bmatrix}te^t-1\\e^t\\te^{2t}\end{bmatrix}$$

矩阵微分与最优化

最简单的最优化问题是求 $f(t)$ 的极大值和极小值

$$\min_{x\in R} f(t)$$

一般称为无约束的最优化问题

相对于n×1向量x的梯度算子记作∇_x

$$\nabla_x = \left[\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \cdots, \frac{\partial}{\partial x_n}\right]^T = \frac{\partial}{\partial x}$$

n×1实向量x为变元的实标量函数的梯度

$$\nabla_x f(x) = \left[\frac{\partial f(x)}{\partial x_1}, \frac{\partial f(x)}{\partial x_2}, \cdots, \frac{\partial f(x)}{\partial x_n}\right]^T = \frac{\partial f(x)}{\partial x}$$

实标量函数$f(A)$为相对于实矩阵 $A=\left[a_{ij}\right]_{m\times n}$ 的梯度

$$\frac{\partial f(A)}{\partial A}=\left[\frac{\partial f(A)}{\partial a_{ij}}\right]_{m\times n}=\nabla_A f(A)$$

例：CDMA系统中，有K个用户，第k个用户的扩频波形向量为 $\boldsymbol{s}_k(t)$ 。假定用户k的信号幅值为 A_k 在t时刻发送比特为 b_k $(+1,-1)$

在基站解扩后，基站的接收信号向量为

$$\boldsymbol{y}=\boldsymbol{RAb}+\boldsymbol{n}$$

其中 $\boldsymbol{A}=\mathbf{diag}\left(A_1,A_2,\cdots,A_K\right),\boldsymbol{b}=\left[b_1,b_2,\cdots,b_K\right]^T$

扩频相关矩阵$\boldsymbol{R}$的元素 $r_{ij}=\int_0^T s_i(t)s_j(t)dt$

设计一个多用户检测器$\boldsymbol{M}=\left[\boldsymbol{m}_1,\boldsymbol{m}_2,\cdots,\boldsymbol{m}_K\right]$，使得

$$\hat{b}_k=\mathbf{sgn}(\boldsymbol{m}_k^T\boldsymbol{y})$$

将K个用户的检测器联合考虑，构造目标函数

$$J(M)=\mathrm{E}\left[\|b-My\|_2^2\right]$$

使其最小化，即可得到最优的盲多用户检测器M

利用矩阵迹的性质，可得

$$\begin{aligned}J(M)&=\mathrm{E}\left\{(b-My)^T(b-My)\right\}\\&=\mathrm{E}\left\{\mathrm{tr}\left[(b-My)(b-My)^T\right]\right\}\\&=\mathrm{tr}\left\{\mathrm{E}\left[(b-My)(b-My)^T\right]\right\}\\&=\mathrm{tr}\left\{cor(b-My)\right\}\end{aligned}$$

其中 $cor(b-My)=\mathrm{E}\left[(b-My)(b-My)^T\right]$是自相关矩阵

在加性噪声与用户信号不相关时有

$$\mathbf{cor}\left(\boldsymbol{b}-\boldsymbol{My}\right)=\boldsymbol{I}+\boldsymbol{M}\left(\boldsymbol{RA}^2\boldsymbol{R}+\sigma^2\boldsymbol{R}\right)\boldsymbol{M}^T-\boldsymbol{ARM}^T-\boldsymbol{MRA}$$

其中加性噪声的方差为 σ^2

于是目标函数可写作

$$\boldsymbol{J}(\boldsymbol{M})=\mathbf{tr}\left\{\boldsymbol{cor}\left(\boldsymbol{b}-\boldsymbol{My}\right)\right\}$$

$$=\mathbf{tr}(\boldsymbol{I})+\mathbf{tr}\left(\boldsymbol{M}\left(\boldsymbol{RA}^2\boldsymbol{R}+\sigma^2\boldsymbol{R}\right)\boldsymbol{M}^T\right)-\mathbf{tr}\left(\boldsymbol{ARM}^T\right)-\mathbf{tr}\left(\boldsymbol{MRA}\right)$$

利用迹函数的微分公式

$$\frac{\partial\,\mathbf{tr}\left(\boldsymbol{M}^T\boldsymbol{B}\right)}{\partial\boldsymbol{M}}=\frac{\partial\,\mathbf{tr}\left(\boldsymbol{BM}^T\right)}{\partial\boldsymbol{M}}=\boldsymbol{B}\qquad\frac{\partial\,\mathbf{tr}\left(\boldsymbol{MB}\right)}{\partial\boldsymbol{M}}=\frac{\partial\,\mathbf{tr}\left(\boldsymbol{BM}\right)}{\partial\boldsymbol{M}}=\boldsymbol{B}^T$$

$$\frac{\partial \operatorname{tr}\left(MDM^T\right)}{\partial M} = M(D+D^T)$$

因为 $D = RA^2R + \sigma^2 R$ 是对称矩阵，所以

$$\frac{\partial J(M)}{\partial M} = 2M\left(RA^2R + \sigma^2 R\right) - 2AR$$

令其为零，即可得

$$M\left(RA^2R + \sigma^2 R\right) = AR$$

如果R非奇异，可得最优的多用户检测器为

$$M = A\left(RA^2 + \sigma^2 I\right)^{-1}$$

[引入]在线性代数中应用Gauss消去法求解n元

线性方程组 $\boldsymbol{Ax}=\boldsymbol{b}$

其中： $\boldsymbol{A}=\left(a_{ij}\right)_{n\times n},\boldsymbol{x}=\left(x_1,x_2,\cdots,x_n\right)^T,\boldsymbol{b}=\left(b_1,b_2,\cdots,b_n\right)^T$

Gauss消去法将系数矩阵化为上三角形矩阵，或将增广矩阵化为上阶梯形矩阵，而后回代求解。

定义4.1 如果n阶矩阵A能够分解为一个下三角矩阵L和一个上三角矩阵U的乘积，则称其为三角分解或LU分解。如果方阵A可分解成$A=LDU$,其中L为一个单位下三角矩阵，D为对角矩阵，则称A可作LDU分解。

定理4.1 矩阵 $A=\left(a_{ij}\right)_{n\times n}$ 的分解式唯一的充要条件为A的顺序主子式 $\Delta_k \neq 0$ 。$A=LDU$，其中L是单位下三角矩阵，U是单位上三角矩阵，D是对角矩阵，并且

$$D=\mathrm{diag}\left(d_1,d_2,\cdots,d_n\right),d_k=\frac{\Delta_k}{\Delta_{k-1}},k=1,2,\cdots,n \quad \left(\Delta_0=1\right)$$

推论 设A是n阶非奇异矩阵，A有三角分解$A=LU$，的充要条件是A的顺序主子式 $\Delta_k \neq 0 \quad k = 1, 2, \cdots, n$

分解原理：以n=4为例

$$\Delta_1(A) = a_{11} : a_{11} \neq 0 \quad \Rightarrow c_{i1} = \frac{a_{i1}}{a_{11}} \quad (i = 2,3,4)$$

$$L_1 = \begin{bmatrix} 1 & & & \\ c_{21} & 1 & & \\ c_{31} & 0 & 1 & \\ c_{41} & 0 & 0 & 1 \end{bmatrix}, L_1^{-1} = \begin{bmatrix} 1 & & & \\ -c_{21} & 1 & & \\ -c_{31} & 0 & 1 & \\ -c_{41} & 0 & 0 & 1 \end{bmatrix}$$

$$L_1^{-1}A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ & a_{22}^{(1)} & a_{23}^{(1)} & a_{24}^{(1)} \\ & a_{32}^{(1)} & a_{33}^{(1)} & a_{34}^{(1)} \\ & a_{42}^{(1)} & a_{43}^{(1)} & a_{44}^{(1)} \end{bmatrix} = A^{(1)}$$

(2) $\Delta_2(A) = \Delta_2(A^{(1)}) = a_{11}a_{22}^{(1)}: \quad a_{22}^{(1)} \neq 0$

$$\Rightarrow c_{i2} = \frac{a_{i2}^{(1)}}{a_{22}^{(1)}} \quad (i = 3,4)$$

$$L_2 = \begin{bmatrix} 1 & & & \\ 0 & 1 & & \\ 0 & c_{32} & 1 & \\ 0 & c_{42} & 0 & 1 \end{bmatrix}, L_2^{-1} = \begin{bmatrix} 1 & & & \\ 0 & 1 & & \\ 0 & -c_{32} & 1 & \\ 0 & -c_{42} & 0 & 1 \end{bmatrix}$$

$$L_2^{-1}A^{(1)} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ & a_{22}^{(1)} & a_{23}^{(1)} & a_{24}^{(1)} \\ & & a_{33}^{(2)} & a_{34}^{(2)} \\ & & a_{43}^{(2)} & a_{44}^{(2)} \end{bmatrix} = A^{(2)}$$

$$(3)\ \Delta_3(A) = \Delta_3(A^{(2)}) = a_{11}a_{22}^{(1)}a_{33}^{(2)} : a_{33}^{(2)} \neq 0$$

$$\Rightarrow c_{43} = \frac{a_{43}^{(2)}}{a_{33}^{(2)}}$$

$$L_3 = \begin{bmatrix} 1 & & & \\ 0 & 1 & & \\ 0 & 0 & 1 & \\ 0 & 0 & c_{43} & 1 \end{bmatrix}, L_3^{-1} = \begin{bmatrix} 1 & & & \\ 0 & 1 & & \\ 0 & 0 & 1 & \\ 0 & 0 & -c_{43} & 1 \end{bmatrix}$$

$$L_3^{-1}A^{(2)} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ & a_{22}^{(1)} & a_{23}^{(1)} & a_{24}^{(1)} \\ & & a_{33}^{(2)} & a_{34}^{(2)} \\ & & & a_{44}^{(3)} \end{bmatrix} = A^{(3)}$$

即： $L_3^{-1}L_2^{-1}L_1^{-1}A = A^{(3)} \quad \Rightarrow A = L_1L_2L_3A^{(3)}$

令 $L = L_1L_2L_3 = \begin{bmatrix} 1 & & & \\ c_{21} & 1 & & \\ c_{31} & c_{32} & 1 & \\ c_{41} & c_{42} & c_{43} & 1 \end{bmatrix}$ 则 $A = LA^{(3)}$

分解 $A^{(3)} = \begin{bmatrix} a_{11} & & & \\ & a_{22}^{(1)} & & \\ & & a_{33}^{(2)} & \\ & & & a_{44}^{(3)} \end{bmatrix} \begin{bmatrix} 1 & * & * & * \\ & 1 & * & * \\ & & 1 & * \\ & & & 1 \end{bmatrix} = DU$

则 $A = LDU$

二、紧凑格式算法： $A = LDU = \tilde{L}U$ (Crout分解)

$$L = \tilde{L} = \begin{bmatrix} l_{11} & & & \\ l_{21} & l_{22} & & \\ \vdots & \vdots & \ddots & \\ l_{n1} & l_{n2} & \cdots & l_{nn} \end{bmatrix}, U = \begin{bmatrix} 1 & u_{12} & \cdots & u_{1n} \\ & 1 & \cdots & u_{2n} \\ & & \ddots & \vdots \\ & & & 1 \end{bmatrix}$$

$(i,1)$元： $a_{i1} = l_{i1} \bullet 1 \quad \Rightarrow l_{i1} = a_{i1} \quad (i = 1, \cdots, n)$

$(1,j)$元： $a_{1j} = l_{11} \bullet u_{1j} \Rightarrow u_{1j} = \dfrac{a_{1j}}{l_{11}} \quad (j = 2, \cdots, n)$

(i,k)元： $a_{ik} = l_{i1} \bullet u_{1k} + \cdots + l_{i,k-1} \bullet u_{k-1,k} + l_{ik} \bullet 1 \quad (i \geq k)$

$$\Rightarrow l_{ik} = a_{ik} - \left(l_{i1} \bullet u_{1k} + \cdots + l_{i,k-1} \bullet u_{k-1,k} \right)$$

(k,j)元：$a_{kj} = l_{k1} \bullet u_{1,j} + \cdots + l_{k,k-1} \bullet u_{k-1,j} + l_{kk} \bullet u_{kj} \quad (j > k)$

$$\Rightarrow u_{kj} = \frac{1}{l_{kk}} \left[a_{kj} - \left(l_{k1} \bullet u_{1j} + \cdots + l_{k,k-1} \bullet u_{k-1,j} \right) \right]$$

计算框图：

$$
\begin{array}{lllll|l}
l_{11} & u_{12} & u_{13} & u_{14} & \cdots & \text{第1框} \\
l_{21} & l_{22} & u_{23} & u_{24} & \cdots & \text{第2框} \\
l_{31} & l_{32} & l_{33} & u_{34} & \cdots & \text{第3框} \\
l_{41} & l_{42} & l_{43} & l_{44} & \cdots & \text{第4框} \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots
\end{array}
$$

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

计算框图:

$$\begin{array}{c|c|c|c} 5 & 2/5 & -4/5 & 0 \\ \cline{2-4} 2 & 1/5 & -2 & 5 \\ \cline{3-4} -4 & -2/5 & 1 & 2 \\ \cline{4-4} 0 & 1 & 2 & -7 \end{array}$$

$$\tilde{L}=\begin{bmatrix} 5 & 0 & 0 & 0 \\ 2 & 1/5 & 0 & 0 \\ -4 & -2/5 & 1 & 0 \\ 0 & 1 & 2 & -7 \end{bmatrix}=\begin{bmatrix} 1 & & & \\ 2/5 & 1 & & \\ -4/5 & -2 & 1 & \\ 0 & 5 & 2 & 1 \end{bmatrix}\begin{bmatrix} 5 & & & \\ & 1/5 & & \\ & & 1 & \\ & & & -7 \end{bmatrix}$$

$$U=\begin{bmatrix} 1 & 2/5 & -4/5 & 0 \\ 0 & 1 & -2 & 5 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix},\qquad A=\tilde{L}U=LDU$$

QR分解

目的：将分解为正交矩阵与上三角矩阵之积.

约定：本节涉及的矩阵为实矩阵，向量为实向量，数为实数.

一、Givens矩阵

$$T_{ij}(c,s) \overset{\Delta}{=} \begin{bmatrix} I & & & & \\ & c & & s & \\ & & I & & \\ & -s & & c & \\ & & & & I \end{bmatrix} \begin{array}{c} \\ (i) \\ \\ (j) \\ \\ \end{array}$$

$$c^2 + s^2 = 1$$

性质：

(1) $T_{ij}^T T_{ij} = I, \left[T_{ij}(c,s)\right]^{-1} = \left[T_{ij}(c,s)\right]^T = T_{ij}(c,-s), \quad \left|T_{ij}\right| = 1$

(2) $$x = \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix}, \quad T_{ij}x \stackrel{\Delta}{=} \begin{bmatrix} \eta_1 \\ \vdots \\ \eta_n \end{bmatrix} \Rightarrow \begin{cases} \eta_i = c\xi_i + s\xi_j \\ \eta_j = -s\xi_i + c\xi_j \\ \eta_k = \xi_k (k \neq i, j) \end{cases}$$

若 $\xi_i^2 + \xi_j^2 \neq 0$，取 $c = \dfrac{\xi_i}{\sqrt{\xi_i^2 + \xi_j^2}}, s = \dfrac{\xi_j}{\sqrt{\xi_i^2 + \xi_j^2}}$

则 $$\eta_i = \sqrt{\xi_i^2 + \xi_j^2} > 0, \eta_j = 0$$

定理3：$x \neq \mathbf{0} \Rightarrow \exists$ 有限个G-矩阵之积T，st．$Tx = |x| e_1$

推论：设非零列向量 $x \in R^n$ 及单位列向量 $z \in R^n$，

则存在有限个Givens矩阵之积，记作T，使得

$$Tx = |x| z$$

例： $x=\begin{bmatrix}3\\4\\5\end{bmatrix}$ 求G-矩阵之积T,使得 $Tx=|x|e_1$

解： $T_{12}(c,s)$ 中， $c=\frac{3}{5}, s=\frac{4}{5}$. $T_{12}x=\begin{bmatrix}5\\0\\5\end{bmatrix}$

$T_{13}(c,s)$ 中, $c=\frac{1}{\sqrt{2}}, s=\frac{1}{\sqrt{2}}$. $T_{13}\left(T_{12}x\right)=\begin{bmatrix}5\sqrt{2}\\0\\0\end{bmatrix}=|x|e_1$

$$T=T_{13}T_{12}=\frac{1}{\sqrt{2}}\begin{bmatrix}1&0&1\\0&\sqrt{2}&0\\-1&0&1\end{bmatrix}\bullet\frac{1}{5}\begin{bmatrix}3&4&0\\-4&3&0\\0&0&5\end{bmatrix}$$

$$=\frac{1}{5\sqrt{2}}\begin{bmatrix}3&4&5\\-4\sqrt{2}&3\sqrt{2}&0\\-3&-4&5\end{bmatrix}$$

$Tx=5\sqrt{2}e_1$

本讲主要内容

- 矩阵的QR分解
- 矩阵的满秩分解
- 矩阵的奇异值分解

Householder矩阵

在平面 R^2 中，将向量 x 映射为关于 e_1 对称的向量y的变换，称为是关于e_1轴的镜像(反射)变换

设 $x = \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix}$ ，有

$$y = \begin{bmatrix} \xi_1 \\ -\xi_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} = \left(I - 2e_2 e_2^T \right) x = Hx$$

其中， $e_2 = [0 \;\; 1]^T$ ，H是正交矩阵，且 $|H| = -1$

Householder矩阵

将向量 x 映射为关于“与单位向量u正交的直线”对称的向量y的变换， $x-y=2u\left(u^T x\right)$

$$y=x-2u\left(u^T x\right) \quad =\left(I-2uu^T\right)x=Hx$$

显然，H是正交矩阵

定义：设单位列向量 $u\in R^n$ ，称 $H=I-2uu^T$ 为**Householder**矩阵(初等反射矩阵)，由**H**矩阵确定的线性变换称为**Householder**变换。

Householder矩阵

$$H_u = I_n - 2uu^T \qquad (\ u \in R^n \text{ 是单位列向量})$$

(1) $H = H^T$ 对称　　(2) $H^T H = I$ 正交

(3) $H^2 = I$ 对合　　(4) $H^{-1} = H$ 自逆

(5) $\det H = -1$ 自逆

验证(5)：

$$\begin{bmatrix} I & 0 \\ -u^T & 1 \end{bmatrix}\begin{bmatrix} I & 2u \\ 0^T & 1 \end{bmatrix}\begin{bmatrix} I-2uu^T & 0 \\ u^T & 1 \end{bmatrix} = \begin{bmatrix} I & 0 \\ -u^T & 1 \end{bmatrix}\begin{bmatrix} I & 2u \\ u^T & 1 \end{bmatrix} = \begin{bmatrix} I & 2u \\ 0^T & -1 \end{bmatrix}$$

$$\begin{vmatrix} I-2uu^T & 0 \\ u^T & 1 \end{vmatrix} = \begin{vmatrix} I & 2u \\ 0^T & -1 \end{vmatrix} = -1$$

定理4：R^n 中 $(n>1), \forall x \neq 0, \forall$ 单位列向量 z

$$\Rightarrow \exists H_u, \text{st } H_u x = |x| z$$

证明：(1) $x = |x| z: n > 1$ 时，取单位向量u使得 $u \perp x$,

于是 $H_u = I - 2uu^T : H_u x = Ix - 2uu^T x = x = |x| z$

(2) $x \neq |x| z$： 取 $u = \dfrac{x - |x| z}{\left| x - |x| z \right|}$， 有

$$H_u x = \left[I - 2\frac{(x - |x| z)(x - |x| z)^T}{\left| x - |x| z \right|^2} \right] x = x - \frac{2(x - |x| z, x)}{\left| x - |x| z \right|^2} (x - |x| z)$$

$$= x - 1 \times (x - |x| z) = |x| z$$

例2：$x=\begin{bmatrix}1\\2\\2\end{bmatrix}$，求H-矩阵$H$ 使得 $Hx=|x|e_1$

解：$|x|=3,\ x-|x|e_1=\begin{bmatrix}-2\\2\\2\end{bmatrix},\quad u=\frac{1}{\sqrt{3}}\begin{bmatrix}-1\\1\\1\end{bmatrix}$

$$H=I-\frac{2}{3}\begin{bmatrix}-1\\1\\1\end{bmatrix}\begin{bmatrix}-1 & 1 & 1\end{bmatrix}=\frac{1}{3}\begin{bmatrix}1 & 2 & 2\\2 & 1 & -2\\2 & -2 & 1\end{bmatrix}$$

$$Hx=3e_1$$

G矩阵与H-矩阵的关系

定理5：G-矩阵 $T_{ij}(c,s) \Rightarrow \exists$H-矩阵$H_u$与$H_v$,st$T_{ij} = H_u H_v$

证明：$c^2+s^2=1 \Rightarrow$取$\theta = \arctan\frac{s}{c}$,则$\cos\theta = c, \sin\theta = s$

$$
T_{ij}(c,s) = \begin{array}{c} \begin{bmatrix} I & & & & \\ & \cos\theta & & \sin\theta & \\ & & I & & \\ & -\sin\theta & & \cos\theta & \\ & & & & I \end{bmatrix} \end{array} \begin{array}{c} \\ (i) \\ \\ (j) \\ \\ \end{array}
$$

$$
v = \begin{bmatrix} 0 & \cdots & 0 & \sin\frac{\theta}{4} & 0 & \cdots & 0 & \cos\frac{\theta}{4} & 0 & \cdots & 0 \end{bmatrix}^T
$$

$$H_v = \begin{bmatrix} I & & & & \\ & 1 & & & \\ & & I & & \\ & & & 1 & \\ & & & & I \end{bmatrix} - 2\begin{bmatrix} O & & & & \\ & \sin^2\frac{\theta}{4} & & \sin\frac{\theta}{4}\cos\frac{\theta}{4} & \\ & & O & & \\ & \sin\frac{\theta}{4}\cos\frac{\theta}{4} & & \cos^2\frac{\theta}{4} & \\ & & & & O \end{bmatrix}$$

$$= \begin{bmatrix} I & & & & \\ & \cos\frac{\theta}{2} & & -\sin\frac{\theta}{2} & \\ & & I & & \\ & -\sin\frac{\theta}{2} & & -\cos\frac{\theta}{2} & \\ & & & & I \end{bmatrix}$$

$$u=\begin{bmatrix}0 & \cdots & 0 & \sin\frac{3\theta}{4} & 0 & \cdots & 0 & \cos\frac{3\theta}{4} & 0 & \cdots & 0\end{bmatrix}^T$$

$$H_u=\begin{bmatrix}I & & & & \\ & \cos\frac{3\theta}{2} & & -\sin\frac{3\theta}{2} & \\ & & I & & \\ & -\sin\frac{3\theta}{2} & & -\cos\frac{3\theta}{2} & \\ & & & & I\end{bmatrix},$$

$$T_{ij}(c,s)=H_uH_v \qquad \#$$

[注] H-矩阵不能由若干个G矩阵的乘积来表示。

因为 $\det H=-1,$ 而 $\det G=1$

例3：G-矩阵 $T_{ij}(0,1)=\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$中, $c=0, s=1 \Rightarrow \theta=\pi/2$

$$H_u=\frac{1}{\sqrt{2}}\begin{bmatrix} -1 & -1 \\ -1 & 1 \end{bmatrix}, H_v=\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ -1 & -1 \end{bmatrix}$$

$$\Rightarrow H_u H_v=\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

四、QR分解

1. Schmidt正交化方法

定理6： $A_{n\times n}$ 可逆 $\Rightarrow$ ∃正交矩阵Q，可逆上三角矩阵R，使得$A=QR$。

证明：$A=(a_1,a_2,\cdots,a_n)$ 可逆 $\Rightarrow a_1,a_2,\cdots,a_n$ 线性无关，

正交化后可得：

$$\begin{cases} b_1=a_1 \\ b_2=a_2-k_{21}b_1 \\ \quad\cdots\cdots \\ b_n=a_n-k_{n,n-1}b_{n-1}-\cdots-k_{n1}b_1 \end{cases} \qquad \begin{cases} a_1=b_1 \\ a_2=k_{21}b_1+b_2 \\ \quad\cdots\cdots \\ a_n=k_{n1}b_1+\cdots+k_{n,n-1}b_{n-1}+b_n \end{cases}$$

$$(a_1,a_2,\cdots,a_n)=(b_1,b_2,\cdots,b_n)K$$

$$=(q_1,q_2,\cdots,q_n)\begin{bmatrix}|b_1| & & & \\ & |b_2| & & \\ & & \ddots & \\ & & & |b_n|\end{bmatrix}\begin{bmatrix}1 & k_{21} & \cdots & k_{n1} \\ & 1 & \cdots & k_{n2} \\ & & \ddots & \vdots \\ & & & 1\end{bmatrix}$$

令 $Q=(q_1,q_2,\cdots,q_n),R=\begin{bmatrix}|b_1| & & & \\ & |b_2| & & \\ & & \ddots & \\ & & & |b_n|\end{bmatrix}\begin{bmatrix}1 & k_{21} & \cdots & k_{n1} \\ & 1 & \cdots & k_{n2} \\ & & \ddots & \vdots \\ & & & 1\end{bmatrix}$

则$A=QR$，其中 $q_i=\dfrac{b_i}{|b_i|}\quad(i=1,2,\cdots,n)$

例4：求 $A=\begin{bmatrix}1&2&2\\2&1&2\\1&2&1\end{bmatrix}$ 的QR分解。

解：$b_1=a_1=\begin{bmatrix}1\\2\\1\end{bmatrix}, b_2=a_2-1\times b_1=\begin{bmatrix}1\\-1\\1\end{bmatrix}, b_3=a_3-\frac{1}{3}b_2-\frac{7}{6}b_1=\begin{bmatrix}1/2\\2\\-1/2\end{bmatrix}$

$$Q=\begin{bmatrix}\frac{1}{\sqrt{6}}&\frac{1}{\sqrt{3}}&\frac{1}{\sqrt{2}}\\\frac{2}{\sqrt{6}}&-\frac{1}{\sqrt{3}}&0\\\frac{1}{\sqrt{6}}&\frac{1}{\sqrt{3}}&-\frac{1}{\sqrt{2}}\end{bmatrix}\qquad R=\begin{bmatrix}\sqrt{6}&&\\&\sqrt{3}&\\&&\frac{1}{\sqrt{2}}\end{bmatrix}\begin{bmatrix}1&1&\frac{7}{6}\\&1&\frac{1}{3}\\&&1\end{bmatrix}=\begin{bmatrix}\sqrt{6}&\sqrt{6}&\frac{7}{\sqrt{6}}\\&\sqrt{3}&\frac{1}{\sqrt{3}}\\&&\frac{1}{\sqrt{2}}\end{bmatrix}$$

定理7： $A_{m\times n}$ 列满秩 $\Rightarrow \exists$ 矩阵 $Q_{m\times n}$ 满足 $Q^H Q = I$,

可逆上三角矩阵 $R_{n\times n}$ ，使得A=QR。

证明：同定理6

2. G-变换方法

定理8： $A_{n\times n}$ 可逆 $\Rightarrow \exists$ 有限个G-矩阵之积T，使得TA为可逆上三角矩阵。

证明：略

2. H-变换方法

定理10：$A_{n\times n}$ 可逆 $\Rightarrow \exists$ 有限个H-矩阵之积S，

使得SA为可逆上三角矩阵。

证明：略

五、化方阵与Hessenberg矩阵相似

上 Hessenberg 矩阵：$F_{上} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ & \ddots & \ddots & \ddots & \vdots \\ & & a_{n-1,n-2} & a_{n-1,n-1} & a_{n-1,n} \\ & & & a_{n,n-1} & a_{nn} \end{bmatrix}$

定理11： $A_{n\times n}$ ，则存在有限个G-矩阵之积Q，使得

$QAQ^{\mathrm{T}} = F_{上}$

定理12： $A_{n\times n}$ ，则存在有限个H-矩阵之积Q，使得

$QAQ^{\mathrm{T}} = F_{上}$

推论： $A_{n\times n}$ 实对称 $\Rightarrow\exists$存在有限个H-矩阵(G-矩阵)之积Q，使得 $QAQ^{\mathrm{T}} =$ “实对称三对角矩阵”

例8: 用H-变换化 $A=\begin{bmatrix}1&0&1\\0&1&2\\1&2&1\end{bmatrix}$ 正交相似于“三对角矩阵”

解: $\beta^{(0)}=\begin{bmatrix}0\\1\end{bmatrix}$: $\beta^{(0)}-\left|\beta^{(0)}\right|e_1=\begin{bmatrix}-1\\1\end{bmatrix}$, $u=\frac{1}{\sqrt{2}}\begin{bmatrix}-1\\1\end{bmatrix}$

$$H_0=\begin{bmatrix}0&1\\1&0\end{bmatrix},\quad Q\triangleq\begin{bmatrix}1&\\&H_0\end{bmatrix}=\begin{bmatrix}1&0&0\\0&0&1\\0&1&0\end{bmatrix}$$

$$QA=\begin{bmatrix}1&0&1\\1&2&1\\0&1&2\end{bmatrix},QAQ^T=\begin{bmatrix}1&1&0\\1&1&2\\0&1&0\end{bmatrix}$$

满秩分解

目的：对$A\in C_r^{m\times r}(r\geq 1)$, 求$F\in C_r^{m\times r}$，及$G\in C_r^{m\times r}$使$A=FG$

分解原理：

$$\mathrm{rank}A=r\Rightarrow A\xrightarrow{\text{行}}\text{阶梯形}\,B=\begin{pmatrix}G\\O\end{pmatrix}:\quad G\in \mathrm{C}_r^{r\times n}$$

$$\Rightarrow\exists\text{有限个初等矩阵之积}\,P_{m\times m},\mathrm{st}.PA=B$$

$$\Rightarrow A=P^{-1}B=\left(F_{m\times r}\,\middle|\,S_{m\times(m-r)}\right)\begin{pmatrix}G\\O\end{pmatrix}=FG:\quad F\in \mathrm{C}_r^{m\times r}$$

例9: $A=\begin{bmatrix} -1 & 0 & 1 & 2 \\ 1 & 2 & -1 & 1 \\ 2 & 2 & -2 & -1 \end{bmatrix}$, 求$A=FG$

解 (1) $(A|I)=\left[\begin{array}{cccc|ccc} -1 & 0 & 1 & 2 & 1 & & \\ 1 & 2 & -1 & 1 & & 1 & \\ 2 & 2 & -2 & -1 & & & 1 \end{array}\right] \to \left[\begin{array}{cccc|ccc} -1 & 0 & 1 & 2 & 1 & & \\ 0 & 2 & 0 & 3 & 1 & 1 & \\ 0 & 0 & 0 & 0 & 1 & -1 & 1 \end{array}\right]$

$$P=\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & -1 & 1 \end{bmatrix}, P^{-1}=\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -2 & 1 & 1 \end{bmatrix}, F=\begin{bmatrix} 1 & 0 \\ -1 & 1 \\ -2 & 1 \end{bmatrix}$$

$G=\begin{bmatrix} -1 & 0 & 1 & 2 \\ 0 & 2 & 0 & 3 \end{bmatrix}$ 满秩分解为$A=FG$

例9: $A=\begin{bmatrix}-1&0&1&2\\1&2&-1&1\\2&2&-2&-1\end{bmatrix}$, 求$A=FG$

解 (2) $(A|I)=\left[\begin{array}{cccc|ccc}-1&0&1&2&1&&\\1&2&-1&1&&1&\\2&2&-2&-1&&&1\end{array}\right]\rightarrow\left[\begin{array}{cccc|ccc}1&0&-1&-2&-1&&\\0&1&0&3/2&1/2&1/2&\\0&0&0&0&1&-1&1\end{array}\right]$

$$A=P^{-1}B=(F\mid S)\begin{pmatrix}I_2&B_{12}\\O&O\end{pmatrix}=(F\mid FB_{12})$$

故 $F=$ "A 的前 2 列" $=\begin{bmatrix}-1&0\\1&2\\2&2\end{bmatrix}$, $G=\begin{bmatrix}1&0&-1&-2\\0&1&0&3/2\end{bmatrix}$

奇异值分解

一、预备知识

(1) $\forall A_{m\times n}, \left(A^{\mathrm{H}}A\right)_{n\times n}$ 是 Hermite（半）正定矩阵.

$$\forall x\neq 0,\ x^{\mathrm{H}}A^{\mathrm{H}}Ax=\left(Ax\right)^{\mathrm{H}}\left(Ax\right)=\left|Ax\right|^{2}\geq 0$$

(2) 齐次方程组 $Ax=0$ 与 $A^{\mathrm{H}}Ax=0$ 同解

若 $Ax=0$，则 $A^{\mathrm{H}}Ax=0$；

反之， $A^{\mathrm{H}}Ax=0$ $\Rightarrow\left|Ax\right|^{2}=\left(Ax\right)^{\mathrm{H}}\left(Ax\right)=x^{\mathrm{H}}(A^{\mathrm{H}}Ax)=0$

$\Rightarrow Ax=0$

(3) $\operatorname{rank} A = \operatorname{rank}(A^{\mathrm{H}} A)$

$$S_1 = \{x \mid Ax = 0\}, \quad S_2 = \{x \mid A^{\mathrm{H}} Ax = 0\}$$

$$S_1 = S_2 \Rightarrow \dim S_1 = \dim S_2 \Rightarrow n - r_A = n - r_{A^{\mathrm{H}} A}$$

$$\Rightarrow r_A = r_{A^{\mathrm{H}} A}$$

(4) $A = O_{m\times n} \Leftrightarrow A^{\mathrm{H}} A = O_{n\times n}$

必要性．左乘即得；

充分性 $r_A = r_{A^{\mathrm{H}} A} = 0 \Rightarrow A = O$

二、正交对角分解

定理15： $A_{n\times n}$ 可逆 $\Rightarrow$ ∃酉矩阵 $U_{n\times n}, V_{n\times n}$ ，使得

$$U^{H}AV=\begin{bmatrix}\sigma_1 & & \\ & \ddots & \\ & & \sigma_n\end{bmatrix}\stackrel{\Delta}{=}D \quad (\sigma_i>0)$$

证：$A^H A$是Hermite正定矩阵，∃酉矩阵$V_{n\times n}$，使得

$$V^{H}\left(A^{H}A\right)V=\mathrm{diag}\left(\lambda_1,\cdots,\lambda_n\right)\stackrel{\Delta}{=}\Lambda \quad (\lambda_i>0)$$

改写为 $D^{-1}V^{H}A^{H}\cdot AVD^{-1}=I \quad (\sigma_i=\sqrt{\lambda_i})$

令 $U=AVD^{-1}$, 则有 $U^H U=I$, 从而U是酉矩阵。

由此可得 $U^H AV=U^H UD=D$

三、奇异值分解

$$A_{n\times n}\in C_r^{m\times n}(r\geq 1)\Rightarrow A^H A\in C_r^{n\times n}$$ 半正定

$A^H A$ 的特征值：$\lambda_1\geq\lambda_2\geq\cdots\geq\lambda_r\geq\lambda_{r+1}=\cdots=\lambda_n=0$

A的奇异值：$\sigma_i=\sqrt{\lambda_i},\quad i=1,2,\cdots n$

特点：（1）A的奇异值个数等于A的列数

（2）A的非零奇异值个数等于 $\mathrm{rank}\,A$

定理16：$A_{n\times n}\in C_r^{m\times n}(r\geq 1),\Sigma_r=\begin{bmatrix}\sigma_1 & & \\ & \ddots & \\ & & \sigma_r\end{bmatrix}\Rightarrow$

存在酉矩阵 $U_{m\times m}$ 及 $V_{n\times n}$ ，使得 $U^H AV=\begin{bmatrix}\Sigma_r & O\\ O & O\end{bmatrix}_{m\times n}\overset{\Delta}{=}D$

[注]：称 $A=UDV^H$ 为A的奇异值分解

(1) U与V不唯一；

(2) U的列为 AA^H 的特征向量，V的列为 $A^H A$的特征向量

(3) 称U的列为A的左奇异向量，称V的列为A的右奇异向量.

例10：称 $A=\begin{bmatrix}1&0&1\\0&1&1\\0&0&0\end{bmatrix}$，求 $A=UDV^T$

解： $AA^T=\begin{bmatrix}1&0&1\\0&1&1\\1&1&2\end{bmatrix}=B, |\lambda I-B|=\lambda(\lambda-1)(\lambda-3)$

$$\lambda_1=3: \quad 3I-B=\begin{bmatrix}2&0&-1\\0&2&-1\\-1&-1&1\end{bmatrix}, \xi_1=\begin{bmatrix}1\\1\\2\end{bmatrix}$$

$$\lambda_2=1: \quad 1I-B=\begin{bmatrix}0&0&-1\\0&0&-1\\-1&-1&-1\end{bmatrix}, \xi_2=\begin{bmatrix}1\\-1\\0\end{bmatrix}$$

$$\lambda_3 = 0: \quad 0I - B = \begin{bmatrix} -1 & 0 & -1 \\ 0 & -1 & -1 \\ -1 & -1 & -2 \end{bmatrix}, \xi_3 = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$$

$$r_A = 2: \quad \Sigma = \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \end{bmatrix}$$

$$V = \begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{2}{\sqrt{6}} & 0 & -\frac{1}{\sqrt{3}} \end{bmatrix}, \quad V_1 = \begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \\ \frac{2}{\sqrt{6}} & 0 \end{bmatrix}$$

$$U_1 = AV_1\Sigma^{-1} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 0 & 0 \end{bmatrix}，\text{取} U_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}，\quad \text{则} U = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$U^{\mathrm{T}}AV = \begin{bmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} = D, \quad A = UDV^{\mathrm{T}}$$

定理17：$A_{n\times n}\in C_r^{m\times n}(r\geq 0)$ 的奇异值分解 $A=U\begin{bmatrix}\Sigma & O\\ O & O\end{bmatrix}V^H$

中，划分 $U=\left(u_1,u_2,\cdots,u_m\right),V=\left(v_1,v_2,\cdots,v_n\right)$，则有

(1) $N(A)=\mathrm{span}\{v_{r+1},v_{r+2},\cdots,v_n\}$;

(2) $R(A)=\mathrm{span}\{u_1,u_2,\cdots,u_r\}$;

(3) $A=\sigma_1 u_1 v_1^{\mathrm{H}}+\sigma_2 u_2 v_2^{\mathrm{H}}+\cdots+\sigma_r u_r v_r^{\mathrm{H}}$

证明： $A=\left(U_1 \mid U_2\right)\begin{pmatrix}\Sigma & O\\ O & O\end{pmatrix}\begin{pmatrix}V_1^{\mathrm{H}}\\ V_2^{\mathrm{H}}\end{pmatrix}=U_1\Sigma V_1^{\mathrm{H}}$

容易验证： $U_1\Sigma V_1^{\mathrm{H}}x=0 \Leftrightarrow V_1^{\mathrm{H}}x=0$

(1) $N(A)=\{x \mid Ax=0\}=\{x \mid U_1 \Sigma V_1^{\mathrm{H}} x=0\}$

$$=\{x \mid V_1^{\mathrm{H}} x=0\}=N(V_1^{\mathrm{H}})=R^{\perp}(V_1)$$

$$=R(V_2)=\mathrm{span}\{v_{r+1},\cdots,v_n\}$$

(2) $R(A)=\{y \mid y=Ax\}=\{y \mid y=U_1(\Sigma V_1^{\mathrm{H}} x)\}$

$$\subset\{y \mid y=U_1 z\}=R(U_1)$$

$$R(U_1)=\{y \mid y=U_1 z\}=\{y \mid y=A(V_1 \Sigma^{-1} z)\}$$

$$\subset\{y \mid y=Ax\}=R(A)$$

故 $R(A)=R(U_1)=\mathrm{span}\{u_1,\cdots,u_r\}$

(3) $A=(u_1,\cdots,u_r)\begin{bmatrix}\sigma_1 & & \\ & \ddots & \\ & & \sigma_r\end{bmatrix}\begin{bmatrix}v_1^{\mathrm{H}} \\ \vdots \\ v_r^{\mathrm{H}}\end{bmatrix}$

$$=\sigma_1 u_1 v_1^{\mathrm{H}}+\cdots+\sigma_r u_r v_r^{\mathrm{H}}$$

四、正交相抵

$A_{m\times n}, B_{m\times n}$，若有酉矩阵 $U_{m\times m}$ 及 $V_{n\times n}$，使 $U^H AV = B$，

称A与B正交相抵。

性质：(1) A与A正交相抵；

(2) A与B正交相抵$\Rightarrow$$B$与$A$正交相抵；

(3) A与B正交相抵，B与C正交相抵A与C

正交相抵

定理18：A与B正交相抵 $\Rightarrow \sigma_A = \sigma_B$

证明： $B = U^H A V \Rightarrow B^H B = \cdots = V^{-1}\left(A^H A\right)V$

$$\Rightarrow \lambda_{B^H B} = \lambda_{A^H A} \geq 0$$

$$\Rightarrow \sigma_A = \sigma_B$$

例： $A^H = A \Rightarrow \sigma_A = \left|\lambda_A\right|$ $\qquad \because \lambda_{A^H A} = \lambda_{A^2} = \left(\lambda_A\right)^2$

$A^H = -A \Rightarrow \sigma_A = \left|\lambda_A\right|$ $\qquad \because \lambda_{A^H A} = \lambda_{(jA)^2} = \left(j\lambda_A\right)^2$

$A^H = -A \Rightarrow \lambda_A$ 为0或纯虚数，$j\lambda_A$ 为实数

矩阵分解的应用

设方程组 $A_{m\times n}x=b$ 有解，则有

(1) $m=n$: $A=LU\Rightarrow Ly=b,\ Ux=y$

(2) $m=n$: $A=QR\Rightarrow Rx=Q^{\mathrm{T}}b$

(3) $A=UDV^{\mathrm{H}}\Rightarrow Dy=U^{\mathrm{H}}b\overset{\mathrm{def}}{=}c,\ V^{\mathrm{H}}x=y$

$$D=\begin{bmatrix}\Sigma & O\\ O & O\end{bmatrix}_{m\times n},\quad \Sigma=\begin{bmatrix}\sigma_1 & & \\ & \ddots & \\ & & \sigma_r\end{bmatrix}$$

$$\begin{bmatrix}\Sigma & O\\ O & O\end{bmatrix}\begin{bmatrix}y_1\\ \vdots\\ y_n\end{bmatrix}=\begin{bmatrix}c_1\\ \vdots\\ c_m\end{bmatrix}$$ （隐含 $c_{r+1}=0,\cdots,c_m=0$）

通解为 $\begin{bmatrix} y_1 \\ \vdots \\ y_r \\ y_{r+1} \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} c_1/\sigma_1 \\ \vdots \\ c_r/\sigma_r \\ k_1 \\ \vdots \\ k_{n-r} \end{bmatrix}$ （$k_1,\cdots,k_{n-r}$ 是任意常数）

$$x = Vy = \left(\frac{c_1}{\sigma_1}v_1 + \cdots + \frac{c_r}{\sigma_r}v_r\right) + \left(k_1 v_{r+1} + \cdots + k_{n-r}v_n\right)$$

[注] $k_1 v_{r+1} + \cdots + k_{n-r}v_n$ 是 $A_{m\times n}x = 0$ 的通解

因为 $A\left(\frac{c_1}{\sigma_1}v_1 + \cdots + \frac{c_r}{\sigma_r}v_r\right) = AV_1\Sigma^{-1}\begin{bmatrix} c_1 \\ \vdots \\ c_r \end{bmatrix} = U_1\begin{bmatrix} c_1 \\ \vdots \\ c_r \end{bmatrix} = \left[U_1 \;\vdots\; U_2\right]c = b$

所以 $\frac{c_1}{\sigma_1}v_1 + \cdots + \frac{c_r}{\sigma_r}v_r$ 是 $A_{m\times n}x = b$ 的一个特解

作业

- P195 3、4
- P225 2、3、4、5
- P233 1、2、4

本讲主要内容

- 投影变换
- 广义逆的存在、性质及构造方法
- 广义逆矩阵的计算方法

投影矩阵

定义：向量空间C^n中，子空间L与M满足$C^n=L\oplus M$,

对 $\forall x\in C^n$, 分解式 $x=y+z, y\in L, z\in M$ 唯一。

称变换 $T_{L,M}(x)=y$ 为沿着M到L的投影

性质(1)：$T_{L,M}$ 是线性变换

性质(2)：$R(T_{L,M})=L, N(T_{L,M})=M$

性质(3)：$\forall x\in L\Rightarrow T_{L,M}(x)=x \quad \forall x\in M\Rightarrow T_{L,M}(x)=\theta$

[注] $T_{L,M}$ 是L中的单位变换

$T_{L,M}$ 是M中的零变换

二、投影矩阵

定义：取线性空间C^n的基为$e_1,e_2,\ldots,e_n$时，元素x与它的坐标“形式一致”。称 $T_{L,M}$ 在该基下的矩阵记为投影矩阵$P_{L,M}$

性质(4)：$T_{L,M}(x)=y \Leftrightarrow P_{L,M}x=y$

$$x\in L \;\Rightarrow T_{L,M}(x)=x \;\Rightarrow P_{L,M}x=x$$

$$x\in M \;\Rightarrow T_{L,M}(x)=\theta \;\Rightarrow P_{L,M}x=\theta$$

预备：$R(A)=\left\{y \mid y=Ax, x\in C^n\right\}, N(A)=\left\{x \mid Ax=0, x\in C^n\right\}$

引理1：$A_{n\times n}, A^2=A \Rightarrow N(A)=R(I-A)$

证明： $A^2=A \Rightarrow A(I-A)=O$

先证 $R(I-A)\subset N(A)$

$$\forall x\in R(I-A) \Rightarrow \exists u\in C^n, \text{st}. x=(I-A)u$$

$$Ax=A(I-A)u=\theta \quad \Rightarrow x\in N(A)$$

先证 $N(A)\subset R(I-A)$： $\forall \alpha\in N(A) \Rightarrow A\alpha=\theta$

$$\alpha=\alpha-A\alpha=(I-A)\alpha\in R(I-A)$$

故 $N(A)=R(I-A)$

定理1： $P_{n\times n}=P_{L,M}\Leftrightarrow P^2=P$

证明： 必要性 $C^n=L\oplus M$

$\forall x\in C^n, x=y+z, y\in L, z\in M$ 唯一 $\Rightarrow P_{L,M}x=y$

$P_{L,M}^2x=P_{L,M}\left(P_{L,M}x\right)=P_{L,M}y=y\ =P_{L,M}x$

充分性 $\forall x\in C^n\Rightarrow x=Px+(I-P)x$

令 $y=Px\in R(P),\ z=(I-P)x\in R(I-P)=N(P)$

则 $C^n=R(P)+N(P)$ ，下证 $R(P)\cap N(P)=\{\theta\}$:

对 $\forall\beta\in R(P)\cap N(P),\ \beta\in R(P)\Rightarrow\exists u\in C^n, \text{st. } \beta=Pu$

$\beta\in N(P)\Rightarrow P\beta=\theta$

故 $\beta=Pu=P^2u=PPu=P\beta=\theta$

于是可得 $C^n = R(P) \oplus N(P)$, 从而有

$$x = y + z, y = Px \in R(P), z = (I - P)x \in N(P)$$ 唯一

因为投影变换 $T_{R(P),N(P)}$满足

$$T_{R(P),N(P)}(x) = y \Rightarrow P_{R(P),N(P)}x = y \quad \left(\forall x \in C^n\right)$$

所以 $P = P_{R(P),N(P)}$

三、投影矩阵的确定方法

$\dim L = r, L$ 的基为 $x_1, \cdots, x_r : X = (x_1, \cdots, x_r)$

$\dim M = n - r, M$ 的基为 $y_1, \cdots, y_{n-r} : Y = (y_1, \cdots, y_{n-r})$

$$\left.\begin{aligned} P_{L,M} x_i = x_i \quad &\Rightarrow P_{L,M} X = X \\ P_{L,M} y_j = \theta \quad &\Rightarrow P_{L,M} Y = O \end{aligned}\right\} \Rightarrow P_{L,M}\left(X \mid Y\right) = \left(X \mid O\right)$$

$$\Rightarrow P_{L,M} = \left(X \mid O\right)\left(X \mid Y\right)^{-1}$$

例1：R^2 中：$\alpha_1=\begin{bmatrix}0\\1\end{bmatrix},\alpha_2=\begin{bmatrix}1\\-1\end{bmatrix},L=L(\alpha_1),M=L(\alpha_2),$

求 $P_{L,M}$

解：$P_{L,M}=\begin{bmatrix}0&0\\1&0\end{bmatrix}\begin{bmatrix}0&1\\1&-1\end{bmatrix}^{-1}=\begin{bmatrix}0&0\\1&1\end{bmatrix}$

例2：$P_{L,M}$与L和M的基的选择无关。

证：L的基 $x_1,\cdots,x_r$ ；另一基$\tilde{x}_1,\cdots,\tilde{x}_r$：

$$X=(x_1,\cdots,x_r),\ \tilde{X}=(\tilde{x}_1,\cdots,\tilde{x}_r)\ \Rightarrow \tilde{X}=XC_{r\times r}$$

M的基$y_1,\cdots,y_{n-r}$；另一基$\tilde{y}_1,\cdots,\tilde{y}_{n-r}$：

$$Y=(y_1,\cdots,y_{n-r}),\ \tilde{Y}=(\tilde{y}_1,\cdots,\tilde{y}_{n-r})\ \Rightarrow \tilde{Y}=YD_{(n-r)\times(n-r)}$$

$$(\tilde{X}\,|\,O)\cdot(\tilde{X}\,|\,\tilde{Y})^{-1}=(XC\,|\,O)\cdot\left[(X\,|\,Y)\begin{pmatrix}C & O\\ O & D\end{pmatrix}\right]^{-1}$$

$$=(XC\,|\,O)\cdot\begin{bmatrix}C^{-1} & O\\ O & D^{-1}\end{bmatrix}\cdot(X\,|\,Y)^{-1}=(X\,|\,O)\cdot(X\,|\,Y)^{-1}$$

四、正交投影变换

欧氏空间 C^n 中，子空间L给定，取 $M = L^{\perp}$，

则 $\mathbf{C}^n = L \oplus M$

正交投影变换$T_L = T_{L,M}$; 正交投影矩阵 $P_L = P_{L,M}$

定理2：方阵 $P = P_L \Leftrightarrow P^2 = P, P^H = P$

四、正交投影矩阵的确定方法

$$\left.\begin{array}{l} L\text{的基为}\quad x_1,\cdots,x_r:\quad X=(x_1,\cdots,x_r) \\ L^{\perp}\quad\text{的基为}\ y_1,\cdots,y_{n-r}:Y=(y_1,\cdots,y_{n-r}) \end{array}\right\}\Rightarrow\begin{cases} X^H Y=O \\ Y^H X=O \end{cases}$$

已求得 $P_L=P_{L,L^{\perp}}=(X|O)\cdot(X|Y)^{-1}$

因为 $(X|Y)^H\cdot(X\mid Y)=\begin{pmatrix} X^H \\ Y^H \end{pmatrix}\cdot(X\mid Y)=\begin{bmatrix} X^H X & O \\ O & Y^H Y \end{bmatrix}$

所以 $(X\mid Y)^{-1}=\begin{bmatrix} (X^H X)^{-1} & O \\ O & (Y^H Y)^{-1} \end{bmatrix}\cdot(X\mid Y)^H=\begin{bmatrix} (X^H X)^{-1} X^H \\ (Y^H Y)^{-1} Y^H \end{bmatrix}$

于是 $P_L=(X|O)\cdot\begin{bmatrix} (X^H X)^{-1} X^H \\ (Y^H Y)^{-1} Y^H \end{bmatrix}=X\cdot(X^H X)^{-1}\cdot X^H$

例3：向量空间 R^3 中，$\alpha=\begin{bmatrix}1\\2\\0\end{bmatrix},\beta=\begin{bmatrix}0\\1\\1\end{bmatrix},L=L(\alpha,\beta)$，

求 P_L

解：$X=\begin{bmatrix}1&0\\2&1\\0&1\end{bmatrix},X^TX=\begin{bmatrix}5&2\\2&2\end{bmatrix},\left(X^TX\right)^{-1}=\frac{1}{6}\begin{bmatrix}2&-2\\-2&5\end{bmatrix}$

$$P_L=X\left(X^TX\right)^{-1}X^T=\frac{1}{6}\begin{bmatrix}2&2&-2\\2&5&1\\-2&1&5\end{bmatrix}$$

[注]：正交投影矩阵 P_L 与子空间 L 的基的选择无关

广义逆矩阵

一、定义与算法

定义：对 $A_{m\times n}$, 若有 $X_{n\times m}$ 满足Penrose方程

（1） $AXA=A$ （2） $XAX=X$

（3） $(AX)^H=AX$ （4） $(XA)^H=XA$

称X为A的M-P逆，记作A^+.(Moore 1920，Penrose1955)

例如 $A_{m\times n}$ 可逆, $X=A^{-1}$ 满足P-方程：$A^+=A^{-1}$

$A=O_{m\times n}, X=O_{n\times m}$ 满足P-方程：$O^+_{m\times n}=O_{n\times m}$

$A=\begin{bmatrix}1\\1\end{bmatrix}, X=\begin{bmatrix}\frac{1}{2} & \frac{1}{2}\end{bmatrix}$ 满足P-方程：$A^+=\begin{bmatrix}\frac{1}{2} & \frac{1}{2}\end{bmatrix}$

例4：$F \in C_r^{m\times r}\ (r \ge 1) \Rightarrow F^{+} = \left(F^{H}F\right)^{-1}F^{H}$，且$F^{+}F = I_r$

$G \in C_r^{r\times n}\ (r \ge 1) \Rightarrow G^{+} = G^{H}\left(GG^{H}\right)^{-1}$，且$GG^{+} = I_r$

验证第一式：令 $F^{+} = \left(F^{H}F\right)^{-1}F^{H}$, 则有

$$FXF = F\left(F^{H}F\right)^{-1}F^{H}F = F$$

$$XFX = \left(F^{H}F\right)^{-1}F^{H}FX = X$$

$$\left(FX\right)^{H} = X^{H}F^{H} = F\left(F^{H}F\right)^{-1}F^{H} = FX$$

$$\left(XF\right)^{H} = I_r^{H} = I_r = XF$$

定理3：$\forall A_{m\times n}, A^{+}$ 存在并唯一

证明： 存在性 $A = O_{m\times n} \Rightarrow A^{+} = O_{n\times m}$

$A \neq O \Rightarrow \mathrm{rank}A \geq 1$：　$A = FG, F \in \mathrm{C}_r^{m\times r}, G \in \mathrm{C}_r^{r\times n}$

令 $X = G^{+}F^{+}$ 则有

$$AXA = FG\cdot G^{+}F^{+}\cdot FG = FG = A$$

$$XAX = G^{+}F^{+}\cdot FG\cdot G^{+}F^{+} = G^{+}F^{+} = X$$

$$(AX)^{\mathrm{H}} = \left(FG\cdot G^{+}F^{+}\right)^{\mathrm{H}} = \left(FF^{+}\right)^{\mathrm{H}} = FF^{+} = F\cdot GG^{+}\cdot F^{+} = AX$$

$$(XA)^{\mathrm{H}} = \left(G^{+}F^{+}\cdot FG\right)^{\mathrm{H}} = \left(G^{+}G\right)^{\mathrm{H}} = G^{+}G = G^{+}\cdot F^{+}F\cdot G = XA$$

$$A^{+} = G^{+}F^{+} = G^{\mathrm{H}}\left(F^{\mathrm{H}}AG^{\mathrm{H}}\right)^{-1}F^{\mathrm{H}}$$

定理3：$\forall A_{m\times n}, A^{+}$ 存在并唯一

证明：唯一性，对$A_{m\times n}$若$X_{n\times m}$与$Y_{n\times m}$都满足P-方程，

则：

$$X = XAX = X\cdot AYA\cdot X = X\cdot (AY)^{H}\cdot (AX)^{H}$$

$$= X\cdot (AXAY)^{H} = X\cdot (AY)^{H} = XAY = X\cdot AYA\cdot Y$$

$$= (XA)^{H}\cdot (YA)^{H}\cdot Y = (YAXA)^{H}\cdot Y = (YA)^{H}\cdot Y = YAY = Y$$

例5：设$A\in C_r^{m\times n}$的奇异值分解为 $A=U\begin{bmatrix}\Sigma_r & O\\ O & O\end{bmatrix}_{m\times n}V^H$

则 $A^{+}=V\begin{bmatrix}\Sigma_r^{-1} & O\\ O & O\end{bmatrix}_{n\times m}U^H$

直接验证即可。

进一步的有

$$A=\left(U_s,U_n\right)\begin{bmatrix}\Sigma_r & O\\ O & O\end{bmatrix}_{m\times n}\left(V_s,V_n\right)^H$$

$$=U_s\Sigma_r V_s^H$$

$$A^{+}=V_s\Sigma_r^{-1}U_s^H$$

广义逆矩阵的分类：对 $A_{m\times n}$ ，若 $X_{n\times m}$ 满足P-方程

(i): 称 X 为 A 的 $\{i\}$-逆，记作 $A^{(i)}$．全体记作 $A\{i\}$

(i),(j): 称 X 为 A 的 $\{i,j\}$-逆，记作 $A^{(i,j)}$．全体记作 $A\{i,j\}$

(i),(j),(k): 称 X 为 A 的 $\{i,j,k\}$-逆，记作 $A^{(i,j,k)}$．全体记作 $A\{i,j,k\}$

(1)～(4)：则X为A^+

合计：15类

常用广义逆矩阵：$A\{1\},A\{1,2\},A\{1,3\},A\{1,4\},A^+$

求 $A^{(1)}, A^{(1,2)}$ 的初等变换方法

$$A \in \mathbb{C}_r^{m\times n}, \quad A \xrightarrow{\text{行}} B \Rightarrow \exists \text{可逆矩阵} Q_{m\times m}, \text{st}. QA = B$$

其中B为拟Hermite标准形, 它的后 $m - r$ 行元素全为零

$$B \xrightarrow{\text{列对换}} \begin{bmatrix} I_r & K \\ O & O \end{bmatrix}_{m\times n} = C \quad \Rightarrow \exists \text{置换矩阵} P_{n\times n}, \text{st}. BP = C$$

于是 $QAP = C \Rightarrow A = Q^{-1}CP^{-1}$

定理14：已知A，P，Q如上所述, 对 $\forall L_{(n-r)\times(m-r)}$，有

$$X = P\begin{bmatrix} I_r & O \\ O & L \end{bmatrix}_{n\times m} Q \in A\{1\},\quad X_0 = P\begin{bmatrix} I_r & O \\ O & O \end{bmatrix}_{n\times m} Q \in A\{1,2\}$$

证明：略

例6：$A=\begin{bmatrix}2&1&0&2\\0&0&1&2\\2&1&1&4\end{bmatrix}$,　　求 $A^{(1)},A^{(1,2)},A^{+}$

解：

$$(A|\ I)\to\left[\begin{array}{cccc|ccc}2&1&0&2&1&0&0\\0&0&1&2&0&1&0\\0&0&0&0&-1&-1&1\end{array}\right]:\ c_1=2,c_2=3$$

$$Q=\begin{bmatrix}1&0&0\\0&1&0\\-1&-1&1\end{bmatrix},\quad P=(e_2,e_3,e_1,e_4)=\begin{bmatrix}0&0&1&0\\1&0&0&0\\0&1&0&0\\0&0&0&1\end{bmatrix}$$

$$A^{(1)}=P\left[\begin{array}{cc|c}1&0&0\\0&1&0\\\hline 0&0&a\\0&0&b\end{array}\right]Q=\begin{bmatrix}-a&-a&a\\1&0&0\\0&1&0\\-b&-b&b\end{bmatrix},\quad A^{(1,2)}=\begin{bmatrix}0&0&0\\1&0&0\\0&1&0\\0&0&0\end{bmatrix}$$

$A = FG$:

$$F = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}, F^{\mathrm{T}} F = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}, F^{+} = \frac{1}{3}\begin{bmatrix} 2 & -1 & 1 \\ -1 & 2 & 1 \end{bmatrix}$$

$$G = \begin{bmatrix} 2 & 1 & 0 & 2 \\ 0 & 0 & 1 & 2 \end{bmatrix}, GG^{\mathrm{T}} = \begin{bmatrix} 9 & 4 \\ 4 & 5 \end{bmatrix}, G^{+} = \frac{1}{29}\begin{bmatrix} 10 & -8 \\ 5 & -4 \\ -4 & 9 \\ 2 & 10 \end{bmatrix}$$

$$A^{+} = G^{+} F^{+} = \frac{1}{87}\begin{bmatrix} 28 & -26 & 2 \\ 14 & -13 & 1 \\ -17 & 22 & 5 \\ -6 & 18 & 12 \end{bmatrix}$$

例7：$A_{m\times n}\neq O$，且 A^+ 已知，记 $B=\begin{bmatrix}A\\A\end{bmatrix}$，求 B^+

解：$\mathrm{rank}A=r\geq 1\Rightarrow A=FG:\quad F\in\mathbf{C}_r^{m\times r},\ G\in\mathbf{C}_r^{r\times n}$

$$B=\begin{pmatrix}FG\\FG\end{pmatrix}=\begin{pmatrix}F\\F\end{pmatrix}G:\quad \begin{pmatrix}F\\F\end{pmatrix}\in\mathbf{C}_r^{2m\times r},G\in\mathbf{C}_r^{r\times n}$$

$$B^+=G^+\begin{pmatrix}F\\F\end{pmatrix}^+=G^+\cdot\left[\left(F^{\mathrm{H}}\middle|F^{\mathrm{H}}\right)\begin{pmatrix}F\\F\end{pmatrix}\right]^{-1}\left(F^{\mathrm{H}}\middle|F^{\mathrm{H}}\right)$$

$$=G^+\cdot\frac{1}{2}\left(F^{\mathrm{H}}F\right)^{-1}\cdot\left(F^{\mathrm{H}}\middle|F^{\mathrm{H}}\right)=\frac{1}{2}\left(G^+F^+\middle|G^+F^+\right)=\frac{1}{2}\left(A^+\middle|A^+\right)$$

二、广义逆矩阵的性质

定理4：$A_{m\times n}, A^{(1)}$ 唯一 $\Leftrightarrow m=n, A$ 可逆，且 $A^{(1)}=A^{-1}$

定理5：$A_{m\times n}, B_{n\times p}, \lambda\in \mathbf{C}, \lambda^{+}=\begin{cases}1/\lambda & (\lambda\neq 0)\\ 0 & (\lambda=0)\end{cases}$

(1) $\left[A^{(1)}\right]^{\mathrm{H}}\in A^{\mathrm{H}}\{1\}$： $AA^{(1)}A=A\Rightarrow A^{\mathrm{H}}(A^{(1)})^{\mathrm{H}}A^{\mathrm{H}}=A^{\mathrm{H}}$

(2) $\lambda^{+}A^{(1)}\in(\lambda A)\{1\}$： $(\lambda A)(\lambda^{+}A^{(1)})(\lambda A)=(\lambda\lambda^{+}\lambda)(AA^{(1)}A)=\lambda A$

(3) $S_{m\times m}$ 和 $T_{n\times n}$ 都可逆 $\Rightarrow T^{-1}A^{(1)}S^{-1}\in(SAT)\{1\}$

(4) $r_{A}\le r_{A^{(1)}}$： $r_{A}=r_{AA^{(1)}A}\le r_{A^{(1)}}$

(5) $AA^{(1)}$ 与 $A^{(1)}A$ 都是幂等矩阵，且 $r_{AA^{(1)}} = r_A = r_{A^{(1)}A}$

因为 $r_A = r_{AA^{(1)}A} \le \begin{cases} r_{AA^{(1)}} \\ r_{A^{(1)}A} \end{cases} \le r_A$

(6) $R(AA^{(1)}) = R(A)$: $R(A) = R(AA^{(1)}A) \subset R(AA^{(1)}) \subset R(A)$

$N(A^{(1)}A) = N(A)$: $N(A) \subset N(A^{(1)}A) \subset N(AA^{(1)}A) = N(A)$

(7) ① $A^{(1)}A = I_n \Leftrightarrow r_A = n$ “A 列满秩”

② $AA^{(1)} = I_m \Leftrightarrow r_A = m$ “A 行满秩”

(8) ① $(AB)(AB)^{(1)}A = A \Leftrightarrow r_{AB} = r_A$

② $B(AB)^{(1)}(AB) = B \Leftrightarrow r_{AB} = r_B$

定理6：$A_{m\times n}, Y \in A\{1\}, \; Z \in A\{1\} \Rightarrow X \triangleq YAZ \in A\{1,2\}$

证明：$AXA = A \cdot YAZ \cdot A = AY(AZA) = AYA = A$

$$XAX = YAZ \cdot A \cdot YAZ = Y(AZA)YAZ = Y \cdot AYA \cdot Z = YAZ = X$$

推论 $A_{m\times n}, Y \in A\{1\} \Rightarrow X \triangleq YAY \in A\{1,2\}$

定理7：设 $X \in A\{1\}$，则 $r_X = r_A \Leftrightarrow X \in A\{1,2\}$

定理8：$Y \triangleq \left(A^{\mathrm{H}} A\right)^{(1)} A^{\mathrm{H}} \in A\{1,2,3\}, \quad Z \triangleq A^{\mathrm{H}}\left(AA^{\mathrm{H}}\right)^{(1)} \in A\{1,2,4\}$

定理9：$A^{+}=A^{(1,4)}AA^{(1,3)}$

定理10：(1) $r_{A^{+}}=r_{A}$ (2) $\left(A^{+}\right)^{+}=A$

(3) $\left(A^{H}\right)^{+}=\left(A^{+}\right)^{H}$ $\left(A^{T}\right)^{+}=\left(A^{+}\right)^{T}$

(4) $\left(A^{H}A\right)^{+}=A^{+}\left(A^{H}\right)^{+}$ $\left(AA^{H}\right)^{+}=\left(A^{H}\right)^{+}A^{+}$

(5) $A^{+}=\left(A^{H}A\right)^{+}A^{H}=A^{H}\left(AA^{H}\right)^{+}$

(6) $R\left(A^{+}\right)=R\left(A^{H}\right), N\left(A^{+}\right)=N\left(A^{H}\right)$

二、M-P逆的等价定义

Moore逆：对 $A_{m\times n}$, 若有 $X_{n\times m}$ 满足 $AX = P_{R(A)}$ 和 $XA = P_{R(X)}$ ，称X为A的Moore逆。

定理11：M-逆与P-逆等价

作业

- P295：1、2、3
- P306：7、8、9