矩 阵 论 讲 稿

讲稿编者： 张 凯 院

使用教材：《矩阵论》（第 2 版）

西北工业大学出版社

程云鹏 等编

辅助教材：《矩阵论导教导学导考》

《矩阵论典型题解析及自测试题》

西北工业大学出版社

张凯院 等编

课时分配：第一章 **17** 学时　　第四章 **8** 学时

第二章 **5** 学时　　第五章 **8** 学时

第三章 **8** 学时　　第六章 **8** 学时

第一章　线性空间与线性变换

§ 1.1　线性空间

一、集合与映射

1．集合：能够作为整体看待的一堆东西．

列举法：$S=\{a_1,a_2,a_3,\cdots\}$

性质法：$S=\{\,a\,|\,a\text{ 所具有的性质 }\}$

相等$(S_1=S_2)$：指下面二式同时成立

$$\forall a\in S_1\Rightarrow a\in S_2,\ \text{即}\ S_1\subseteq S_2$$

$$\forall b\in S_2\Rightarrow b\in S_1,\ \text{即}\ S_2\subseteq S_1$$

交：$S_1\cap S_2=\{a\,|\,a\in S_1\text{ 且 }a\in S_2\}$

并：$S_1\cup S_2=\{a\,|\,a\in S_1\text{ 或 }a\in S_2\}$

和：$S_1+S_2=\{a=a_1+a_2\ |\,a_1\in S_1,a_2\in S_2\}$

例 1　$S_1=\{A=\begin{bmatrix}a_{11}&0\\a_{21}&a_{22}\end{bmatrix}\,|\,a_{ij}\in\mathbf{R}\}$

$S_2=\{A=\begin{bmatrix}a_{11}&a_{12}\\0&a_{22}\end{bmatrix}\,|\,a_{ij}\in\mathbf{R}\}$，$S_1\neq S_2$

$S_1\cap S_2=\{A=\begin{bmatrix}a_{11}&0\\0&a_{22}\end{bmatrix}\,|\,a_{11},a_{22}\in\mathbf{R}\}$

$S_1\cup S_2=\{A=\begin{bmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{bmatrix}\,|\,a_{12}a_{21}=0,\,a_{ij}\in\mathbf{R}\}$

$S_1+S_2=\{A=\begin{bmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{bmatrix}\,|\,a_{ij}\in\mathbf{R}\}$

2．数域：关于四则运算封闭的数的集合．

例如：实数域$\mathbf{R}$，复数域$\mathbf{C}$，有理数域$\mathbf{Q}$，等等．

3．映射：设集合S_1与S_2，若对任意的$a\in S_1$，按照法则σ，对应唯一的

$b \in S_2$, 记作 $\sigma(a)=b$. 称 σ 为由 S_1 到 S_2 的映射；称 b 为 a 的象，a 为 b 的象源.

变换：当 $S_1 = S_2$ 时，称映射 σ 为 S_1 上的变换.

例 2　$S=\{A=(a_{ij})_{n\times n} \mid a_{ij}\in \mathrm{R}\}\quad (n\geq 2)$.

映射 σ_1：$\sigma_1(A)=\det A \qquad (S\to \mathrm{R})$

变换 σ_2：$\sigma_2(A)=(\det A)I_n \qquad (S\to S)$

二、线性空间及其性质

1. 线性空间：集合 V 非空，给定数域 K，若在 V 中

(Ⅰ) 定义的加法运算封闭，即

$\forall\, x, y\in V$, 对应唯一元素 $(x+y)\in V$，且满足

(1) 结合律：$x+(y+z)=(x+y)+z \quad (\forall z\in V)$

(2) 交换律：$x+y=y+x$

(3) 有零元：$\exists\,\theta\in V$, 使得 $x+\theta=x \quad (\forall x\in V)$

(4) 有负元：$\forall x\in V, \exists\,(-x)\in V$, 使得 $x+(-x)=\theta$.

(Ⅱ) 定义的数乘运算封闭，即

$\forall\, x\in V, \forall\, k\in K$, 对应唯一 元素 $(kx)\in V$，且满足

(5) 数对元素分配律：$k(x+y)=kx+ky \quad (\forall y\in V)$

(6) 元素对数分配律：$(k+l)x=kx+lx \quad (\forall l\in K)$

(7) 数因子结合律：$k(lx)=(kl)x \quad (\forall l\in K)$

(8) 有单位数：单位数 $1\in K$, 使得 $1x=x$.

则称 V 为 K 上的线性空间.

例 3　$K=\mathrm{R}$ 时，R^n—向量空间；　　$\mathrm{R}^{m\times n}$—矩阵空间

$P_n[t]$—多项式空间；　$C[a,b]$—函数空间

$K=\mathbf{C}$时，$\mathbf{C}^n$—复向量空间；　$\mathbf{C}^{m\times n}$—复矩阵空间

例 4　集合$\mathbf{R}^+=\{m \mid m是正实数\}$，数域$\mathbf{R}=\{k \mid k是实数\}$.

加法：$m,n\in\mathbf{R}^+$，$m\oplus n=mn$

数乘：$m\in\mathbf{R}^+, k\in\mathbf{R}$，$k\otimes m=m^k$

验证$\mathbf{R}^+$是$\mathbf{R}$上的线性空间.

证　加法封闭，且(1)～(2)成立.

(3)　$m\oplus\theta=m\Rightarrow m\theta=m\Rightarrow\theta=1$

(4)　$m\oplus(-m)=\theta\Rightarrow m(-m)=1\Rightarrow(-m)=1/m$

数乘封闭，(5)～(8)成立．故$\mathbf{R}^+$是$\mathbf{R}$上的线性空间.

例 5　集合$\mathbf{R}^2=\{\alpha=(\xi_1,\xi_2)\mid\xi_i\in\mathbf{R}\}$，数域$\mathbf{R}$．设$\beta=(\eta_1,\eta_2),k\in\mathbf{R}$.

运算方式 1　加法：　$\alpha+\beta=(\xi_1+\eta_1,\xi_2+\eta_2)$

数乘：　$k\alpha=(k\xi_1,k\xi_2)$

运算方式 2　加法：　$\alpha\oplus\beta=(\xi_1+\eta_1,\xi_2+\eta_2+\xi_1\eta_1)$

数乘：　$k\circ\alpha=(k\xi_1,k\xi_2+\frac{1}{2}k(k-1)\xi_1^2)$

可以验证$\mathbf{R}^2(+\cdot)$与$\mathbf{R}^2(\oplus\circ)$都是$\mathbf{R}$上的线性空间.

[注]　在$\mathbf{R}^2(\oplus\circ)$中，$\theta=(0,0)$，$-\alpha=(-\xi_1,-\xi_2+\xi_1^2)$.

Th1　线性空间V中的零元素唯一，负元素也唯一.

证　设θ_1与θ_2都是V的零元素，则$\theta_1=\theta_1+\theta_2=\theta_2+\theta_1=\theta_2$

设x_1与x_2都是x的负元素，则由$x+x_1=\theta$及$x+x_2=\theta$可得

$$
\begin{aligned}
x_1&=x_1+\theta=x_1+(x+x_2)=(x_1+x)+x_2\\
&=(x+x_1)+x_2=\theta+x_2=x_2+\theta=x_2
\end{aligned}
$$

例 6 在线性空间 V 中，下列结论成立.

$\mathbf{0}\boldsymbol{x}=\theta$： $\mathbf{1}\boldsymbol{x}+\mathbf{0}\boldsymbol{x}=(\mathbf{1}+\mathbf{0})\boldsymbol{x}=\mathbf{1}\boldsymbol{x}\Rightarrow\mathbf{0}\boldsymbol{x}=\theta$

$k\theta=\theta$： $k\boldsymbol{x}+k\theta=k(\boldsymbol{x}+\theta)=k\boldsymbol{x}\Rightarrow k\theta=\theta$

$(-1)\boldsymbol{x}=(-\boldsymbol{x})$： $(-1)\boldsymbol{x}=(-1)\boldsymbol{x}+[\boldsymbol{x}+(-\boldsymbol{x})]=[(-1)\boldsymbol{x}+1\boldsymbol{x}]+(-\boldsymbol{x})=(-\boldsymbol{x})$

2．减法运算：线性空间 V 中， $\boldsymbol{x}-\boldsymbol{y}=\boldsymbol{x}+(-\boldsymbol{y})$.

3．线性组合： $\boldsymbol{x},\boldsymbol{x}_i\in V$, *若存在* $c_i\in K$，使 $\boldsymbol{x}=c_1\boldsymbol{x}_1+\cdots+c_m\boldsymbol{x}_m$，则称 $\boldsymbol{x}$

是 $\boldsymbol{x}_1,\cdots,\boldsymbol{x}_m$ 的线性组合，或者 $\boldsymbol{x}$ 可由 $\boldsymbol{x}_1,\cdots,\boldsymbol{x}_m$ 线性表示.

4．线性相关：若有 $c_1,\cdots,c_m$ 不全为零，使得 $c_1\boldsymbol{x}_1+\cdots+c_m\boldsymbol{x}_m=\theta$，则称

$\boldsymbol{x}_1,\cdots,\boldsymbol{x}_m$ 线性相关.

5．线性无关：仅当 $c_1,\cdots,c_m$ 全为零时，才有 $c_1\boldsymbol{x}_1+\cdots+c_m\boldsymbol{x}_m=\theta$，则称

$\boldsymbol{x}_1,\cdots,\boldsymbol{x}_m$ 线性无关.

[注] 在 $\mathbf{R}^2(\oplus\circ)$ 中，$\alpha_1=(1,1)$，$\alpha_2=(2,2)$ 线性无关；

$\alpha_1=(1,1)$，$\alpha_2=(2,3)$ 线性相关.　（自证）

三、基与坐标

1．基与维数：线性空间 V 中，若元素组 $\boldsymbol{x}_1,\cdots,\boldsymbol{x}_n$ 满足

(1) $\boldsymbol{x}_1,\cdots,\boldsymbol{x}_n$ 线性无关；

(2) $\forall\ \boldsymbol{x}\in V$ 都可由 $\boldsymbol{x}_1,\cdots,\boldsymbol{x}_n$ 线性表示.

称 $\boldsymbol{x}_1,\cdots,\boldsymbol{x}_n$ 为 V 的一个基，n 为 V 的维数，记作 $\dim V=n$，或者 V^n.

例 7　矩阵空间 $\mathbf{R}^{m\times n}$ 中，易见

(1) $E_{ij}\ (i=1,2,\cdots,m;\ j=1,2,\cdots,n)$ 线性无关；

(2) $A=(a_{ij})_{m\times n}=\sum_{i=1}^{m}\sum_{j=1}^{n}a_{ij}E_{ij}$.

故 $E_{ij}\ (i=1,2,\cdots,m;\ j=1,2,\cdots,n)$ 是 $\mathbf{R}^{m\times n}$ 的一个基，$\dim\mathbf{R}^{m\times n}=mn$.

2． 坐标：给定线性空间 V^n 的基 $x_1,\cdots,x_n$，当 $x\in V^n$ 时，有

$x=\xi_1x_1+\cdots+\xi_nx_n$．称 $\xi_1,\cdots,\xi_n$ 为 x 在给定基 $x_1,\cdots,x_n$ 下的坐标，记作列向量 $\alpha=(\xi_1,\cdots,\xi_n)^{\mathrm T}$．

例 8　矩阵空间 $\mathbf{R}^{2\times 2}$ 中，设 $A=(a_{ij})_{2\times 2}$．

(1)　取基 $E_{11},E_{12},E_{21},E_{22}$，$A=a_{11}E_{11}+a_{12}E_{12}+a_{21}E_{21}+a_{22}E_{22}$

坐标为 $\alpha=(a_{11},a_{12},a_{21},a_{22})^{\mathrm T}$

(2)　取基 $B_1=\begin{bmatrix}1&1\\1&1\end{bmatrix}$，$B_2=\begin{bmatrix}0&1\\1&1\end{bmatrix}$，$B_3=\begin{bmatrix}0&0\\1&1\end{bmatrix}$，$B_4=\begin{bmatrix}0&0\\0&1\end{bmatrix}$

$$\begin{aligned}A&=a_{11}(B_1-B_2)+a_{12}(B_2-B_3)+a_{21}(B_3-B_4)+a_{22}B_4\\&=a_{11}B_1+(a_{12}-a_{11})B_2+(a_{21}-a_{12})B_3+(a_{22}-a_{21})B_4\end{aligned}$$

坐标为 $\beta=(a_{11},a_{12}-a_{11},a_{21}-a_{12},a_{22}-a_{21})^{\mathrm T}$

[注]　一个元素在两个不同的基下的坐标可能相同，也可能不同.

例如：$A=E_{22}$ 在上述两个基下的坐标都是 $(0,0,0,1)^{\mathrm T}$；

$A=E_{11}$ 在上述两个基下的坐标不同.

Th2　线性空间 V^n 中，元素在给定基下的坐标唯一.

证　设 V^n 的基为 $x_1,\cdots,x_n$，对于 $x\in V^n$，若

$$x=\xi_1x_1+\cdots+\xi_nx_n=\eta_1x_1+\cdots+\eta_nx_n$$

则有　$(\xi_1-\eta_1)x_1+\cdots+(\xi_n-\eta_n)x_n=\theta$

因为 $x_1,\cdots,x_n$ 线性无关，所以 $\xi_i-\eta_i=0$，即 $\xi_i=\eta_i\ (i=1,2,\cdots,n)$.

故 x 的坐标唯一.

例 9　设线性空间 V^n 的基为 $x_1,\cdots,x_n$，元素 y_j 在该基下的坐标为 $\alpha_j\ (j=1,2,\cdots,m)$，则元素组 $y_1,\cdots,y_m$ 线性相关（线性无关）$\Leftrightarrow$ 向量组 $\alpha_1,\cdots,\alpha_m$ 线性相关（线性无关）.

证　对于数组$k_1,\cdots,k_m$，因为

$$k_1y_1+\cdots+k_my_m=(x_1,\cdots,x_n)(k_1\alpha_1+\cdots+k_m\alpha_m)=\theta$$

等价于$k_1\alpha_1+\cdots+k_m\alpha_m=\theta$，所以结论成立.

四、基变换与坐标变换

1．基变换：设线性空间V^n的基(Ⅰ)为$x_1,\cdots,x_n$，基(Ⅱ)为$y_1,\cdots,y_n$，则

$$\begin{cases} y_1=c_{11}x_1+c_{21}x_2+\cdots+c_{n1}x_n \\ y_2=c_{12}x_1+c_{22}x_2+\cdots+c_{n2}x_n \\ \cdots\cdots\cdots \\ y_n=c_{1n}x_1+c_{2n}x_2+\cdots+c_{nn}x_n \end{cases} \qquad C=\begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{bmatrix}$$

写成矩阵乘法形式为　$(y_1,\cdots,y_n)=(x_1,\cdots,x_n)C$

称上式为基变换公式，C为由基(Ⅰ)改变为基(Ⅱ)的过渡矩阵.

[注] 过渡矩阵C一定可逆. 否则C的n个列向量线性相关，从而$y_1,\cdots,y_n$线性相关（例 9）. 矛盾！由此可得

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

称C^{-1}为由基(Ⅱ)改变为基(Ⅰ)的过渡矩阵.

2．坐标变换：设$x\in V^n$在两个基下的坐标分别为α和β，则有

$$x=\xi_1x_1+\cdots+\xi_nx_n=(x_1,\cdots,x_n)\alpha$$

$$x=\eta_1y_1+\cdots+\eta_ny_n=(y_1,\cdots,y_n)\beta=(x_1,\cdots,x_n)C\beta$$

由定理 2 可得$\alpha=C\beta$，或者$\beta=C^{-1}\alpha$，称为坐标变换公式.

例 10　矩阵空间$\mathbf{R}^{2\times 2}$中，取基

(Ⅰ) $A_1=\begin{bmatrix}1&0\\0&1\end{bmatrix}$, $A_2=\begin{bmatrix}1&0\\0&-1\end{bmatrix}$, $A_3=\begin{bmatrix}0&1\\1&0\end{bmatrix}$, $A_4=\begin{bmatrix}0&1\\-1&0\end{bmatrix}$

(Ⅱ) $B_1=\begin{bmatrix}1&1\\1&1\end{bmatrix}$, $B_2=\begin{bmatrix}1&1\\1&0\end{bmatrix}$, $B_3=\begin{bmatrix}1&1\\0&0\end{bmatrix}$, $B_4=\begin{bmatrix}1&0\\0&0\end{bmatrix}$

(1) 求由基(Ⅰ)改变为基(Ⅱ)的过渡矩阵；

(2) 求由基(Ⅱ)改变为基(Ⅰ)的坐标变换公式.

解　采用中介法求过渡矩阵.

基(0)：$E_{11}=\begin{bmatrix}1&0\\0&0\end{bmatrix}$，$E_{12}=\begin{bmatrix}0&1\\0&0\end{bmatrix}$，$E_{21}=\begin{bmatrix}0&0\\1&0\end{bmatrix}$，$E_{22}=\begin{bmatrix}0&0\\0&1\end{bmatrix}$

(0)→(Ⅰ)：$(A_1,A_2,A_3,A_4)=(E_{11},E_{12},E_{21},E_{22})C_1$

(0)→(Ⅱ)：$(B_1,B_2,B_3,B_4)=(E_{11},E_{12},E_{21},E_{22})C_2$

$$C_1=\begin{bmatrix}1&1&0&0\\0&0&1&1\\0&0&1&-1\\1&-1&0&0\end{bmatrix},\quad C_2=\begin{bmatrix}1&1&1&1\\1&1&1&0\\1&1&0&0\\1&0&0&0\end{bmatrix}$$

(Ⅰ)→(Ⅱ)：$(B_1,B_2,B_3,B_4)=(A_1,A_2,A_3,A_4)C_1^{-1}C_2$

$$C=C_1^{-1}C_2=\frac{1}{2}\begin{bmatrix}1&0&0&1\\1&0&0&-1\\0&1&1&0\\0&1&-1&0\end{bmatrix}C_2=\frac{1}{2}\begin{bmatrix}2&1&1&1\\0&1&1&1\\2&2&1&0\\0&0&1&0\end{bmatrix}$$

$$\begin{bmatrix}\xi_1\\\xi_2\\\xi_3\\\xi_4\end{bmatrix}=C\begin{bmatrix}\eta_1\\\eta_2\\\eta_3\\\eta_4\end{bmatrix}=\frac{1}{2}\begin{bmatrix}2\eta_1+\eta_2+\eta_3+\eta_4\\\eta_2+\eta_3+\eta_4\\2\eta_1+2\eta_2+\eta_3\\\eta_3\end{bmatrix}$$

五、线性子空间

1．定义：线性空间V中，若子集V_1非空，且对V中的线性运算封闭，即

(1) $\forall\ x,y\in V_1\Rightarrow x+y\in V_1$

(2) $\forall\ x\in V_1,\forall\ k\in K\Rightarrow kx\in V_1$

称V_1为V的线性子空间，简称为子空间.

[注] (1) 子空间 V_1 也是线性空间，而且 $\dim V_1 \le \dim V$.

(2) $\{\theta\}$ 是 V 的线性子空间，规定 $\dim\{\theta\} = 0$.

(3) 子空间 V_1 的零元素就是 V 的零元素.

例 11　线性空间 V 中，子集 V_1 是 V 的子空间 $\Leftrightarrow$

对 $\forall\, x, y \in V_1, \forall\, k, l \in K$, 有 $kx + ly \in V_1$.

证　充分性. $k = l = 1$： $\forall\, x, y \in V_1 \Rightarrow x + y \in V_1$

$l = 0$： $\forall\, x \in V_1, \forall\, k \in K \Rightarrow kx = kx + 0y \in V_1$

故 V_1 是 V 的子空间.

必要性. $\forall\, x \in V_1, \forall\, k \in K \Rightarrow kx \in V_1$　　（数乘封闭）

$\forall\, y \in V_1, \forall\, l \in K \Rightarrow ly \in V_1$　　（数乘封闭）

故 $kx + ly \in V_1$　　（加法封闭）

例 12　在线性空间 V 中，设 $x_i \in V\ (i = 1, 2, \cdots, m)$，则

$$V_1 = \{x = k_1 x_1 + \cdots + k_m x_m \mid k_i \in K\}$$

是 V 的子空间，称 V_1 为由 $x_1, \cdots, x_m$ 生成的子空间.

证　$\forall\, x \in V_1 \Rightarrow x = k_1 x_1 + \cdots + k_m x_m$

$\forall\, y \in V_1 \Rightarrow y = l_1 x_1 + \cdots + l_m x_m$

$\forall\, k, l \in K$： $kx + ly = (kk_1 + ll_1)x_1 + \cdots + (kk_m + ll_m)x_m \in V_1$

根据例 11 知，V_1 是 V 的子空间.

[注] (1) 将 V_1 记作 $\mathrm{span}\{x_1, \cdots, x_m\}$ 或者 $L(x_1, \cdots, x_m)$.

(2) 元素组 $x_1, \cdots, x_m$ 的最大无关组是 $L(x_1, \cdots, x_m)$ 的基；

(3) 若线性空间 V^n 的基为 $x_1, \cdots, x_n$，则 $V^n = L(x_1, \cdots, x_n)$.

2．矩阵的值域（列空间）：

划分 $A=(a_{ij})_{m\times n}=(\beta_1,\cdots,\beta_n)\in \mathbf{C}^{m\times n}$（$\beta_j\in\mathbf{C}^m$），

称 $R(A)=L(\beta_1,\cdots,\beta_n)$ 为矩阵 A 的值域（列空间）.

易见 $\dim R(A)=\operatorname{rank}A$.

例 13　矩阵 A 的值域 $R(A)=\{\beta=Ax \mid x\in\mathbf{C}^n\}$.

证　$\forall\ \beta\in$左, 有 $\beta=k_1\beta_1+\cdots+k_n\beta_n=(\beta_1,\cdots,\beta_n)\begin{bmatrix}k_1\\ \vdots\\ k_n\end{bmatrix}=Ax\in$右

$\forall\ \beta\in$右, 有 $\beta=Ax=(\beta_1,\cdots,\beta_n)\begin{bmatrix}k_1\\ \vdots\\ k_n\end{bmatrix}=k_1\beta_1+\cdots+k_n\beta_n\in$左

3．矩阵的零空间：

设 $A\in\mathbf{C}^{m\times n}$，称 $N(A)=\{x \mid Ax=0, x\in\mathbf{C}^n\}$ 为矩阵 A 的零空间.

易见 $\dim N(A)=n-\operatorname{rank}A$.

Th3　线性空间 V^n 中, 设子空间 V_1 的基为 $x_1,\cdots,x_m\ (m<n)$, 则存在

$x_{m+1},\cdots,x_n\in V^n$, 使得 $x_1,\cdots,x_m,x_{m+1},\cdots,x_n$ 为 V^n 的基.

证　$m<n\Rightarrow\exists\, x_{m+1}\in V^n$不能由 $x_1,\cdots,x_m$线性表示

$\Rightarrow x_1,\cdots,x_m,x_{m+1}$线性无关

若 $m+1=n$，则 $x_1,\cdots,x_m,x_{m+1}$ 是 V^n 的基;

否则，$m+1<n\Rightarrow\exists\, x_{m+2}\in V^n$不能由 $x_1,\cdots,x_m,x_{m+1}$ 线性表示

$\Rightarrow x_1,\cdots,x_m,x_{m+1},x_{m+2}$线性无关

若 $m+2=n$，则 $x_1,\cdots,x_m,x_{m+1},x_{m+2}$ 是 V^n 的基;

否则，$m+2<n\Rightarrow\cdots\cdots$.

依此类推, 即得所证.

六、子空间的交与和

1． 子空间的交：$V_1 \cap V_2 = \{x \mid x \in V_1 \text{且} x \in V_2\}$

Th4 设 V_1, V_2 是线性空间 V 的子空间，则 $V_1 \cap V_2$ 是 V 的子空间.

证　$\theta \in V_1, \theta \in V_2 \Rightarrow \theta \in V_1 \cap V_2 \Rightarrow V_1 \cap V_2$ 非空

$$\forall x, y \in V_1 \cap V_2 \Rightarrow \begin{cases} x, y \in V_1 \Rightarrow x + y \in V_1 \\ x, y \in V_2 \Rightarrow x + y \in V_2 \end{cases} \Rightarrow x + y \in V_1 \cap V_2$$

$$\forall k \in K, \forall x \in V_1 \cap V_2 \Rightarrow \begin{cases} x \in V_1 \Rightarrow kx \in V_1 \\ x \in V_2 \Rightarrow kx \in V_2 \end{cases} \Rightarrow kx \in V_1 \cap V_2$$

所以 $V_1 \cap V_2$ 是 V 的子空间.

2． 子空间的和：　$V_1 + V_2 = \{x = x_1 + x_2 \mid x_1 \in V_1, x_2 \in V_2\}$

Th5 设 V_1, V_2 是线性空间 V 的子空间，则 $V_1 + V_2$ 是 V 的子空间.

证　$\theta \in V_1, \theta \in V_2 \Rightarrow \theta = \theta + \theta \in V_1 + V_2 \Rightarrow V_1 + V_2$ 非空

$$\forall x, y \in V_1 + V_2 \Rightarrow \begin{cases} x = x_1 + x_2, x_1 \in V_1, x_2 \in V_2 \\ y = y_1 + y_2, y_1 \in V_1, y_2 \in V_2 \end{cases}$$

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

$$\Rightarrow x + y \in V_1 + V_2$$

$$\forall k \in K, \forall x \in V_1 + V_2 \Rightarrow x = x_1 + x_2, x_1 \in V_1, x_2 \in V_2$$

$$\Rightarrow kx = kx_1 + kx_2,\ kx_1 \in V_1, kx_2 \in V_2$$

$$\Rightarrow kx \in V_1 + V_2$$

所以 $V_1 + V_2$ 是 V 的子空间.

[注]　$V_1 \cup V_2$ 不一定是 V 的子空间.

例如：在 $\mathbf{R}^2$ 中，$V_1 = L(e_1)$ 与 $V_2 = L(e_2)$ 的并集为

$$V_1 \cup V_2 = \{\alpha = (\xi_1, \xi_2) \mid \xi_1 \cdot \xi_2 = 0, \xi_i \in \mathbf{R}\}$$

易见 $e_1, e_2 \in V_1 \cup V_2$，但 $e_1 + e_2 = (1,1) \notin V_1 \cup V_2$，故加法运算不封闭.

Th6 设V_1,V_2是线性空间V的有限维子空间，则

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

证　记 $\dim V_1=n_1$， $\dim V_2=n_2$， $\dim V_1\cap V_2=m$

欲证 $\dim(V_1+V_2)=n_1+n_2-m$

(1) $m=n_1$： $(V_1\cap V_2)\subset V_1\Rightarrow V_1\cap V_2=V_1$

$$(V_1\cap V_2)\subset V_2\Rightarrow V_1\subset V_2\Rightarrow V_1+V_2=V_2$$

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

(2) $m=n_2$： $(V_1\cap V_2)\subset V_2\Rightarrow V_1\cap V_2=V_2$

$$(V_1\cap V_2)\subset V_1\Rightarrow V_2\subset V_1\Rightarrow V_1+V_2=V_1$$

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

(3) $m<n_1,m<n_2$： 设$V_1\cap V_2$的基为$x_1,\cdots,x_m$，那么

扩充为V_1的基： $x_1,\cdots,x_m,y_1,\cdots,y_{n_1-m}$　　（Ⅰ）

扩充为V_2的基： $x_1,\cdots,x_m,z_1,\cdots,z_{n_2-m}$　　（Ⅱ）

考虑元素组： $x_1,\cdots,x_m,\ y_1,\cdots,y_{n_1-m},z_1,\cdots,z_{n_2-m}$　　(Ⅲ)

因为 $V_1=L$（Ⅰ）， $V_2=L$（Ⅱ），所以 $V_1+V_2=L$(Ⅲ) （自证）.

下面证明元素组(Ⅲ)线性无关：

设数组$k_1,\cdots,k_m,\ p_1,\cdots,p_{n_1-m},q_1,\cdots,q_{n_2-m}$使得

$$k_1x_1+\cdots+k_mx_m+p_1y_1+\cdots+p_{n_1-m}y_{n_1-m}+q_1z_1+\cdots+q_{n_2-m}z_{n_2-m}=\theta$$

由 $$x=\begin{cases}k_1x_1+\cdots+k_mx_m+p_1y_1+\cdots+p_{n_1-m}y_{n_1-m}\in V_1\\-(q_1z_1+\cdots+q_{n_2-m}z_{n_2-m})\in V_2\end{cases}\qquad(*)$$

得 $x\in V_1\cap V_2\Rightarrow x=l_1x_1+\cdots+l_mx_m$

结合(*)中第二式得

$$l_1x_1+\cdots+l_mx_m+q_1z_1+\cdots+q_{n_2-m}z_{n_2-m}=\theta$$

（Ⅱ）线性无关$\Rightarrow l_1=\cdots=l_m=0, q_1=\cdots=q_{n_2-m}=0$

结合(*)中第一式得

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

（Ⅰ）线性无关$\Rightarrow k_1=\cdots=k_m=0,\ p_1=\cdots=p_{n_1-m}=0$

故元素组(Ⅲ)线性无关，从而是V_1+V_2的一个基.

因此 $\dim(V_1+V_2)=n_1+n_2-m$.

3．子空间的直和：

$$V_1+V_2=\{x=x_1+x_2 \mid \text{唯一}\, x_1\in V_1, \text{唯一}\, x_2\in V_2\}$$

记作：$V_1+V_2=V_1\oplus V_2$

Th7 设V_1,V_2是线性空间V的子空间，则V_1+V_2是直和$\Leftrightarrow V_1\cap V_2=\{\theta\}$.

证　充分性．已知$V_1\cap V_2=\{\theta\}$：对于$\forall\, z\in V_1+V_2$，若

$$\begin{cases} z=x_1+x_2,\ x_1\in V_1,\ x_2\in V_2 \\ z=y_1+y_2,\ y_1\in V_1,\ y_2\in V_2 \end{cases}$$

则有 $(x_1-y_1)+(x_2-y_2)=\theta,\ x_1-y_1\in V_1, x_2-y_2\in V_2$

$$\begin{aligned} &\Rightarrow x_1-y_1=-(x_2-y_2)\in V_1\cap V_2 \\ &\Rightarrow x_1-y_1=\theta,\ x_2-y_2=\theta \\ &\Rightarrow x_1=y_1,\ x_2=y_2 \end{aligned}$$

故$z\in V_1+V_2$的分解式唯一，从而$V_1+V_2=V_1\oplus V_2$.

必要性．若$V_1\cap V_2\ne\{\theta\}$，则有$\theta\ne x\in V_1\cap V_2$．对于$\theta\in V_1+V_2$，有

$$\begin{aligned} &\theta=\theta+\theta, \quad \theta\in V_1,\theta\in V_2 \\ &\theta=x+(-x), \quad x\in V_1,(-x)\in V_2 \end{aligned}$$

即$\theta\in V_1+V_2$有两种不同的分解式．这与V_1+V_2是直和矛盾.

故$V_1\cap V_2=\{\theta\}$.

推论 1　V_1+V_2是直和$\Leftrightarrow \dim(V_1+V_2)=\dim V_1+\dim V_2$

推论 2　设V_1+V_2是直和，V_1的基为$x_1,\cdots,x_k$，V_2的基为$y_1,\cdots,y_l$，

则V_1+V_2的基为$x_1,\cdots,x_k,y_1,\cdots,y_l$.

证　因为 $V_1+V_2=L(x_1,\cdots,x_k,y_1,\cdots,y_l)$，且

$$\dim(V_1+V_2)=\dim V_1+\dim V_2=k+l$$

所以$x_1,\cdots,x_k,y_1,\cdots,y_l$线性无关, 故$x_1,\cdots,x_k,y_1,\cdots,y_l$是$V_1+V_2$的基.

§1.2　线性变换及其矩阵

一、线性变换

1．定义　线性空间 V，数域 K，T 是 V 中的变换．若对 $\forall\, x,y\in V$，$\forall\, k,l\in K$，

都有　$T(kx+ly)=k(Tx)+l(Ty)$，　称 T 是 V 中的线性变换．

性质 (1)　$T\theta=T(0x+0y)=0(Tx)+0(Ty)=\theta$

(2)　$T(-x)=T((-1)x+0y)=(-1)(Tx)+0(Ty)=-(Tx)$

(3)　$x_1,\cdots,x_m\in V$ 线性相关 $\Rightarrow$ $Tx_1,\cdots,Tx_m$ 线性相关

(4)　$x_1,\cdots,x_m\in V$ 线性无关时，不能推出 $Tx_1,\cdots,Tx_m$ 线性无关．

(5)　T 是线性变换 $\Leftrightarrow$ $T(x+y)=Tx+Ty$，$T(kx)=k(Tx)$

$$(\forall\, x,y\in V\,,\forall\, k\in K\,)$$

例 1　矩阵空间 $\mathbf{R}^{n\times n}$，给定矩阵 $B_{n\times n}$，则变换 $TX=BX+XB$　$(\forall X\in\mathbf{R}^{n\times n})$

是 $\mathbf{R}^{n\times n}$ 的线性变换．

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

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

Th8　设 T 是线性空间 V 的线性变换，则 $R(T)$和 $N(T)$都是 V 的子空间．

证　(1) V 非空 $\Rightarrow$ $R(T)$ 非空．

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

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

$y_1+y_2=Tx_1+Tx_2=T(x_1+x_2)\in R(T)$　　　$(\because\ x_1+x_2\in V)$

$k\,y_1=k(Tx_1)=T(k\,x_1)\in R(T)$　　　$(\because\ \forall k\in K,\,kx_1\in V)$

故 $R(T)$是 V 的子空间．

(2)　$\theta\in V,\,T\theta=\theta\Rightarrow\theta\in N(T)$，即 $N(T)$ 非空．

$\forall x, y \in N(T) \Rightarrow T(x+y) = Tx + Ty = \theta$，即 $x + y \in N(T)$.

$\forall x \in N(T), \forall k \in K \Rightarrow T(kx) = k(Tx) = \theta$，即 $kx \in N(T)$.

故 $N(T)$是 V 的子空间.

[注]　定义：T 的秩 $= \dim R(T)$，T 的亏 $= \dim N(T)$

例 2 设线性空间 V^n 的基为 $x_1,\cdots,x_n$，　T 是 V^n 的线性变换，则

$$R(T) = L(Tx_1,\cdots,Tx_n)，\quad \dim R(T) + \dim N(T) = n$$

证　(1) 先证 $R(T) \subset L(Tx_1,\cdots,Tx_n)$：$\forall\, y \in R(T) \Rightarrow \exists\, x \in V^n, \text{st } y = Tx$

$x = c_1x_1 + \cdots + c_nx_n \Rightarrow y = c_1(Tx_1) + \cdots + c_n(Tx_n) \in L(Tx_1,\cdots,Tx_n)$

再证 $R(T) \supset L(Tx_1,\cdots,Tx_n)$：

$\forall\, y \in L(Tx_1,\cdots,Tx_n) \Rightarrow \exists\; c_1,\cdots,c_n$, st $y = c_1(Tx_1) + \cdots + c_n(Tx_n)$

$x_i \in V^n \Rightarrow Tx_i \in R(T) \Rightarrow y = c_1(Tx_1) + \cdots + c_n(Tx_n) \in R(T)$

(2) 设 $\dim N(T) = m$，且 $N(T)$ 的基为 $y_1,\cdots,y_m$，扩充为 V^n 的基：

$$y_1,\cdots,y_m,y_{m+1},\cdots,y_n$$

则　$R(T) = L(Ty_1,\cdots,Ty_m,Ty_{m+1},\cdots,Ty_n) = L(Ty_{m+1},\cdots,Ty_n)$

设数组 $k_{m+1},\cdots,k_n$ 使得 $k_{m+1}(Ty_{m+1}) + \cdots + k_n(Ty_n) = \theta$，则

$$T(k_{m+1}y_{m+1} + \cdots + k_ny_n) = \theta$$

因为 T 是线性变换，所以 $k_{m+1}y_{m+1} + \cdots + k_ny_n \in N(T)$，故

$$k_{m+1}y_{m+1} + \cdots + k_ny_n = l_1y_1 + \cdots + l_my_m$$

即　$$(-l_1)y_1 + \cdots + (-l_m)y_m + k_{m+1}y_{m+1} + \cdots + k_ny_n = \theta$$

因为 $y_1,\cdots,y_m,y_{m+1},\cdots,y_n$ 线性无关，所以 $k_{m+1} = 0,\cdots,k_n = 0$. 因此 $Ty_{m+1},\cdots,Ty_n$ 线性无关，从而 $\dim R(T) = n - m$，即 $\dim R(T) + m = n$.

例 3 向量空间 $\mathbf{R}^4$ 中，$x = (\xi_1,\xi_2,\xi_3,\xi_4)$，线性变换 T 为

$$Tx=(\xi_1+\xi_2-3\xi_3-\xi_4, 3\xi_1-\xi_2-3\xi_3+4\xi_4,\ 0,\ 0)$$

求 $R(T)$ 和 $N(T)$ 的基与维数.

解 (1) 取 $\mathbf{R}^4$ 的简单基 e_1,e_2,e_3,e_4， 计算

$Te_1=(1,3,0,0)$，$Te_2=(1,-1,0,0)$，$Te_3=(-3,-3,0,0)$，$Te_4=(-1,4,0,0)$

该基象组的一个最大线性无关组为 Te_1,Te_2.

故 $\dim R(T)=2$，且 $R(T)$的一个基为 Te_1,Te_2.

(2) 记 $A=\begin{bmatrix}1 & 1 & -3 & -1\\ 3 & -1 & -3 & 4\end{bmatrix}$，则 $N(T)=\{x \mid Tx=\theta\}=\{x \mid A\begin{bmatrix}\xi_1\\ \vdots\\ \xi_4\end{bmatrix}=0\}$

$A\begin{bmatrix}\xi_1\\ \vdots\\ \xi_4\end{bmatrix}=0$ 的基础解系为 $\begin{bmatrix}3\\3\\2\\0\end{bmatrix}$，$\begin{bmatrix}-3\\7\\0\\4\end{bmatrix}$.

故 $\dim N(T)=2$，且 $N(T)$的一个基为(3, 3, 2, 0)，(－3, 7, 0, 4).

4．单位变换：线性空间 V 中，定义变换 T 为 $Tx=x \quad (\forall x\in V)$，

则 T 是线性变换，记作 T_e .

5．零变换：线性空间 V 中，定义变换 T 为 $Tx=\theta \quad (\forall x\in V)$，

则 T 是线性变换，记作 T_0 .

6．线性变换的运算：线性空间 V，数域 K，线性变换 T_1 与 T_2 .

(1) 相等：若 $T_1x=T_2x\ (\forall x\in V)$，称 $T_1=T_2$.

(2) 加法：定义变换 T 为 $Tx=T_1x+T_2x \quad (\forall x\in V)$，

则 T 是线性变换，记作 $T=T_1+T_2$.

负变换：定义变换 T 为 $Tx=-(T_1x) \quad (\forall x\in V)$，

则 T 是线性变换， 记作 $T=-T_1$.

(3) 数乘：给定$k \in K$，定义变换 T 为 $Tx = k(T_1 x) \quad (\forall x \in V)$，

则 T 是线性变换， 记作$T = kT_1$.

[注] 集合 $\mathrm{Hom}(V, V) \overset{\mathrm{def}}{=} \{T \mid T$是数域 K上的线性空间 V 的线性变换$\}$

按照线性运算(2)和(3)构成数域 K 上的线性空间，称为 V 的同态.

(4) 乘法：定义变换 T 为 $Tx = T_1(T_2 x) \quad (\forall x \in V)$，

则 T 是线性变换， 记作$T = T_1 T_2$.

7．逆变换：设 T 是线性空间 V 的线性变换，若 V 的线性变换 S 满足

$$(ST)x = (TS)x = x \quad (\forall x \in V)$$

则称 T 为可逆变换，且 S 为 T 的逆变换，记作 $T^{-1} = S$.

8．幂变换：设 T 是线性空间 V 的线性变换，

则$T^m \overset{\mathrm{def}}{=} T^{m-1}T \quad (m = 2, 3, \cdots)$也是 V 的线性变换.

9．多项式变换：设 T 是线性空间 V 的线性变换，多项式

$$f(t) = a_0 + a_1 t + \cdots + a_m t^m \quad (a_i \in K)$$

则 $f(T) = a_0 T_e + a_1 T + \cdots + a_m T^m$ 也是 V 的线性变换.

二、线性变换的矩阵表示

1．线性变换在给定基下的矩阵

设线性空间V^n的基为 $x_1, \cdots, x_n$，T 是V^n的线性变换，则$Tx_i \in V^n$，且有

$$\begin{cases} Tx_1 = a_{11}x_1 + a_{21}x_2 + \cdots + a_{n1}x_n \\ Tx_2 = a_{12}x_1 + a_{22}x_2 + \cdots + a_{n2}x_n \\ \qquad \cdots\cdots \\ Tx_n = a_{1n}x_1 + a_{2n}x_2 + \cdots + a_{nn}x_n \end{cases} \qquad A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

写成矩阵乘法形式 $T(x_1, \cdots, x_n) \overset{\mathrm{def}}{=} (Tx_1, \cdots, Tx_n) = (x_1, \cdots, x_n)A$

称 A 为线性变换 T 在基 $x_1,\cdots,x_n$ 下的矩阵.

[注] (1) 给定 V^n 的基 $x_1,\cdots,x_n$ 和线性变换 T 时，矩阵 A 唯一.

(2) 给定 V^n 的基 $x_1,\cdots,x_n$ 和矩阵 A 时，基象组 $Tx_1,\cdots,Tx_n$ 确定.

$\forall x\in V^n \Rightarrow x=c_1x_1+\cdots+c_nx_n$，定义变换

$$Tx=c_1(Tx_1)+\cdots+c_n(Tx_n)$$

则 T 是线性变换. 因此线性变换 T 与方阵 A 是一一对应关系.

例 4 线性空间 $P_n[t]$ 的线性变换为 $T\,f(t)=f'(t)\quad(\forall f(t)\in P_n[t])$.

基(I): $f_0=1, f_1=t, f_2=\dfrac{t^2}{2!},\cdots,f_n=\dfrac{t^n}{n!}$

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

记 T 在基(I)下的矩阵为 A_1，T 在基(II)下的矩阵为 A_2. 因为

$$Tf_0=0, Tf_1=f_0, Tf_2=f_1,\cdots,Tf_n=f_{n-1}$$

$$Tg_0=0, Tg_1=g_0, Tg_2=2g_1,\cdots,Tg_n=ng_{n-1}$$

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

易见 $A_1\neq A_2\ (n\geq 2)$.

例 5 线性空间 V^n 中，设线性变换 T 在基 $x_1,\cdots,x_n$ 下的矩阵为 A，则

$\dim R(T)=\operatorname{rank}A$，$\dim N(T)=n-\operatorname{rank}A$.

证 $\operatorname{rank}A=m\ \Leftrightarrow\ A$ 的列向量组 $\beta_1,\cdots,\beta_n$ 中最大无关组含 m 个向量

$\Leftrightarrow$ 元素组 $Tx_1,\cdots,Tx_n$ 中最大无关组含 m 个向量

$\Leftrightarrow \dim R(T)=\dim L(Tx_1,\cdots,Tx_n)=m$

由例 **2** 知另一结论成立.

2. 线性运算的矩阵表示（将线性变换运算转化为矩阵运算）

Th9 设线性空间 V^n 的基为 $x_1,\cdots,x_n$，线性变换 T_1 与 T_2 的矩阵 A 与 B，则

(1) T_1+T_2 在该基下的矩阵为 $A+B$.

(2) kT_1 在该基下的矩阵为 kA.

(3) T_1T_2 在该基下的矩阵为 AB.

(4) T_1^{-1} 在该基下的矩阵为 A^{-1}.

证　$T_1(x_1,\cdots,x_n)=(x_1,\cdots,x_n)A,\ T_2(x_1,\cdots,x_n)=(x_1,\cdots,x_n)B$

(1) 略. **(2)** 略.

(3) 先证：$\forall C=(c_{ij})_{n\times m},\ T[(x_1,\cdots,x_n)C]=[T(x_1,\cdots,x_n)]C$

$$\text{左}=T\left[\sum c_{i1}x_i,\cdots,\sum c_{im}x_i\right]=\left[\sum c_{i1}(Tx_i),\cdots,\sum c_{im}(Tx_i)\right]$$

$$=(Tx_1,\cdots,Tx_n)C=\text{右}$$

由此可得 $(T_1T_2)(x_1,\cdots,x_n)=T_1[T_2(x_1,\cdots,x_n)]=T_1[(x_1,\cdots,x_n)B]$

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

(4) 记 $T_1^{-1}=T_2$，则 $T_1T_2=T_2T_1=T_e\overset{(3)}{\Rightarrow}AB=BA=I\Rightarrow B=A^{-1}$.

3. 象与原象坐标间的关系

Th10 线性空间 V^n 的基为 $x_1,\cdots,x_n$，线性变换 T 在该基下的矩阵为 A，

$x\in V^n$ 的坐标为 $\begin{pmatrix}\xi_1\\ \vdots\\ \xi_n\end{pmatrix}$，$Tx$ 的坐标为 $\begin{pmatrix}\eta_1\\ \vdots\\ \eta_n\end{pmatrix}$，则 $\begin{pmatrix}\eta_1\\ \vdots\\ \eta_n\end{pmatrix}=A\begin{pmatrix}\xi_1\\ \vdots\\ \xi_n\end{pmatrix}$.

证　$x=\xi_1x_1+\cdots+\xi_nx_n$

$$Tx=\xi_1(Tx_1)+\cdots+\xi_n(Tx_n)=(Tx_1,\cdots,Tx_n)\begin{pmatrix}\xi_1\\ \vdots\\ \xi_n\end{pmatrix}=(x_1,\cdots,x_n)A\begin{pmatrix}\xi_1\\ \vdots\\ \xi_n\end{pmatrix}$$

由定理 2 知 $\begin{pmatrix}\eta_1\\ \vdots\\ \eta_n\end{pmatrix}=A\begin{pmatrix}\xi_1\\ \vdots\\ \xi_n\end{pmatrix}$.

4．线性变换在不同基下矩阵之间的关系

Th11 线性空间 V^n 的基(I)：$x_1,\cdots,x_n$，基(II)：$y_1,\cdots,y_n$

线性变换 T：$T(x_1,\cdots,x_n)=(x_1,\cdots,x_n)A$

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

由基(I)到基(II)的过渡矩阵为 C，则 $B=C^{-1}AC$.

证 因为 $T(y_1\cdots,y_n)=T(x_1,\cdots x_n)C=(x_1,\cdots,x_n)AC=(y_1,\cdots,y_n)C^{-1}AC$

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

所以 $B=C^{-1}AC$.

三、线性变换的特征值与特征向量

1．定义 线性空间 V，线性变换 T，若 $\lambda_0\in K$ 及 $\theta\neq x\in V$ 满足 $Tx=\lambda_0 x$，

称 λ_0 为 T 的特征值，x 为 T 的对应于 λ_0 的特征向量（元素）.

2．算法 设线性空间 V^n 的基为 $x_1,\cdots,x_n$，线性变换 T 的矩阵为 $A_{n\times n}$.

T 的特征值为 λ_0，对应的特征向量为 x.

x 的坐标为 $\alpha=\begin{pmatrix}\xi_1\\ \vdots\\ \xi_n\end{pmatrix}$，$Tx$ 的坐标为 $A\alpha$，$\lambda_0 x$ 的坐标为 $\lambda_0\alpha$.

因为 $Tx=\lambda_0 x\Leftrightarrow A\alpha=\lambda_0\alpha$，所以 T 的特征值与 A 的特征值相同；

T 的对应于 λ_0 的特征向量的坐标就是 A 的对应于 λ_0 的特征向量.

例 6 设 $B=\begin{bmatrix}1&1\\0&1\end{bmatrix}$，线性空间 $V=\{X=(x_{ij})_{2\times2}\mid x_{11}+x_{22}=0,x_{ij}\in\mathrm{R}\}$，

线性变换为 $TX=B^{\mathrm T}X-X^{\mathrm T}B\ (X\in V)$，求 T 的特征值与特征向量.

解　$X \in V \Rightarrow X = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & -x_{11} \end{bmatrix} = \begin{bmatrix} x_{11} & 0 \\ 0 & -x_{11} \end{bmatrix} + \begin{bmatrix} 0 & x_{12} \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ x_{21} & 0 \end{bmatrix}$

$$= x_{11}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} + x_{12}\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + x_{21}\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$

可得 V 的简单基为　$X_1 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},\ X_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix},\ X_3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$

由公式求得　$TX_1 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix},\ TX_2 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix},\ TX_3 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$

故 T 在简单基下的矩阵为

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

A 的特征值与线性无关的特征向量为

$$\lambda_1 = \lambda_2 = 0,\ \alpha_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \alpha_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix};\quad \lambda_3 = 2, \alpha_3 = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$$

T 的特征值与线性无关的特征向量为

$$\lambda_1 = \lambda_2 = 0, Y_1 = (X_1, X_2, X_3)\alpha_1 = \begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix}$$

$$Y_2 = (X_1, X_2, X_3)\alpha_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$$\lambda_3 = 2, Y_3 = (X_1, X_2, X_3)\alpha_3 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

例 7　线性空间 V，线性变换 T，$V_{\lambda_0} = \{x \mid Tx = \lambda_0 x, x \in V\}$ 是 V 的子空间.

证　$\theta \in V, T\theta = \lambda_0\theta \Rightarrow \theta \in V_{\lambda_0}$，即 V_{λ_0} 非空.

$$\forall x, y \in V_{\lambda_0} \Rightarrow T(x+y) = Tx + Ty = \lambda_0 x + \lambda_0 y = \lambda_0(x+y)$$

$$\Rightarrow x + y \in V_{\lambda_0}$$

$$\forall k \in K, \forall x \in V_{\lambda_0} \Rightarrow T(kx) = k(Tx) = k(\lambda_0 x) = \lambda_0(kx) \Rightarrow kx \in V_{\lambda_0}$$

故V_{λ_0}是 V 的子空间.

[注] 若λ_0是线性变换的特征值，则称V_{λ_0}为 T 的特征子空间.

3．矩阵的迹：$A=(a_{ij})_{n\times n}$, $\mathrm{tr}A\overset{\Delta}{=}\sum_{i=1}^{n}a_{ii}$.

Th12 $A_{m\times n}, B_{n\times m}\Rightarrow \mathrm{tr}(AB)=\mathrm{tr}(BA)$.

证 $A=(a_{ij})_{m\times n}$, $B=(b_{ij})_{n\times m}$， $AB\overset{\Delta}{=}(u_{ij})_{m\times m}$， $BA\overset{\Delta}{=}(v_{ij})_{n\times n}$：

$$u_{ii}=(a_{i1},\cdots,a_{in})\begin{pmatrix}b_{1i}\\ \vdots\\ b_{ni}\end{pmatrix}=\sum_{k=1}^{n}a_{ik}b_{ki}\ ,\quad v_{kk}=(b_{k1},\cdots,b_{km})\begin{pmatrix}a_{1k}\\ \vdots\\ a_{mk}\end{pmatrix}=\sum_{i=1}^{m}b_{ki}a_{ik}$$

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

Th13 若 A 相似于 B ，则$\mathrm{tr}A=\mathrm{tr}B$.

证 由$B=P^{-1}AP$可得 $\mathrm{tr}B=\mathrm{tr}(P^{-1}AP)=\mathrm{tr}((AP)P^{-1})=\mathrm{tr}A$

[注] 因为相似矩阵有相同的特征值（Th14 -- 线性代数课程结论）

所以线性变换的特征值与线性空间中基的选取无关

4．三角相似

Th17 $A_{n\times n}$相似于上三角矩阵.

证 归纳法．$n=1$ 时，$A=(a_{11})$是上三角矩阵$\Rightarrow A$ 相似于上三角矩阵.

假设 $n=k-1$ 时定理成立，下证 $n=k$ 时定理也成立.

$A_{k\times k}$的特征值为$\lambda_1,\lambda_2,\cdots,\lambda_k$，对应$\lambda_1$的特征向量为$x_1\Rightarrow Ax_1=\lambda_1x_1$.

扩充x_1为C^k的基：$x_1,x_2,\cdots,x_k$（列向量）

$P_1=(x_1,x_2,\cdots,x_k)$可逆，$AP_1=(Ax_1,Ax_2,\cdots,Ax_k)$

$Ax_j\in\mathrm{C}^k\Rightarrow Ax_j=b_{1j}x_1+b_{2j}x_2+\cdots+b_{kj}x_k\ \ (j=2,\cdots,k)$

$$AP_1=(x_1,x_2,\cdots,x_k)\begin{bmatrix}\lambda_1 & b_{12} & \cdots & b_{1k}\\ 0 & b_{22} & \cdots & b_{2k}\\ \vdots & \vdots & & \vdots\\ 0 & b_{k2} & \cdots & b_{kk}\end{bmatrix}$$

$$P_1^{-1}AP_1=\left[\begin{array}{c|ccc}\lambda_1 & b_{12} & \cdots & b_{1k}\\ \hline 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{array}\right]$$

A_1 的特征值为 $\lambda_2,\cdots,\lambda_k$，由假设知，存在 $k-1$ 阶可逆矩阵 Q 使得

$$Q^{-1}A_1Q=\begin{bmatrix}\lambda_2 & \cdots & *\\ & \ddots & \vdots\\ & & \lambda_k\end{bmatrix},\quad P_2\overset{\Delta}{=}\left[\begin{array}{c|ccc}1 & 0 & \cdots & 0\\ \hline 0 & & & \\ \vdots & & Q & \\ 0 & & & \end{array}\right]$$

$$P\overset{\Delta}{=}P_1P_2\Rightarrow P^{-1}AP=\cdots=\begin{bmatrix}\lambda_1 & * & \cdots & *\\ & \lambda_2 & \ddots & \vdots\\ & & \ddots & *\\ & & & \lambda_k\end{bmatrix}$$

由归纳法原理，对任意 n，定理成立.

5．Hamilton-Cayley 定理

Th18　设 $A_{n\times n},\varphi(\lambda)\overset{\Delta}{=}\det(\lambda I-A)=\lambda^n+a_1\lambda^{n-1}+\cdots+a_{n-1}\lambda+a_n$，则

$$\varphi(A)\overset{\Delta}{=}A^n+a_1A^{n-1}+\cdots+a_{n-1}A+a_nI=O_{n\times n}$$

证　A 的特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_n\Rightarrow\varphi(\lambda)=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n)$.

由 Th17 知，存在可逆矩阵 $P_{n\times n}$，使得 $P^{-1}AP=\begin{bmatrix}\lambda_1 & \cdots & *\\ & \ddots & \vdots\\ & & \lambda_n\end{bmatrix}$.

$$\varphi(P^{-1}AP)=(P^{-1}AP-\lambda_1I)(P^{-1}AP-\lambda_2I)\cdots(P^{-1}AP-\lambda_nI)$$

$$
=\begin{bmatrix} 0 & * & \cdots & * \\ & \lambda_2-\lambda_1 & \ddots & \vdots \\ & & \ddots & * \\ & & & \lambda_n-\lambda_1 \end{bmatrix}\begin{bmatrix} \lambda_1-\lambda_2 & * & \cdots & * \\ & 0 & \ddots & \vdots \\ & & \ddots & * \\ & & & \lambda_n-\lambda_2 \end{bmatrix}\cdots
$$

$$
\begin{bmatrix} \lambda_1-\lambda_n & * & \cdots & * \\ & \ddots & \ddots & \vdots \\ & & \lambda_{n-1}-\lambda_n & * \\ & & & 0 \end{bmatrix}
$$

$$
=\begin{bmatrix} 0 & 0 & * & \cdots & * \\ 0 & 0 & * & \cdots & * \\ 0 & 0 & * & \ddots & \vdots \\ \vdots & \vdots & & \ddots & * \\ 0 & 0 & & & * \end{bmatrix}\begin{bmatrix} \lambda_1-\lambda_3 & * & * & \cdots & * \\ & \lambda_2-\lambda_3 & * & \cdots & * \\ & & 0 & \ddots & \vdots \\ & & & \ddots & * \\ & & & & \lambda_n-\lambda_3 \end{bmatrix}\cdots
$$

$$
\begin{bmatrix} \lambda_1-\lambda_n & * & \cdots & * \\ & \ddots & \ddots & \vdots \\ & & \lambda_{n-1}-\lambda_n & * \\ & & & 0 \end{bmatrix}=O
$$

即 $P^{-1}\varphi(A)P=O \Rightarrow \varphi(A)=O$.

[注] (1) $|A|\neq 0 \Rightarrow a_n\neq 0,\ A^{-1}=-\dfrac{1}{a_n}\left(A^{n-1}+a_1A^{n-2}+\cdots+a_{n-2}A+a_{n-1}I\right)$

(2) $A^n\in \operatorname{span}\{A^{n-1},\cdots,A,I\}$

例 8 $A=\begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & 1 \\ 0 & -1 & 2 \end{bmatrix}$，计算 $A^{100}+2A^{50}$.

解 $f(\lambda)=\lambda^{100}+2\lambda^{50}$，$\varphi(\lambda)=\det(\lambda I-A)=(\lambda-1)^2(\lambda-2)$

φ 除 f ： $f(\lambda)=\varphi(\lambda)g(\lambda)+\left(b_0+b_1\lambda+b_2\lambda^2\right)$

$f'(\lambda)=[\varphi(\lambda)g(\lambda)]'+\left(b_1+2b_2\lambda\right)$

由 $f(1)=3,\ f'(1)=200,\ f(2)=2^{100}+2^{51}$ 可得

$$\left.\begin{array}{l} b_0+b_1+b_2=3 \\ b_1+2b_2=200 \\ b_0+2b_1+4b_2=2^{100}+2^{51} \end{array}\right\} \Rightarrow \left\{\begin{array}{l} b_0=2^{100}+2^{51}-400 \\ b_1=-2^{101}-2^{52}+606 \\ b_2=2^{100}+2^{51}-203 \end{array}\right.$$

$$\varphi(A)=O \Rightarrow f(A)=b_0I+b_1A+b_2A^2$$

6．最小多项式：以 $A_{n\times n}$ 为根，且次数最低的首 1 多项式，记作 $m(\lambda)$.

$$f(\lambda)=1 \Rightarrow f(A)=I\neq O \Rightarrow \partial m(\lambda)\geq 1$$

$$\varphi(\lambda)=\det(\lambda I-A),\ \text{Th18} \Rightarrow \varphi(A)=O \Rightarrow \partial m(\lambda)\leq n$$

例 9 $A=\begin{bmatrix} 3 & -3 & 2 \\ -1 & 5 & -2 \\ -1 & 3 & 0 \end{bmatrix}$, $\varphi(\lambda)=(\lambda-2)^2(\lambda-4)$

$$f_1(\lambda)=\lambda+k\ (\forall k\in \mathbf{R}):\ f_1(A)=A+kI\neq O \Rightarrow \partial m(\lambda)>1$$

$$f_2(\lambda)=(\lambda-2)(\lambda-4):\ f_2(A)=(A-2I)(A-4I)=O \Rightarrow m(\lambda)=f_2(\lambda)$$

Th19 (1) 多项式 $f(\lambda)$ 满足 $f(A)=O \Rightarrow m(\lambda)\mid f(\lambda)$;

(2) $m(\lambda)$ 唯一.

证 (1) 反证法.

$$m(\lambda)\nmid f(\lambda) \Rightarrow f(\lambda)=m(\lambda)g(\lambda)+r(\lambda)$$

$$r(\lambda)\neq 0 \text{ 且 } \partial r(\lambda)<\partial m(\lambda)$$

$$\Rightarrow f(A)=m(A)g(A)+r(A)$$

$$f(A)=O,\ m(A)=O \Rightarrow r(A)=O,\ \partial r(\lambda)<\partial m(\lambda)$$

$\Rightarrow m(\lambda)$ 不是 A 的最小多项式，矛盾！

(2) 设 $m(\lambda)$ 与 $\tilde{m}(\lambda)$ 都是 A 的最小多项式，则

$$\left.\begin{array}{l} \tilde{m}(A)=O \Rightarrow m(\lambda)\mid \tilde{m}(\lambda) \\ m(A)=O \Rightarrow \tilde{m}(\lambda)\mid m(\lambda) \end{array}\right\} \overset{\text{首1}}{\Rightarrow} m(\lambda)=\tilde{m}(\lambda)$$

Th20 $m(\lambda)$ 与 $\varphi(\lambda)$ 的零点相同（不计重数）.

证　Th19 $\Rightarrow m(\lambda)$ 的零点是 $\varphi(\lambda)$ 的零点．再设 λ_0 是 $\varphi(\lambda)$ 的零点，则有

$$Ax = \lambda_0 x\ (x \neq 0) \Rightarrow m(A)x = m(\lambda_0)x$$

$m(A) = O \Rightarrow m(\lambda_0)x = 0 \Rightarrow m(\lambda_0) = 0$，故 λ_0 也是 $m(\lambda)$ 的零点．

[注]　Th20 $\Rightarrow m(\lambda)$ 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．

例如：$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \varphi(\lambda) = (\lambda - 1)^2, m(\lambda) \neq (\lambda - 1)$．

7．最小多项式求法

Th21　对 $A_{n\times n}$，设 $\lambda I - A$ 的第 i 行第 j 列元素的余子式为 $M_{ij}(\lambda)$，则

$$m(\lambda) = \frac{\det(\lambda I - A)}{d(\lambda)} \qquad \left(d(\lambda) = \max_{i,j} \lfloor M_{ij}(\lambda) \rfloor \right)$$

例 10　设 $A = \begin{bmatrix} 3 & -3 & 2 \\ -1 & 5 & -2 \\ -1 & 3 & 0 \end{bmatrix}$，求 $m(\lambda)$．

解　$\lambda I - A = \begin{bmatrix} \lambda - 3 & 3 & -2 \\ 1 & \lambda - 5 & 2 \\ 1 & -3 & \lambda \end{bmatrix}$，$\varphi(\lambda) = \det(\lambda I - A) = (\lambda - 2)^2(\lambda - 4)$

$M_{11} = \lambda^2 - 5\lambda + 6$，$M_{21} = 3(\lambda - 2)$，　$M_{31} = 2(\lambda - 2)$

$M_{12} = \lambda - 2$，　$M_{22} = \lambda^2 - 3\lambda + 2$，$M_{32} = 2(\lambda - 2)$

$M_{13} = -(\lambda - 2)$，　$M_{23} = -3(\lambda - 2)$，　$M_{33} = \lambda^2 - 8\lambda + 12$

$$d(\lambda) = \lambda - 2,\ m(\lambda) = \frac{\varphi(\lambda)}{d(\lambda)} = (\lambda - 2)(\lambda - 4) = \lambda^2 - 6\lambda + 8$$

例 11　$A_{n\times n}$ 相似于 $B_{n\times n} \Rightarrow m_A(\lambda) = m_B(\lambda)$．

证　$B = P^{-1}AP \Rightarrow A = PBP^{-1}$

取 $f(\lambda) = m_A(\lambda)$，则 $f(A) = m_A(A) = O$，从而有

$$f(B) = f(P^{-1}AP) = P^{-1}f(A)P = O$$

$$\overset{\text{Th19}}{\Rightarrow} \boldsymbol{m}_{B}(\lambda) \mid \boldsymbol{f}(\lambda), \text{即 } \boldsymbol{m}_{B}(\lambda) \mid \boldsymbol{m}_{A}(\lambda) \qquad ①$$

取 $\boldsymbol{g}(\lambda)=\boldsymbol{m}_B(\lambda)$，则 $\boldsymbol{g}(\boldsymbol{B})=\boldsymbol{m}_B(\boldsymbol{B})=\boldsymbol{O}$，从而有

$$\boldsymbol{g}(\boldsymbol{A})=\boldsymbol{g}\left(\boldsymbol{P}\boldsymbol{B}\boldsymbol{P}^{-1}\right)=\boldsymbol{P}\,\boldsymbol{g}(\boldsymbol{B})\boldsymbol{P}^{-1}=\boldsymbol{O}$$

$$\overset{\text{Th19}}{\Rightarrow} \boldsymbol{m}_{A}(\lambda) \mid \boldsymbol{g}(\lambda), \text{即 } \boldsymbol{m}_{A}(\lambda) \mid \boldsymbol{m}_{B}(\lambda) \qquad ②$$

①+②得：$\boldsymbol{m}_A(\lambda)=\boldsymbol{m}_B(\lambda)$.

四、对角矩阵

Th24 在线性空间 V^n 中，线性变换 T 在某基下的矩阵为对角矩阵

$\Leftrightarrow T$ 有 n 个线性无关的特征向量（元素）.

证 必要性．设 V^n 的基为 $x_1,\cdots,x_n$，且 $T(x_1,\cdots,x_n)=(x_1,\cdots,x_n)\Lambda$，

$\Lambda=\mathrm{diag}(\lambda_1,\cdots,\lambda_n)$，则有

$$(Tx_1,\cdots,Tx_n)=(x_1,\cdots,x_n)\begin{bmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{bmatrix}=(\lambda_1x_1,\cdots,\lambda_nx_n)$$

$\Rightarrow Tx_j=\lambda_jx_j\quad(j=1,2,\cdots,n)$

$\Rightarrow x_1,\cdots,x_n$ 是 T 的 n 个线性无关的特征向量

充分性．设 T 有 n 个线性无关的特征向量 $y_1,\cdots,y_n$，即

$$T\,y_j=\lambda_jy_j,\ j=1,2,\cdots,n$$

取 $y_1,\cdots,y_n$ 为 V^n 的基，则有

$$T(y_1,\cdots,y_n)=(Ty_1,\cdots,Ty_n)=(\lambda_1y_1,\cdots,\lambda_ny_n)$$

$$=(y_1,\cdots,y_n)\begin{bmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{bmatrix}$$

Th25 $A_{n\times n}$ 相似于对角矩阵 $\Leftrightarrow A$ 有 n 个线性无关的特征向量（列向量）.

证 A 相似于 $\Lambda=\mathrm{diag}(\lambda_1,\cdots,\lambda_n)$

$\Leftrightarrow$ 存在可逆矩阵 $P=(x_1,\cdots,x_n)$，使得 $P^{-1}AP=\Lambda$

$\Leftrightarrow A(x_1,\cdots,x_n)=(x_1,\cdots,x_n)\Lambda\Leftrightarrow Ax_j=\lambda_jx_j,j=1,2,\cdots,n$

$\Leftrightarrow A$ 有 n 个线性无关的特征向量 $x_1,\cdots,x_n$

Th26 $A_{n\times n}$ 有 n 个互异的特征值 $\Rightarrow A$ 相似于对角矩阵.

算法：线性空间 V^n 的基 $x_1,\cdots,x_n$，线性变换 T 在该基下的矩阵 A 相似于

$\boldsymbol{\Lambda}=\mathbf{diag}(\lambda_1,\cdots,\lambda_n)$，确定$V^n$的新基$y_1,\cdots,y_n$，使得$T$在新基下的矩阵为$\Lambda$．求$P$使$P^{-1}AP=\Lambda$，令$(y_1,\cdots,y_n)=(x_1,\cdots,x_n)P$，则有

$$T(y_1,\cdots,y_n)=T(x_1,\cdots,x_n)P=(x_1,\cdots,x_n)AP$$
$$=(y_1,\cdots,y_n)P^{-1}AP=(y_1,\cdots,y_n)\Lambda$$

例 12　在$\mathbf{R}^{2\times2}$中，给定$B=\begin{bmatrix}0&1\\4&0\end{bmatrix}$，线性变换为　$TX=XB$　$(\forall X\in\mathbf{R}^{2\times2})$，求$\mathbf{R}^{2\times2}$的一个基，使线性变换$T$在该基下的矩阵为对角矩阵．

解　取$\mathbf{R}^{2\times2}$的简单基$E_{11},E_{12},E_{21},E_{22}$，求得$T$在该基下的矩阵为

$$A=\begin{bmatrix}0&4&0&0\\1&0&0&0\\0&0&0&4\\0&0&1&0\end{bmatrix}$$

求P使得$P^{-1}AP=\Lambda$：

$$\Lambda=\begin{bmatrix}2&&&\\&2&&\\&&-2&\\&&&-2\end{bmatrix},\quad P=\begin{bmatrix}2&0&2&0\\1&0&-1&0\\0&2&0&2\\0&1&0&-1\end{bmatrix}$$

由$(B_1,B_2,B_3,B_4)=(E_{11},E_{12},E_{21},E_{22})P$可得

$$B_1=\begin{bmatrix}2&1\\0&0\end{bmatrix},B_2=\begin{bmatrix}0&0\\2&1\end{bmatrix},B_3=\begin{bmatrix}2&-1\\0&0\end{bmatrix},B_4=\begin{bmatrix}0&0\\2&-1\end{bmatrix}$$

故T在基B_1,B_2,B_3,B_4下的矩阵为Λ．

五、不变子空间

线性空间V，子空间V_1，线性变换T.

若对$\forall x\in V_1$，有$Tx\in V_1$，称V_1是T的不变子空间.

[注]　V_1是T的不变子空间时，可将T看作V_1中的线性变换.

例 ① 子空间$V_{\lambda_0}=\{x\,|\,Tx=\lambda_0 x, x\in V\}$是 T 的不变子空间.

$\because \forall x\in V_{\lambda_0}\Rightarrow Tx=\lambda_0 x\in V_{\lambda_0}$

② 子空间 $R(T)$是 T 的不变子空间.

$\because \forall x\in R(T)\subset V\Rightarrow Tx\in R(T)$

③ 子空间 $N(T)$是 T 的不变子空间.

$\because \forall x\in N(T)\Rightarrow Tx=\theta\in N(T)$

④ V_1与 V_2是 T 的不变子空间$\Rightarrow V_1\cap V_2, V_1+V_2$亦是 T 的不变子空间.

$$1^\circ\quad \forall x\in V_1\cap V_2\Rightarrow\begin{Bmatrix} x\in V_1, Tx\in V_1\\ x\in V_2, Tx\in V_2\end{Bmatrix}\Rightarrow Tx\in V_1\cap V_2$$

$$2^\circ\quad \forall x\in V_1+V_2\Rightarrow x=x_1+x_2, x_1\in V_1, x_2\in V_2$$

$$\Rightarrow Tx=Tx_1+Tx_2, Tx_1\in V_1, Tx_2\in V_2$$

$$\Rightarrow Tx\in V_1+V_2$$

Th27 线性空间V^n，线性变换T，V_1与V_2是T的不变子空间，且 $V^n=V_1\oplus V_2$. T 在 V_1 的基 $x_1,\cdots,x_{n_1}$ 下的矩阵为 A_1，T 在 V_2 的基 $y_1,\cdots,y_{n_2}$ 下的矩阵为 A_2. 则 T 在V^n的基 $x_1,\cdots,x_{n_1},y_1,\cdots,y_{n_2}$ 下的矩阵为 $A=\begin{bmatrix} A_1 & O\\ O & A_2\end{bmatrix}$.

证 因为 $(Tx_1,\cdots,Tx_{n_1})=(x_1,\cdots,x_{n_1})A_1$，$(Ty_1,\cdots,Ty_{n_2})=(y_1,\cdots,y_{n_2})A_2$

所以
$$\begin{aligned}T(x_1,\cdots,x_{n_1},y_1,\cdots,y_{n_2})&=[(Tx_1,\cdots,Tx_{n_1}),(Ty_1,\cdots,Ty_{n_2})]\\&=[(x_1,\cdots,x_{n_1})A_1,(y_1,\cdots,y_{n_2})A_2]\\&=[(x_1,\cdots,x_{n_1}),(y_1,\cdots,y_{n_2})]\begin{bmatrix} A_1 & O\\ O & A_2\end{bmatrix}\\&=(x_1,\cdots,x_{n_1},y_1,\cdots,y_{n_2})A\end{aligned}$$

[注]　若 T 在 V^n 的基 $x_1,\cdots,x_{n_1},y_1,\cdots,y_{n_2}$ 下的矩阵 $A=\begin{bmatrix} A_1 & O \\ O & A_2 \end{bmatrix}$，则 $V_1 \stackrel{\Delta}{=} L(x_1,\cdots,x_{n_1})$ 与 $V_2 \stackrel{\Delta}{=} L(y_1,\cdots,y_{n_2})$ 都是 T 的不变子空间，且 $V^n = V_1 \oplus V_2$.

六、Jordan 标准形

1．$\lambda-$矩阵：$A(\lambda)=(a_{ij}(\lambda))_{n\times n}, a_{ij}(\lambda)$是$\lambda$的多项式.

$A(\lambda)$的秩：$A(\lambda)$中不恒等于零的子式的最高阶数.

$\lambda-$矩阵的初等变换：	行变换	列变换
(1) 对调：	$r_i \leftrightarrow r_j$	$c_i \leftrightarrow c_j$
(2) 数乘$(k\neq 0)$：	kr_i	kc_i
(3) 倍加（$p(\lambda)$是多项式）：	$r_i + p(\lambda)r_j$	$c_i + p(\lambda)c_j$

2．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：　$d_k(\lambda)=\dfrac{D_k(\lambda)}{D_{k-1}(\lambda)}$　$(D_0(\lambda)=1, k=1,2,\cdots,n)$

初等因子：　$d_k(\lambda)$的不可约因式

[注] 考虑$\lambda-$矩阵$\lambda I - A$可得A的最小多项式$m(\lambda)=d_n(\lambda)=\dfrac{D_n(\lambda)}{D_{n-1}(\lambda)}$

例 13　$A=\begin{bmatrix} -1 & 1 & 0 \\ -4 & 3 & 0 \\ 1 & 0 & 2 \end{bmatrix}$，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 $d_1(\lambda)=1,\quad d_2(\lambda)=1,\quad d_3(\lambda)=(\lambda-1)^2(\lambda-2)$.

全体初等因子为 $(\lambda-1)^2$，$\lambda-2$.

3．初等变换法求初等因子

$$A(\lambda)\to\begin{bmatrix} f_1(\lambda) & & \\ & \ddots & \\ & & f_n(\lambda)\end{bmatrix}\quad (f_k(\lambda)\text{是首1多项式})$$

$f_k(\lambda)$ 的不可约因式为 $A(\lambda)$ 的初等因子

例如：在例 13 中

$$\lambda I-A=\begin{bmatrix}\lambda+1 & -1 & 0\\ 4 & \lambda-3 & 0\\ -1 & 0 & \lambda-2\end{bmatrix}\xrightarrow{c_1\leftrightarrow c_2}\begin{bmatrix}-1 & \lambda+1 & 0\\ \lambda-3 & 4 & 0\\ 0 & -1 & \lambda-2\end{bmatrix}$$

$$\xrightarrow{r_2+(\lambda-3)r_1}\begin{bmatrix}-1 & \lambda+1 & 0\\ 0 & (\lambda-1)^2 & 0\\ 0 & -1 & \lambda-2\end{bmatrix}\xrightarrow[c_2-(\lambda+1)c_1]{(-1)c_1}\begin{bmatrix}1 & 0 & 0\\ 0 & (\lambda-1)^2 & 0\\ 0 & -1 & \lambda-2\end{bmatrix}$$

$$\xrightarrow{r_2\leftrightarrow r_3}\begin{bmatrix}1 & 0 & 0\\ 0 & -1 & \lambda-2\\ 0 & (\lambda-1)^2 & 0\end{bmatrix}\xrightarrow{r_3+(\lambda-1)^2r_2}\begin{bmatrix}1 & 0 & 0\\ 0 & -1 & \lambda-2\\ 0 & 0 & (\lambda-1)^2(\lambda-2)\end{bmatrix}$$

$$\xrightarrow[c_3-(\lambda-2)c_2]{(-1)c_2}\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & (\lambda-1)^2(\lambda-2)\end{bmatrix}$$

于是　$f_1(\lambda)=1, f_2(\lambda)=1, f_3(\lambda)=(\lambda-1)^2(\lambda-2)$.

故 $\lambda I-A$ 的全体初等因子为 $(\lambda-1)^2,\lambda-2$.

[注]　设 $A=(a_{ij})_{n\times n}$，称 $\lambda I-A$ 的行列式因子（不变因子，初等因子）为 A 的行列式因子（不变因子，初等因子）.

4．Jordan 标准形

设 $A=(a_{ij})_{n\times n}$ 的全体初等因子为

$$(\lambda-\lambda_1)^{m_1},\cdots,(\lambda-\lambda_i)^{m_i},\cdots,(\lambda-\lambda_s)^{m_s}$$

则有　　$\varphi(\lambda)=\det(\lambda I-A)=D_n(\lambda)=D_{n-1}(\lambda)d_n(\lambda)=\cdots$

$$=D_0(\lambda)d_1(\lambda)\cdots d_n(\lambda)=(\lambda-\lambda_1)^{m_1}\cdots(\lambda-\lambda_i)^{m_i}\cdots(\lambda-\lambda_s)^{m_s}$$

而且　　$m_1+\cdots+m_i+\cdots+m_s=n$

对于第 i 个初等因子 $(\lambda-\lambda_i)^{m_i}$ 构造 m_i 阶 Jordan 块矩阵 J_i，以及准对角矩阵 J 如下：

$$J_i=\begin{bmatrix}\lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i\end{bmatrix}_{m_i\times m_i},\quad J=\begin{bmatrix}J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_s\end{bmatrix}$$

称 J 为矩阵 A 的 **Jordan 标准形**.

Th29　设矩阵 A 的 Jordan 标准形为 J，则存在可逆矩阵 P，使得

$$P^{-1}AP=J\ .$$

例如：在例 13 中，A 的 Jordan 标准形为 $J=\left[\begin{array}{cc|c}1 & 1 & \\ & 1 & \\ \hline & & 2\end{array}\right]$.

[注] 若 A 的全体互异特征值为 $\lambda_1,\cdots,\lambda_l$，$m_i$ 表示 A 的 Jordan 标准形中含 λ_i 的 Jordan 块的最高阶数，则 $m(\lambda)=(\lambda-\lambda_1)^{m_1}\cdots(\lambda-\lambda_l)^{m_l}$.

5．特征向量分析法求初等因子

设 $\varphi(\lambda)=\det(\lambda I-A)$ 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（证明略去）

$\Leftrightarrow$ 对应特征值 λ_0 有 k 个线性无关的特征向量

$\Leftrightarrow$ $k=n-\mathrm{rank}(\lambda_0 I-A)$

例 14 求 $A=\begin{bmatrix}1&&&\\2&1&&\\1&3&2&\\2&3&1&1\end{bmatrix}$ 的 Jordan 标准形.

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 的全体初等因子为 $(\lambda-1)^2,\lambda-1,\lambda-2$.

A 的 Jordan 标准形为 $J=\begin{bmatrix}1&1&&\\&1&&\\&&1&\\&&&2\end{bmatrix}$.

[注] 在例 14 中，将 $a_{33}=2$，$a_{43}=1$ 改作 $a_{33}=1$，$a_{43}=0$ 时，此法失效.

6．相似变换矩阵的求法

仅适用于初等因子组中 $\lambda_i\neq\lambda_j(i\neq j)$ 的情形.

$$P=(P_1,\cdots,P_s),\ P_i=\left(X_1^{(i)},\cdots,X_{m_i}^{(i)}\right)$$

$$AP=PJ\Leftrightarrow AP_i=P_iJ_i,\ i=1,2,\cdots,s$$

$$\left(AX_1^{(i)},AX_2^{(i)},\cdots,AX_{m_i}^{(i)}\right)=\left(\lambda_iX_1^{(i)},X_1^{(i)}+\lambda_iX_2^{(i)},\cdots,X_{m_i-1}^{(i)}+\lambda_iX_{m_i}^{(i)}\right)$$

$$\begin{cases}(\lambda_iI-A)X_1^{(i)}=0 & X_1^{(i)}\text{ 是 }(\lambda_iI-A)X=0\text{ 的非零解}\\(\lambda_iI-A)X_2^{(i)}=-X_1^{(i)} & X_2^{(i)}\text{ 是 }(\lambda_iI-A)X=-X_1^{(i)}\text{ 的一个解}\\\cdots\cdots\cdots & \cdots\cdots\cdots\\(\lambda_iI-A)X_{m_i}^{(i)}=-X_{m_i-1}^{(i)} & X_{m_i}^{(i)}\text{ 是 }(\lambda_iI-A)X=-X_{m_i-1}^{(i)}\text{ 的一个解}\end{cases}$$

可以证明：$X_1^{(i)},X_2^{(i)},\cdots,X_{m_i}^{(i)}$ 线性无关.

在例 13 中，$\lambda_1=1,m_1=2$，求 $X_1^{(1)},X_2^{(1)}$：

$$\lambda_1I-A=\begin{bmatrix}2&-1&0\\4&-2&0\\-1&0&-1\end{bmatrix},\quad X_1^{(1)}=\begin{bmatrix}1\\2\\-1\end{bmatrix}$$

$$\left[\lambda_1 I - A,\ -X_1^{(1)}\right] = \begin{bmatrix} 2 & -1 & 0 & -1 \\ 4 & -2 & 0 & -2 \\ -1 & 0 & -1 & 1 \end{bmatrix},\quad X_2^{(1)} = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$$

$\lambda_2 = 2, m_2 = 1$，求 $X_1^{(2)}$：

$$\lambda_2 I - A = \begin{bmatrix} 3 & -1 & 0 \\ 4 & -1 & 0 \\ -1 & 0 & 0 \end{bmatrix},\quad X_1^{(2)} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.\ \text{故 } P = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -1 & 1 \end{bmatrix}.$$

例 15　解线性微分方程组 $\begin{cases} \xi_1'(t) = -\xi_1 + \xi_2 \\ \xi_2'(t) = -4\xi_1 + 3\xi_2 \\ \xi_3'(t) = \xi_1 \quad + 2\xi_3 \end{cases}$.

解　$x(t) = \begin{bmatrix} \xi_1(t) \\ \xi_2(t) \\ \xi_3(t) \end{bmatrix}, x'(t) = \begin{bmatrix} \xi_1'(t) \\ \xi_2'(t) \\ \xi_3'(t) \end{bmatrix}, A = \begin{bmatrix} -1 & 1 & 0 \\ -4 & 3 & 0 \\ 1 & 0 & 2 \end{bmatrix}$：$x'(t) = Ax(t)$

已求得 $P = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -1 & 1 \end{bmatrix}$，使得 $P^{-1}AP = J = \left[\begin{array}{cc|c} 1 & 1 & \\ & 1 & \\ \hline & & 2 \end{array}\right]$，则有

$$P^{-1}x'(t) = P^{-1}APP^{-1}x(t) \Rightarrow \left[P^{-1}x(t)\right]' = J\left[P^{-1}x(t)\right]$$

$$y(t) \overset{\Delta}{=} P^{-1}x(t) = \begin{bmatrix} \eta_1(t) \\ \eta_2(t) \\ \eta_3(t) \end{bmatrix} \Rightarrow y'(t) = J\,y(t)$$

$$\begin{cases} \eta_1'(t) = \eta_1 + \eta_2 \\ \eta_2'(t) = \quad \eta_2 \\ \eta_3'(t) = \quad 2\eta_3 \end{cases} \Rightarrow \begin{cases} \eta_2(t) = c_2 e^t \\ \eta_1'(t) = \eta_1 + c_2 e^t \Rightarrow \eta_1(t) = c_1 e^t + c_2 t e^t \\ \eta_3(t) = c_3 e^{2t} \end{cases}$$

$$x(t) = P\,y(t) = \begin{bmatrix} \eta_1 \\ 2\eta_1 + \eta_2 \\ -\eta_1 - \eta_2 + \eta_3 \end{bmatrix} \Rightarrow \begin{cases} \xi_1(t) = c_1 e^t + c_2 t e^t \\ \xi_2(t) = 2c_1 e^t + c_2(2t+1)e^t \\ \xi_3(t) = -c_1 e^t - c_2(t+1) + c_3 e^{2t} \end{cases}$$

$(c_1, c_2, c_3$ 为任意常数)

［注］$\eta_1(t) = e^t\{c_1 + \int_0^t e^{-\tau}\eta_2(\tau)d\tau\}$

求线性变换在给定基下的矩阵——方法总结：

给定线性空间 V^n 的基 $x_1,\cdots,x_n$，设线性变换 T 在该基下的矩阵为 A .

一、直接法

(1) 计算基象组 $T(x_1),\cdots,T(x_n)$，并求出 $T(x_j)$ 在基 $x_1,\cdots,x_n$ 下的坐标（列向量）$\beta_j\ (j=1,2,\cdots,n)$；

(2) 写出 T 在给定基 $x_1,\cdots,x_n$ 下的矩阵 $A=(\beta_1,\cdots,\beta_n)$.

二、中介法

(1) 选取 V^n 的简单基，记作 $\varepsilon_1,\cdots,\varepsilon_n$；

（要求 V^n 中元素在该基下的坐标能够直接写出）

(2) 写出由简单基改变为给定基的过渡矩阵 C（采用直接法）；

(3) 计算基象组 $T(\varepsilon_1),\cdots,T(\varepsilon_n)$，并写出 $T(\varepsilon_j)$ 在简单基 $\varepsilon_1,\cdots,\varepsilon_n$ 下的坐标（列向量）$\beta_j\ (j=1,2,\cdots,n)$，以及 T 在简单基下的矩阵 $B=(\beta_1,\cdots,\beta_n)$；

(4) 计算 T 在给定基 $x_1,\cdots,x_n$ 下的矩阵 $A=C^{-1}BC$.

三、混合法

(1) 选取 V^n 的简单基，记作 $\varepsilon_1,\cdots,\varepsilon_n$；

(2) 写出由简单基改变为给定基的过渡矩阵 C（采用直接法），则有

$$(x_1,\cdots,x_n)=(\varepsilon_1,\cdots,\varepsilon_n)C$$

(3) 计算基象组 $T(x_1),\cdots,T(x_n)$，并写出 $T(x_j)$ 在在简单基 $\varepsilon_1,\cdots,\varepsilon_n$ 下的坐标（列向量）$\beta_j\ (j=1,2,\cdots,n)$，以及矩阵 $B=(\beta_1,\cdots,\beta_n)$，则有

$$\begin{aligned}T(x_1,\cdots,x_n)&=(\,T(x_1),\cdots,T(x_n)\,)\\&=(\varepsilon_1,\cdots,\varepsilon_n)B=(x_1,\cdots,x_n)C^{-1}B\end{aligned}$$

(4) 计算 T 在给定基 $x_1,\cdots,x_n$ 下的矩阵 $A=C^{-1}B$.

§1.3 欧氏空间与酉空间

一、欧氏空间

1. 内积：线性空间 V，数域 $\mathbf{R}$，对 $\forall x, y \in V$，定义实数 (x, y)，且满足

(1) 交换律　$(x, y) = (y, x)$

(2) 分配律　$(x, y + z) = (x, y) + (x, z), \quad \forall z \in V$

(3) 齐次性　$(kx, y) = k(x, y), \quad \forall k \in \mathbf{R}$

(4) 非负性　$(x, x) \geq 0, (x, x) = 0 \Leftrightarrow x = \theta$

称实数 (x, y) 为 x 与 y 的内积.

例 ① 线性空间 $\mathbf{R}^n$ 中：$x = (\xi_1, \cdots, \xi_n), y = (\eta_1, \cdots, \eta_n)$

内积 1：$(x, y) \stackrel{\Delta}{=} \xi_1\eta_1 + \cdots + \xi_n\eta_n$

内积 2：$(x, y)_h \stackrel{\Delta}{=} h(\xi_1\eta_1 + \cdots + \xi_n\eta_n), \quad h > 0$

② 线性空间 $\mathbf{R}^{m\times n}$ 中：$A = (a_{ij})_{m\times n}, B = (b_{ij})_{m\times n}$

内积：$(A, B) \stackrel{\Delta}{=} \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}b_{ij} = \mathrm{tr}(AB^{\mathrm{T}})$

③ 线性空间 $C[a, b]$ 中：$f(t), g(t)$ 是区间 $[a, b]$ 上的连续函数

内积：$(f(t), g(t)) \stackrel{\Delta}{=} \int_a^b f(t)g(t)dt$

2. 欧氏空间：定义了内积运算的实线性空间.

设欧氏空间 V^n 的基为 $x_1, \cdots, x_n$，对 $\forall x, y \in V^n$ 有

$$\left.\begin{aligned} x &= \xi_1 x_1 + \cdots + \xi_n x_n \\ y &= \eta_1 x_1 + \cdots + \eta_n x_n \end{aligned}\right\} \Rightarrow (x, y) = \sum_{i,j=1}^{n} \xi_i \eta_j (x_i, x_j)$$

令　$a_{ij} = (x_i, x_j)$　（$i, j = 1, 2, \cdots, n$）

则称 $A = (a_{ij})_{n\times n}$ 为基 $x_1, \cdots, x_n$ 的度量矩阵（**Gram Matrix**），此时有

$$(x,y)=\sum_{i,j=1}^{n}a_{ij}\xi_i\eta_j=(\xi_1,\cdots,\xi_n)A\begin{bmatrix}\eta_1\\ \vdots\\ \eta_n\end{bmatrix}$$

① A 对称：$(x_i,x_j)=(x_j,x_i)\Rightarrow a_{ij}=a_{ji}$

② A 正定：$\forall\xi_1,\cdots,\xi_n\in\mathbf{R}$ 不全为零 $\Rightarrow x\overset{\Delta}{=}\xi_1x_1+\cdots+\xi_nx_n\neq\theta$

$$(x,x)>0\Rightarrow \text{二次型}(\xi_1,\cdots,\xi_n)A\begin{bmatrix}\xi_1\\ \vdots\\ \xi_n\end{bmatrix}>0，\text{即 } A \text{ 正定.}$$

③ V^n 中不同基的度量矩阵是合同的：

基(I): $x_1,\cdots,x_n$；基(II): $y_1,\cdots,y_n$

基(I)到基(II)的过渡矩阵为 C：$(y_1,\cdots,y_n)=(x_1,\cdots,x_n)C$

$$y_i=c_{1i}x_1+\cdots+c_{ni}x_n，\ y_j=c_{1j}x_1+\cdots+c_{nj}x_n$$

$$b_{ij}=(y_i,y_j)=(c_{1i},\cdots,c_{ni})A\begin{bmatrix}c_{1j}\\ \vdots\\ c_{nj}\end{bmatrix}$$

$$(b_{i1},\cdots,b_{in})=(c_{1i},\cdots,c_{ni})AC$$

$$B=\begin{bmatrix}b_{11}&\cdots&b_{1n}\\ \vdots&&\vdots\\ b_{n1}&\cdots&b_{nn}\end{bmatrix}=\begin{bmatrix}c_{11}&\cdots&c_{n1}\\ \vdots&&\vdots\\ c_{1n}&\cdots&c_{nn}\end{bmatrix}AC=C^{\mathrm{T}}AC$$

④ “内积值”的计算与基的选取无关：$\forall x,y\in V^n$

$$\left.\begin{aligned}x&=(x_1,\cdots,x_n)\alpha_1\\ x&=(y_1,\cdots,y_n)\alpha_2=(x_1,\cdots,x_n)C\alpha_2\end{aligned}\right\}\Rightarrow\alpha_1=C\alpha_2$$

$$\left.\begin{aligned}y&=(x_1,\cdots,x_n)\beta_1\\ y&=(y_1,\cdots,y_n)\beta_2=(x_1,\cdots,x_n)C\beta_2\end{aligned}\right\}\Rightarrow\beta_1=C\beta_2$$

基(I)下：$(x,y)=\alpha_1^{\mathrm{T}}A\beta_1$

基(II)下：$(x,y)=\alpha_2^{\mathrm{T}}B\beta_2=\alpha_2^{\mathrm{T}}C^{\mathrm{T}}AC\beta_2=\alpha_1^{\mathrm{T}}A\beta_1$

3．元素的模：欧氏空间 $V,\forall x\in V$，称 $|x|\overset{\Delta}{=}\sqrt{(x,x)}$ 为 x 的模（长度）.

① $\forall k\in\mathbf{R},\ |kx|=|k|\,|x|$

② $|(x,y)|\le|x|\,|y|$；　$x\ne\theta,y\ne\theta$ 时，等号成立 $\Leftrightarrow x,y$ 线性相关.

证　$\forall t\in\mathbf{R},\ x-ty\in V\Rightarrow|x-ty|^2=(x-ty,x-ty)\ge 0$

$(y,y)t^2-2(x,y)t+(x,x)\ge 0$

$\varDelta\le 0:4(x,y)^2-4(y,y)(x,x)\le 0\Rightarrow|(x,y)|\le|x||y|$

充分性．已知 x,y 线性相关，且 $x\ne\theta$，所以 $y=kx$，从而

$$|(x,y)|=|(x,kx)|=|k||(x,x)|=|k||x||x|=|x||y|$$

必要性．已知 $|(x,y)|=|x||y|$，则

$(x,y)\ge 0$ 时，取　$t=\dfrac{|x|}{|y|}\Rightarrow(x-ty,x-ty)=\cdots=0\Rightarrow x-ty=\theta$

$(x,y)<0$ 时，取　$t=-\dfrac{|x|}{|y|}\Rightarrow(x-ty,x-ty)=\cdots=0\Rightarrow x-ty=\theta$

故 x,y 线性相关.

③ $|x+y|\le|x|+|y|$

④ $|x-y|\ge\big||x|-|y|\big|$

4．元素之间的夹角：欧氏空间 V 中，$x\ne\theta,y\ne\theta$，称

$\phi\overset{\varDelta}{=}\arccos\dfrac{(x,y)}{|x||y|}\in[0,\pi]$ 为 x 与 y 之间的夹角.

二、正交性

欧氏空间 V 中，若 $(x,y)=0$，称 x 与 y 正交，记作 $x\perp y$；

若 $x_1,\cdots,x_m$ 满足 $(x_i,x_j)=0\ (i\ne j)$，称 $x_1\cdots,x_m$ 为正交元素组.

Th 31　$x\perp y\Rightarrow|x+y|^2=|x|^2+|y|^2$.

Th 32　$x_1,\cdots,x_m$ 两两正交且非零 $\Rightarrow x_1,\cdots,x_m$ 线性无关.

[注]　欧氏空间 V^n 中，两两正交的非零元素的个数不超过 n.

1． 正交基：欧氏空间 V^n 中，若 $x_1,\cdots,x_n$ 两两正交且非零，称 $x_1,\cdots,x_n$ 为 V^n 的正交基；若还有 $|x_i|=1\ (i=1,2,\cdots,n)$，称 $x_1,\cdots,x_n$ 为 V^n 的标准正交基.

Schmidt 正交化方法：设欧氏空间 V^n 的一个基为 $x_1,\cdots,x_n$，构造元素组 $y_1,\cdots,y_n$，使满足 $y_i\perp y_j\ (i\neq j)$ 且 $y_i\neq\theta\ (i=1,2,\cdots,n)$.

$y_1=x_1$,　　　$y_1\neq\theta$

$y_2=x_2+k_{21}y_1$,　　$y_2\neq\theta$

$$(y_2,y_1)=0\Rightarrow 0=(x_2,y_1)+k_{21}(y_1,y_1)\Rightarrow k_{21}=-\frac{(x_2,y_1)}{(y_1,y_1)}$$

$y_3=x_3+k_{32}y_2+k_{31}y_1$,　　$y_3\neq\theta$

$$(y_3,y_1)=0\Rightarrow k_{31}=-\frac{(x_3,y_1)}{(y_1,y_1)}$$

$$(y_3,y_2)=0\Rightarrow k_{32}=-\frac{(x_3,y_2)}{(y_2,y_2)}$$

$\cdots\cdots\cdots\cdots\cdots\cdots$

$y_n=x_n+k_{n,n-1}y_{n-1}+\cdots+k_{n1}y_1$,　　$y_n\neq\theta$

$$(y_n,y_j)=0\Rightarrow k_{nj}=-\frac{(x_n,y_j)}{(y_j,y_j)},\quad j=1,2,\cdots,n-1$$

因为 $y_1,\cdots,y_n$ 两两正交且非零，所以线性无关，从而是 V^n 的正交基.

Th 33 欧氏空间 $V^n\ (n\geq 1)$ 存在标准正交基.

证　对 V^n 的基 $x_1,\cdots,x_n$ 进行正交化，可得正交基 $y_1,\cdots,y_n$；

再进行单位化，可得标准正交基 $z_1,\cdots,z_n$　$(z_j=\dfrac{1}{|y_j|}y_j)$.

例 1 标准正交基的特征：设欧氏空间 V^n 的标准正交基为 $x_1,\cdots,x_n$，

且 $x=\xi_1x_1+\cdots+\xi_nx_n$，$y=\eta_1x_1+\cdots+\eta_nx_n$，则有

(1) 基 $\boldsymbol{x}_1,\cdots,\boldsymbol{x}_n$ 的度量矩阵 $\boldsymbol{A}=\boldsymbol{I}$.

(2) $\xi_i=(\boldsymbol{x},\boldsymbol{x}_i)$, $\eta_j=(\boldsymbol{y},\boldsymbol{x}_j)$.

(3) $(\boldsymbol{x},\boldsymbol{y})=\xi_1\eta_1+\cdots+\xi_n\eta_n$.

2．子空间的正交性：欧氏空间V的子空间 V_1，给定 $\boldsymbol{y}\in V$，若$\forall \boldsymbol{x}\in V_1$，

都有 $\boldsymbol{y}\perp\boldsymbol{x}$，则称 $\boldsymbol{y}$ 正交于 V_1，记作 $\boldsymbol{y}\perp V_1$.

例 2　欧氏空间V^n，子空间V_1的基为 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_m$，则

$$\boldsymbol{y}\perp V_1 \Leftrightarrow \boldsymbol{y}\perp\boldsymbol{x}_j\,,\ j=1,2,\cdots,m$$

证 必要性．（略）

充分性． $\forall \boldsymbol{x}\in V_1 \Rightarrow \boldsymbol{x}=k_1\boldsymbol{x}_1+\cdots+k_m\boldsymbol{x}_m$

$$\boldsymbol{y}\perp\boldsymbol{x}_j \Rightarrow (\boldsymbol{y},\boldsymbol{x}_j)=0 \Rightarrow (\boldsymbol{y},\boldsymbol{x})=0 \quad \text{即} \quad \boldsymbol{y}\perp V_1$$

例 3 欧氏空间V，子空间 V_1，则$V_1^{\perp}\overset{\Delta}{=}\{\boldsymbol{y}\mid \boldsymbol{y}\in V \text{ 且 } \boldsymbol{y}\perp V_1\}$是$V$的子空间.

证　$\theta\in V_1^{\perp} \Rightarrow V_1^{\perp}$非空．$\forall \boldsymbol{y},\boldsymbol{z}\in V_1^{\perp}, \forall \boldsymbol{x}\in V_1, \forall k\in \mathbf{R}$，有

$$(\boldsymbol{y}+\boldsymbol{z},\boldsymbol{x})=(\boldsymbol{y},\boldsymbol{x})+(\boldsymbol{z},\boldsymbol{x})=0 \Rightarrow (\boldsymbol{y}+\boldsymbol{z})\perp V_1:\ (\boldsymbol{y}+\boldsymbol{z})\in V_1^{\perp}$$

$$(k\boldsymbol{y},\boldsymbol{x})=k(\boldsymbol{y},\boldsymbol{x})=0 \qquad \Rightarrow (k\boldsymbol{y})\perp V_1:\quad (k\boldsymbol{y})\in V_1^{\perp}$$

故 $V_1^{\perp}$是V的子空间（称$V_1^{\perp}$为V_1的正交补）.

Th 34　设欧氏空间V^n，子空间V_1，则$V^n=V_1\oplus V_1^{\perp}$.

证　若$V_1=\{\theta\}$，则 $V_1^{\perp}=V^n \Rightarrow V^n=\{\theta\}\oplus V^n=V_1\oplus V_1^{\perp}$

若$V_1\neq\{\theta\}$，记$\dim V_1\overset{\Delta}{=}m\le n$，设 V_1 的标准正交基为 $\boldsymbol{x}_1,\cdots,\boldsymbol{x}_m$.

① 先证$V^n=V_1+V_1^{\perp}$：只需证明 $V^n\subset V_1+V_1^{\perp}$即可.

$$\forall\, \boldsymbol{x}\in V^n, \text{记 } a_i=(\boldsymbol{x}_i,\boldsymbol{x}), i=1,2,\cdots,m\,,\quad \boldsymbol{y}\overset{\Delta}{=}a_1\boldsymbol{x}_1+\cdots+a_m\boldsymbol{x}_m\in V_1$$

因为 $(\boldsymbol{x}-\boldsymbol{y},\boldsymbol{x}_i)=(\boldsymbol{x},\boldsymbol{x}_i)-(\boldsymbol{y},\boldsymbol{x}_i)=a_i-a_i=0$，所以由例 2 可得

$$(\boldsymbol{x}-\boldsymbol{y})\perp V_1 \Rightarrow \boldsymbol{z}\overset{\Delta}{=}\boldsymbol{x}-\boldsymbol{y}\in V_1^{\perp}$$

故　$x = y + z,\ y \in V_1,\ z \in V_1^{\perp}$，即　$x \in V_1 + V_1^{\perp}$.

② 再证 $V_1 \cap V_1^{\perp} = \{\theta\}$：$\forall\, x \in V_1 \cap V_1^{\perp} \Rightarrow \begin{cases} x \in V_1 \\ x \in V_1^{\perp} \end{cases} \Rightarrow (x, x) = 0 \Rightarrow x = \theta$

①+②：结论成立.

Th 35 设 $A = (a_{ij})_{m\times n} \in \mathbf{R}^{m\times n}$，则

(1) $[R(A)]^{\perp} = N(A^{\mathrm{T}})$, 且 $R(A) \oplus N(A^{\mathrm{T}}) = \mathbf{R}^m$；

(2) $[R(A^{\mathrm{T}})]^{\perp} = N(A)$, 且 $R(A^{\mathrm{T}}) \oplus N(A) = \mathbf{R}^n$.

证　划分 $A = (\beta_1, \cdots, \beta_n),\ \beta_j \in \mathbf{R}^m$（列向量）

$$V_1 \overset{\Delta}{=} R(A) = L(\beta_1, \cdots, \beta_n) \subset \mathbf{R}^m$$

$$\begin{aligned} V_1^{\perp} &= \{ y \mid y \in \mathbf{R}^m \text{且}\ y \perp (k_1\beta_1 + \cdots + k_n\beta_n) \} \subset \mathbf{R}^m \\ &= \{ y \mid y \in \mathbf{R}^m \text{且}\ y \perp \beta_j, j = 1, 2, \cdots, n \} \\ &= \{ y \mid y \in \mathbf{R}^m \text{且}\ \beta_j^{\mathrm{T}} y = 0, j = 1, 2, \cdots, n \} \\ &= \{ y \mid y \in \mathbf{R}^m \text{且}\ A^{\mathrm{T}} y = 0 \} = N(A^{\mathrm{T}}) \end{aligned}$$

由 Th34 可得　$\mathbf{R}^m = V_1 \oplus V_1^{\perp} = R(A) \oplus N(A^{\mathrm{T}})$.

对 A^{T} 应用上述结果可得　$[R(A^{\mathrm{T}})]^{\perp} = N(A)$

再由 Th34 可得　$\mathbf{R}^n = R(A^{\mathrm{T}}) \oplus N(A)$.

三、正交变换与正交矩阵

欧氏空间 V 中，若线性变换 T 满足 $(Tx, Tx) = (x, x)\quad (\forall x \in V)$,

称 T 为正交变换.

Th36 欧氏空间 V，线性变换 T.

$$T\text{是正交变换} \Leftrightarrow \forall x, y \in V, (Tx, Ty) = (x, y).$$

证　充分性. 取 $y = x$，则 $\forall x \in V, (Tx, Tx) = (x, x)$.

必要性. T 是正交变换：$\forall x, y \in V \Rightarrow x - y \in V$

$$(T(x-y),T(x-y))=(x-y,x-y)\Rightarrow\cdots\Rightarrow(Tx,Ty)=(x,y)$$

Th 37　欧氏空间 V^n 的标准正交基为 $x_1,\cdots,x_n$

线性变换 T ：$T(x_1,\cdots,x_n)=(x_1,\cdots,x_n)A$

则 T 是正交变换 $\Leftrightarrow$ A 是正交矩阵.

证　必要性．T 是正交变换：设 $A=(a_{ij})_{n\times n}$，则

$Tx_i=a_{1i}x_1+\cdots+a_{ni}x_n$，$Tx_j=a_{1j}x_1+\cdots+a_{nj}x_n$

$$(Tx_i,Tx_j)=(x_i,x_j)\Rightarrow\sum_{k=1}^{n}a_{ki}a_{kj}=\delta_{ij}\Rightarrow A^{\mathrm{T}}A=I$$

充分性．A 是正交矩阵：

$$\forall x\in V\Rightarrow x=(x_1,\cdots,x_n)\begin{bmatrix}\xi_1\\ \vdots\\ \xi_n\end{bmatrix},Tx=(x_1,\cdots,x_n)A\begin{bmatrix}\xi_1\\ \vdots\\ \xi_n\end{bmatrix}$$

$$(Tx,Tx)=(\xi_1,\cdots,\xi_n)A^{\mathrm{T}}A\begin{bmatrix}\xi_1\\ \vdots\\ \xi_n\end{bmatrix}=(\xi_1,\cdots,\xi_n)\begin{bmatrix}\xi_1\\ \vdots\\ \xi_n\end{bmatrix}=(x,x)$$

[注] 欧氏空间 V^n 的标准正交基具有如下性质：

(1) $x_1,\cdots,x_n$ 是标准正交基，T 是正交变换 $\Rightarrow$ $Tx_1,\cdots,Tx_n$ 是标准正交基

(2) $\left.\begin{array}{l}x_1,\cdots,x_n \text{ 和 } y_1,\cdots,y_n \text{ 都是标准正交基}\\ (y_1,\cdots,y_n)=(x_1,\cdots,x_n)C\end{array}\right\}\Rightarrow C$ 是正交矩阵

四、对称变换与对称矩阵

欧氏空间 V 中，若线性变换 T 满足 $(Tx,y)=(x,Ty)$ $(\forall x,y\in V)$，

称 T 是对称变换.

Th 38　欧氏空间 V^n 的标准正交基为 $x_1,\cdots,x_n$

线性变换 T ：$T(x_1,\cdots,x_n)=(x_1,\cdots,x_n)A$

则 T 是对称变换 $\Leftrightarrow$ A 是对称矩阵.

证　设 $A=(a_{ij})_{n\times n}$，则 $Tx_i=a_{1i}x_1+\cdots+a_{ni}x_n \Rightarrow (Tx_i,x_j)=a_{ji}$

$$Tx_j=a_{1j}x_1+\cdots+a_{nj}x_n \Rightarrow (x_i,Tx_j)=a_{ij}$$

必要性．T 是对称变换 $\Rightarrow (Tx_i,x_j)=(x_i,Tx_j) \Rightarrow a_{ji}=a_{ij}$，即 $A^{\mathrm{T}}=A$

充分性．A 是对称矩阵，则有

$$\forall x\in V^n \Rightarrow x=(x_1,\cdots,x_n)\begin{bmatrix}\xi_1\\ \vdots\\ \xi_n\end{bmatrix}, Tx=(x_1,\cdots,x_n)A\begin{bmatrix}\xi_1\\ \vdots\\ \xi_n\end{bmatrix}$$

$$\forall y\in V^n \Rightarrow y=(x_1,\cdots,x_n)\begin{bmatrix}\eta_1\\ \vdots\\ \eta_n\end{bmatrix}, Ty=(x_1,\cdots,x_n)A\begin{bmatrix}\eta_1\\ \vdots\\ \eta_n\end{bmatrix}$$

$$(Tx,y)=(\xi_1,\cdots,\xi_n)A^{\mathrm{T}}\cdot\begin{bmatrix}\eta_1\\ \vdots\\ \eta_n\end{bmatrix}=(\xi_1,\cdots,\xi_n)\cdot A\begin{bmatrix}\eta_1\\ \vdots\\ \eta_n\end{bmatrix}=(x,Ty)$$

Th 39　$A\in\mathbf{R}^{n\times n}$ 且 $A^{\mathrm{T}}=A \Rightarrow \lambda_A\in\mathbf{R}$．

证　设 $Ax=\lambda x\ (x\neq\theta)$，则 $x^{\mathrm{H}}Ax=\begin{cases}x^{\mathrm{H}}(Ax)=\lambda(x^{\mathrm{H}}x)\\(Ax)^{\mathrm{H}}x=\bar{\lambda}(x^{\mathrm{H}}x)\end{cases}$.

故 $(\lambda-\bar{\lambda})(x^{\mathrm{H}}x)=0 \Rightarrow \lambda=\bar{\lambda}\ \ (\because x^{\mathrm{H}}x>0)$，即 $\lambda\in\mathbf{R}$．

[注]　因为 $(\lambda I-A)x=0$ 是实系数齐次线性方程组，所以可求得非零解向量 $x\in\mathbf{R}^n$．因此，约定实对称矩阵的特征向量为实向量.

Th 40　设实对称矩阵 A 的特征值 $\lambda_1\neq\lambda_2$，对应的特征向量分别为 x_1 和 x_2，则 $(x_1,x_2)=0$．

证　$\left.\begin{aligned}Ax_1=\lambda_1x_1\\Ax_2=\lambda_2x_2\end{aligned}\right\}\Rightarrow x_1^{\mathrm{T}}Ax_2=\begin{cases}x_1^{\mathrm{T}}(Ax_2)=\lambda_2(x_1^{\mathrm{T}}x_2)\\(Ax_1)^{\mathrm{T}}x_2=\lambda_1(x_1^{\mathrm{T}}x_2)\end{cases}$

$$\Rightarrow(\lambda_2-\lambda_1)(x_1^{\mathrm{T}}x_2)=0 \Rightarrow x_1^{\mathrm{T}}x_2=0，即 (x_1,x_2)=0.$$

五、酉空间简介

1．复内积　线性空间 V，复数域 K，对 $\forall x,y\in V$，定义复数 (x,y)，且满足

(1) $(x,y)=\overline{(y,x)}$

(2) $(x,y+z)=(x,y)+(x,z),\forall z\in V$

(3) $(kx,y)=k(x,y),\forall k\in K$

(4) $(x,x)\geq 0,(x,x)=0\Leftrightarrow x=\theta$

称复数(x, y)为 x 与 y 的复内积.

2． 酉空间　定义了复内积运算的复线性空间.

① $(x,ky)=\bar{k}(x,y)$

② 基的度量矩阵为 **Hermite** 正定矩阵

③ $(x,y)=(\xi_1,\cdots,\xi_n)A\begin{bmatrix}\overline{\eta_1}\\ \vdots\\ \overline{\eta_n}\end{bmatrix}$

④ 酉变换：$(Tx, Tx)=(x, x)$　$(\forall x\in V)$

T 是酉变换$\Leftrightarrow T$ 在标准正交基下的矩阵 A 是酉矩阵，即 $A^{\mathrm{H}}A=I$.

⑤ **Hermite-**变换：$(Tx, y)=(x, Ty)$　$(\forall x,y\in V)$

T 是 **Hermite-**变换$\Leftrightarrow$

T 在标准正交基下的矩阵 A 是 **Hermite** 矩阵，即 $A^{\mathrm{H}}=A$.

Th41 (1) 设 $A_{n\times n}$ 的特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_n$，则存在酉矩阵 $P_{n\times n}$，使得

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

(2) 设 $A\in\mathbf{R}^{n\times n}$ 且 $\lambda_A\in\mathbf{R}$，则存在正交矩阵 $Q_{n\times n}$，使得 $Q^{\mathrm{T}}AQ$ 为上三角矩阵.

证　在 **Th17** 的证明过程中将“扩充 x_1 为 $\mathbf{C}^k$ 的基”改为

“扩充 $\boldsymbol{x}_1$ 为 $\mathbf{C}^k$ 的标准正交基”即可.

3. 正规矩阵：指 $A_{n\times n}$ 满足 $A^{\mathrm{H}}A = AA^{\mathrm{H}}$.

例如　$\mathbf{C}^{n\times n}$ 中：① $A^{\mathrm{H}} = A \Rightarrow A$ 正规；② $A^{\mathrm{H}}A = I \Rightarrow A$ 正规.

$\mathbf{R}^{n\times n}$ 中：① $A^{\mathrm{T}} = A \Rightarrow A$ 正规；② $A^{\mathrm{T}}A = I \Rightarrow A$ 正规.

$$B = \begin{bmatrix} 5+4\mathrm{j} & 1+6\mathrm{j} \\ 1+6\mathrm{j} & 5+4\mathrm{j} \end{bmatrix}:\quad B^{\mathrm{H}}B = \begin{bmatrix} 78 & 58 \\ 58 & 78 \end{bmatrix} = BB^{\mathrm{H}} \Rightarrow B\text{ 是正规矩阵}$$

但是 $B^{\mathrm{H}} \neq B,\ B^{\mathrm{H}}B \neq I$.

Th42 **(1)**　$A \in \mathbf{C}^{n\times n}$, A 正规 $\Leftrightarrow \exists$ 酉矩阵 $P_{n\times n}$ ，使得 $P^{\mathrm{H}}AP = \Lambda$ ；

(2)　$A \in \mathbf{R}^{n\times n}$ 且 $\lambda_A \in \mathbf{R}$ ，A 正规 $\Leftrightarrow \exists$ 正交矩阵 $Q_{n\times n}$ ，使得 $Q^{\mathrm{T}}AQ = \Lambda$.

证　**(1)** 充分性.　$A = P\Lambda P^{\mathrm{H}},\ A^{\mathrm{H}} = P\overline{\Lambda}P^{\mathrm{H}}$

$$A^{\mathrm{H}}A = P\overline{\Lambda}\Lambda P^{\mathrm{H}} = P\Lambda\overline{\Lambda}P^{\mathrm{H}} = AA^{\mathrm{H}}$$

必要性.　$A^{\mathrm{H}}A = AA^{\mathrm{H}}$:

$$\text{th41(1)} \Rightarrow \exists\text{ 酉矩阵 } P \text{ 使得 } P^{\mathrm{H}}AP = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ & b_{22} & \cdots & b_{2n} \\ & & \ddots & \vdots \\ & & & b_{nn} \end{bmatrix} \stackrel{\Delta}{=} B$$

$$B^{\mathrm{H}}B = P^{\mathrm{H}}A^{\mathrm{H}}AP = P^{\mathrm{H}}AA^{\mathrm{H}}P = BB^{\mathrm{H}}$$

$$\begin{bmatrix} \bar b_{11} & & & \\ \bar b_{12} & \bar b_{22} & & \\ \vdots & \vdots & \ddots & \\ \bar b_{1n} & \bar b_{2n} & \cdots & \bar b_{nn} \end{bmatrix}\begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ & b_{22} & \cdots & b_{2n} \\ & & \ddots & \vdots \\ & & & b_{nn} \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ & b_{22} & \cdots & b_{2n} \\ & & \ddots & \vdots \\ & & & b_{nn} \end{bmatrix}\begin{bmatrix} \bar b_{11} & & & \\ \bar b_{12} & \bar b_{22} & & \\ \vdots & \vdots & \ddots & \\ \bar b_{1n} & \bar b_{2n} & \cdots & \bar b_{nn} \end{bmatrix}$$

比较第一行第一列元素可得

$$|b_{11}|^2 = |b_{11}|^2 + |b_{12}|^2 + \cdots + |b_{1n}|^2 \Rightarrow b_{12} = 0,\ b_{13} = 0, \cdots, b_{1n} = 0$$

一般地，有

$$\left.\begin{array}{ll} i=1: & b_{12}=0, b_{13}=0,\cdots,b_{1n}=0 \\ i=2: & \qquad b_{23}=0,\cdots,b_{2n}=0 \\ \cdots\cdots & \\ i=n-1: & \qquad\qquad b_{n-1,n}=0 \end{array}\right\} \Rightarrow B=\begin{bmatrix} b_{11} & & & \\ & b_{22} & & \\ & & \ddots & \\ & & & b_{nn} \end{bmatrix}$$

(2) 利用 th41(2)可得.

例如，对上述 B，可求得酉矩阵 $P=\dfrac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$，使得

$$P^{\mathrm{H}}BP=\begin{bmatrix} 6+10\mathrm{j} & \\ & 4-2\mathrm{j} \end{bmatrix}$$

推论 1　$A_{n\times n}$ 实对称 $\Rightarrow \exists$ 正交矩阵 $Q_{n\times n}$，使得 $Q^{\mathrm{T}}AQ=\Lambda$.

推论 2　欧氏空间 V^n，对称变换 $T\Rightarrow$

$\exists$ 标准正交基 $y_1,\cdots,y_n$，使得 $T(y_1,\cdots,y_n)=(y_1,\cdots,y_n)\Lambda$.

证　设 V^n 的标准正交基为 $x_1,\cdots,x_n$，T 在该基下的矩阵为 A，则

A 是实对称矩阵 $\Rightarrow \exists$ 正交矩阵 $Q_{n\times n}$ 使得 $Q^{\mathrm{T}}AQ=\Lambda$

构造 V^n 的标准正交基：$(y_1,\cdots,y_n)=(x_1,\cdots,x_n)Q$　（为什么？）

则 $T(y_1,\cdots,y_n)=T(x_1,\cdots,x_n)Q=(x_1,\cdots,x_n)AQ$

$$=(y_1,\cdots,y_n)Q^{-1}AQ=(y_1,\cdots,y_n)\Lambda$$

推论 3　A 实对称 $\Rightarrow A$ 有 n 个线性无关的特征向量（$\mathbf{R}^n$ 中）.

推论 4　T 是欧氏空间 V^n 的对称变换 $\Rightarrow$

T 有 n 个线性无关的特征向量（V^n 中）

谱分解　$A_{n\times n}$ 是 Hermite 矩阵 $\overset{\text{Th42}}{\Rightarrow} \exists$ 酉矩阵 $P_{n\times n}$，使得 $A=P\Lambda P^{\mathrm{H}}$

划分 $P\overset{\Delta}{=}(p_1,\cdots,p_n)$，则有

$$A=(p_1,\cdots,p_n)\begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}P^{\mathrm{H}}=(\lambda_1 p_1,\cdots,\lambda_n p_n)\begin{bmatrix} p_1^{\mathrm{H}} \\ \vdots \\ p_n^{\mathrm{H}} \end{bmatrix}$$

$$=\lambda_1\left(\boldsymbol{p}_1\boldsymbol{p}_1^{\mathrm{H}}\right)+\cdots+\lambda_n\left(\boldsymbol{p}_n\boldsymbol{p}_n^{\mathrm{H}}\right)$$

① 矩阵组 $\boldsymbol{B}_1=\boldsymbol{p}_1\boldsymbol{p}_1^{\mathrm{H}},\cdots,\boldsymbol{B}_n=\boldsymbol{p}_n\boldsymbol{p}_n^{\mathrm{H}}$ 线性无关；

② $\mathbf{rank}\,\boldsymbol{B}_j=1,\ j=1,2,\cdots,n$.

典型题解析

例 1　设多项式空间 $P[t]_3$ 的两个基为

(Ⅰ)　$f_1(t)=1$，　$f_2(t)=1+t$，　$f_3(t)=1+t+t^2$，　$f_4(t)=1+t+t^2+t^3$

(Ⅱ)　$g_1(t)=1+t^2+t^3$，　$g_2(t)=t+t^2+t^3$，　$g_3(t)=1+t+t^2$，　$g_4(t)=1+t+t^3$

(1) 求由基(Ⅰ)改变为基(Ⅱ)的过渡矩阵；

(2) 求 $P[t]_3$ 中在基(Ⅰ)和基(Ⅱ)下有相同坐标的全体多项式.

解　**(1)** 采用中介基法求过渡矩阵：取 $P[t]_3$ 的简单基为 $1, t, t^2, t^3$，写出由简单基改变为基(Ⅰ)和基(Ⅱ) 过渡矩阵

$$C_1=\begin{bmatrix}1&1&1&1\\ &1&1&1\\ & &1&1\\ & & &1\end{bmatrix},\quad C_2=\begin{bmatrix}1&0&1&1\\0&1&1&1\\1&1&1&0\\1&1&0&1\end{bmatrix}$$

由基(Ⅰ)改变为基(Ⅱ)的过渡矩阵为

$$C=C_1^{-1}C_2=\begin{bmatrix}1&-1&0&0\\-1&0&0&1\\0&0&1&-1\\1&1&0&1\end{bmatrix}$$

(2) 设 $f(t)\in P[t]_3$ 在基(Ⅰ)和基(Ⅱ)下的坐标分别为

$$\alpha=(\xi_1,\xi_2,\xi_3,\xi_4)^{\mathrm{T}},\quad \beta=(\eta_1,\eta_2,\eta_3,\eta_4)^{\mathrm{T}}$$

由坐标变换公式 $\alpha=C\beta$ 及题设 $\alpha=\beta$ 可得　$(I-C)\alpha=0$

该齐次线性方程组的通解为　$\alpha=k(0,0,1,0)^{\mathrm{T}}\ (\forall k\in\mathbf{R})$

在基(Ⅰ)和基(Ⅱ)下有相同坐标的全体多项式

$$\begin{aligned}f(t)&=(f_1(t),f_2(t),f_3(t),f_4(t))\alpha\\&=k\,f_3(t)=k+kt+kt^2\ (\forall k\in\mathbf{R})\end{aligned}$$

例 2　设线性空间 V^3 的线性变换 T 在基 $\alpha_1,\alpha_2,\alpha_3$ 下的矩阵为 $A=\begin{bmatrix}1&2&2\\2&1&2\\2&2&1\end{bmatrix}$，

证明：V^3 的子空间 $W=L(\alpha_2-\alpha_1,\alpha_3-\alpha_1)$ 是 T 的不变子空间.

证法 1　由 $T(\alpha_1,\alpha_2,\alpha_3)=(\alpha_1,\alpha_2,\alpha_3)A$ 知

$$T(\alpha_1)=\alpha_1+2\alpha_2+2\alpha_3，\quad T(\alpha_2)=2\alpha_1+\alpha_2+2\alpha_3$$

$$T(\alpha_3)=2\alpha_1+2\alpha_2+\alpha_3$$

对任意的 $\alpha\in W$，存在常数 k_1,k_2，使得

$$\alpha=k_1(\alpha_2-\alpha_1)+k_2(\alpha_3-\alpha_1)=-(k_1+k_2)\alpha_1+k_1\alpha_2+k_2\alpha_3$$

$$T(\alpha)=-(k_1+k_2)T(\alpha_1)+k_1T(\alpha_2)+k_2T(\alpha_3)$$

$$=-k_1(\alpha_2-\alpha_1)-k_2(\alpha_3-\alpha_1)\in W$$

故 W 是 T 的不变子空间.

证法 2　对任意的 $\alpha\in W$，存在常数 k_1,k_2，使得

$$\alpha=k_1(\alpha_2-\alpha_1)+k_2(\alpha_3-\alpha_1)=(\alpha_1,\alpha_2,\alpha_3)\begin{bmatrix}-k_1-k_2\\k_1\\k_2\end{bmatrix}$$

$$T(\alpha)=(\alpha_1,\alpha_2,\alpha_3)A\begin{bmatrix}-k_1-k_2\\k_1\\k_2\end{bmatrix}=(\alpha_1,\alpha_2,\alpha_3)\begin{bmatrix}k_1+k_2\\-k_1\\-k_2\end{bmatrix}=-\alpha\in W$$

故 W 是 T 的不变子空间.

例 3　设矩阵空间 $\mathbf{R}^{2\times2}$ 的子空间为

$$V=\left\{X=\left(x_{ij}\right)_{2\times2}\,\middle|\,x_{11}+x_{12}+x_{21}=0,\ x_{ij}\in\mathbf{R}\right\}$$

V 中的线性变换为 $T(X)=X+X^{\mathrm{T}}\ \left(\forall\,X\in V\right)$，求 V 的一个基，使 T 在该基下的矩阵为对角矩阵.

解　**(1)** 先求 V 的简单基

$$X \in V \Rightarrow X = \begin{bmatrix} -x_{12}-x_{21} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} = \begin{bmatrix} -x_{12} & x_{12} \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} -x_{21} & 0 \\ x_{21} & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & x_{22} \end{bmatrix}$$

$$= x_{12}\begin{bmatrix} -1 & 1 \\ 0 & 0 \end{bmatrix} + x_{21}\begin{bmatrix} -1 & 0 \\ 1 & 0 \end{bmatrix} + x_{22}\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

容易验证 $X_1 = \begin{bmatrix} -1 & 1 \\ 0 & 0 \end{bmatrix}$，$X_2 = \begin{bmatrix} -1 & 0 \\ 1 & 0 \end{bmatrix}$，$X_3 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ 线性无关，故

X_1, X_2, X_3 是 V 的简单基.

(2) 由公式计算

$$T(X_1) = X_1 + X_1^{\mathrm{T}} = \begin{bmatrix} -2 & 1 \\ 1 & 0 \end{bmatrix} = 1X_1 + 1X_2 + 0X_3$$

$$T(X_2) = X_2 + X_2^{\mathrm{T}} = \begin{bmatrix} -2 & 1 \\ 1 & 0 \end{bmatrix} = 1X_1 + 1X_2 + 0X_3$$

$$T(X_3) = X_3 + X_3^{\mathrm{T}} = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix} = 0X_1 + 0X_2 + 2X_3$$

T 在基 X_1, X_2, X_3 下的矩阵为 $A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$

求 P 使得 $P^{-1}AP = \Lambda$： $\Lambda = \begin{bmatrix} 0 & & \\ & 2 & \\ & & 2 \end{bmatrix}$， $P = \begin{bmatrix} 1 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

由 $(Y_1, Y_2, Y_3) = (X_1, X_2, X_3)P$ 求得

$$Y_1 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, Y_2 = \begin{bmatrix} -2 & 1 \\ 1 & 0 \end{bmatrix}, Y_3 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

故 T 在基 Y_1, Y_2, Y_3 下的矩阵为 Λ .

例 4　设欧氏空间 $\mathbf{R}^{2\times 2}$ 的内积定义为

$$(A, B) = \sum_{i=1}^{2}\sum_{j=1}^{2} a_{ij}b_{ij} \quad \left(A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix},\ B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \right)$$

选取 $A_1 = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$， $A_2 = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$，构造子空间 $W = L(A_1, A_2)$.

(1) 求 $W^{\perp}$ 的一个基；

(2) 利用已知的子空间 W 和 $W^{\perp}$ 的基，求 $\mathbf{R}^{2\times 2}$ 的一个标准正交基.

解 **(1)** 设 $X=\begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix}\in W^{\perp}$，则有 $A_1\perp X, A_2\perp X$，即

$$\begin{cases} x_1+x_2=0 \\ x_2+x_3+x_4=0 \end{cases}$$

该齐次线性方程组的基础解系为（凑正交——为(2)作准备）

$$\xi_1=(1,-1,1,0)^{\mathrm{T}},\quad \xi_2=(1,-1,-2,3)^{\mathrm{T}}$$

故 $W^{\perp}$ 的一个基为 $A_3=\begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}$，$A_4=\begin{bmatrix} 1 & -1 \\ -2 & 3 \end{bmatrix}$.

(2) 对 W 的基正交化：易见 A_1,A_2 线性无关，正交化可得

$$B_1=A_1$$

$$B_2=A_2-\frac{(A_2,B_1)}{(B_1,B_1)}B_1=A_2-\frac{1}{2}B_1=\begin{bmatrix} -1/2 & 1/2 \\ 1 & 1 \end{bmatrix}$$

对 B_1,B_2 单位化可得 W 的标准正交基：

$$C_1=\frac{1}{\sqrt{(B_1,B_1)}}B_1=\frac{1}{\sqrt{2}}B_1=\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$$

$$C_2=\frac{1}{\sqrt{(B_2,B_2)}}B_2=\frac{1}{\sqrt{5/2}}B_2=\frac{1}{\sqrt{10}}\begin{bmatrix} -1 & 1 \\ 2 & 2 \end{bmatrix}$$

$W^{\perp}$ 的正交基为 A_3,A_4，单位化可得 $W^{\perp}$ 的标准正交基：

$$C_3=\frac{1}{\sqrt{(A_3,A_3)}}A_3=\frac{1}{\sqrt{3}}A_3=\frac{1}{\sqrt{3}}\begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}$$

$$C_4=\frac{1}{\sqrt{(A_4,A_4)}}A_4=\frac{1}{\sqrt{15}}A_4=\frac{1}{\sqrt{15}}\begin{bmatrix} 1 & -1 \\ -2 & 3 \end{bmatrix}$$

于是，$\mathbf{R}^{2\times 2}=W\oplus W^{\perp}$ 的标准正交基为 C_1,C_2,C_3,C_4.

例 5　给定欧氏空间V^n的标准正交基$x_1,x_2,\cdots,x_n$，设T是V^n的正交变换，$W=L(x_1,x_2,\cdots,x_r)$是T的不变子空间，证明：V^n的子空间$W^{\perp}=\{y\,|\,y\in V^n,\ y\perp W\}$也是$T$的不变子空间.

证　因为T是V^n的正交变换，而$x_1,x_2,\cdots,x_n$是V^n的标准正交基

所以$T(x_1),T(x_2),\cdots,T(x_n)$也是$V^n$的标准正交基（定理 37 之证明）

$W=L(x_1,x_2,\cdots,x_r)\ \Rightarrow\ W^{\perp}=L(x_{r+1},\cdots,x_n)$

W 是T的不变子空间 $\Rightarrow$ $T(x_1),\cdots,T(x_r)$是W 的标准正交基

$T(x_{r+1}),\cdots,T(x_n)\perp W\Rightarrow\ T(x_{r+1}),\cdots,T(x_n)\in W^{\perp}$

$\forall\, x\in W^{\perp}\ \Rightarrow x=k_{r+1}x_{r+1}+\cdots+k_nx_n$

$\Rightarrow T(x)=k_{r+1}T(x_{r+1})+\cdots+k_nT(x_n)\in W^{\perp}$

故$W^{\perp}$是T的不变子空间.

例 6　设线性空间V中的线性变换T满足$T^2=T$，$R(T)$表示T的值域, $N(T)$表示T的核, T_e表示V中的单位变换，证明：$N(T)=R(T_e-T)$.

证　先证$N(T)\subset R(T_e-T)$：$\forall\alpha\in N(T)$, 有

$T(\alpha)=0\Rightarrow\alpha=T_e(\alpha)=T_e(\alpha)-T(\alpha)=(T_e-T)(\alpha)\in R(T_e-T)$

再证$R(T_e-T)\subset N(T)$：$\forall\alpha\in R(T_e-T)$, 存在$\beta\in V$, 使得

$\alpha=(T_e-T)(\beta)=\beta-T(\beta)$

$T(\alpha)=T(\beta)-T[T(\beta)]=T(\beta)-T^2(\beta)=T(\beta)-T(\beta)=0$

故$\alpha\in N(T)$.

第二章　范数理论及其应用

§2.1　向量范数

一、$\mathbf{C}^n$ 中向量序列的收敛性

设 $\boldsymbol{x}^{(k)}=\left(\xi_1^{(k)},\cdots,\xi_n^{(k)}\right)$，若 $\lim\limits_{k\to\infty}\xi_i^{(k)}=\xi_i\ (i=1,2,\cdots,n)$，称 $\{\boldsymbol{x}^{(k)}\}$收敛于 $\boldsymbol{x}=(\xi_1,\cdots,\xi_n)$，记作 $\lim\limits_{k\to\infty}\boldsymbol{x}^{(k)}=\boldsymbol{x}$ 或 $\boldsymbol{x}^{(k)}\to\boldsymbol{x}\ (k\to\infty)$.

[注] 判断一个向量序列收敛等价于判断 n 个数列同时收敛.

用模刻划：$\mathbf{C}^n$ 中向量 $\boldsymbol{x}$ 的模为 $|\boldsymbol{x}|=\left(|\xi_1|^2+\cdots+|\xi_n|^2\right)^{\frac{1}{2}}$

$$\lim_{k\to\infty}\boldsymbol{x}^{(k)}=\boldsymbol{x}\Leftrightarrow\lim_{k\to\infty}\xi_i^{(k)}=\xi_i\ (i=1,2,\cdots,n)\Leftrightarrow\lim_{k\to\infty}\left|\xi_i^{(k)}-\xi_i\right|^2=0\ (i=1,2,\cdots,n)$$

$$\Leftrightarrow\lim_{k\to\infty}\left(\left|\xi_1^{(k)}-\xi_1\right|^2+\cdots+\left|\xi_n^{(k)}-\xi_n\right|^2\right)^{1/2}=0\Leftrightarrow\lim_{k\to\infty}\left|\boldsymbol{x}^{(k)}-\boldsymbol{x}\right|=0$$

二、线性空间 V 的向量范数

线性空间 V，数域 K，$\forall \boldsymbol{x}\in V$，定义实数$\|\boldsymbol{x}\|$，且满足

(1) $\|\boldsymbol{x}\|\ge 0;\ \|\boldsymbol{x}\|=0\Leftrightarrow \boldsymbol{x}=\theta$

(2) $\|k\boldsymbol{x}\|=|k|\|\boldsymbol{x}\|,\ \forall k\in K$

(3) $\|\boldsymbol{x}+\boldsymbol{y}\|\le\|\boldsymbol{x}\|+\|\boldsymbol{y}\|,\ \forall \boldsymbol{y}\in V$

称$\|\boldsymbol{x}\|$为向量 $\boldsymbol{x}$ 的范数.

例 1 欧氏空间V中，$\|\boldsymbol{x}\|\overset{\Delta}{=}|\boldsymbol{x}|=\sqrt{(\boldsymbol{x},\boldsymbol{x})}$ 是一种向量范数.

例 2 线性空间$\mathbf{C}^n$中，$\|\boldsymbol{x}\|_p\overset{\Delta}{=}\left(\sum\limits_{i=1}^{n}|\xi_i|^p\right)^{\frac{1}{p}}\ (1\le p<\infty)$是向量范数.

证 1° 略．2° 略．3° 设 $\boldsymbol{x}=(\xi_1,\cdots,\xi_n),\ \boldsymbol{y}=(\eta_1,\cdots,\eta_n)$，则

$p=1$：$\|\boldsymbol{x}+\boldsymbol{y}\|_1=\sum|\xi_i+\eta_i|\le\sum(|\xi_i|+|\eta_i|)=\|\boldsymbol{x}\|_1+\|\boldsymbol{y}\|_1$；

$p>1$：$\boldsymbol{x}+\boldsymbol{y}=\theta$时，结论成立；$\boldsymbol{x}+\boldsymbol{y}\ne\theta$时，应用 Hölder 不等式

$$\sum|a_ib_i|\le\left(\sum|a_i|^p\right)^{\frac{1}{p}}\left(\sum|b_i|^q\right)^{\frac{1}{q}}\qquad (\,p>1,q>1,\frac{1}{p}+\frac{1}{q}=1\,)$$

可得（利用 $(p-1)q=p$ ）

$$\left(\|x+y\|_p\right)^p=\sum|\xi_i+\eta_i|^p=\sum\left(|\xi_i+\eta_i||\xi_i+\eta_i|^{p-1}\right)$$

$$\le\sum|\xi_i||\xi_i+\eta_i|^{p-1}+\sum|\eta_i||\xi_i+\eta_i|^{p-1}$$

$$\le\left(\sum|\xi_i|^p\right)^{\frac{1}{p}}\left(\sum\left(|\xi_i+\eta_i|^{p-1}\right)^q\right)^{\frac{1}{q}}+\left(\sum|\eta_i|^p\right)^{\frac{1}{p}}\left(\sum\left(|\xi_i+\eta_i|^{p-1}\right)^q\right)^{\frac{1}{q}}$$

$$=\left(\|x\|_p+\|y\|_p\right)\left(\|x+y\|_p\right)^{p-1}\qquad\left(\frac{1}{q}=\frac{1}{p}\cdot(p-1)\right)$$

故 $\|x+y\|_p\le\|x\|_p+\|y\|_p$

因此，当 $1\le p<+\infty$ 时，$\|x\|_p$ 是向量 x 的范数.

特例：1－范数　$\|x\|_1=\sum|\xi_i|$

2－范数　$\|x\|_2=\left(\sum|\xi_i|^2\right)^{\frac{1}{2}}$

∞ －范数　$\|x\|_\infty\overset{\Delta}{=}\lim\limits_{p\to+\infty}\|x\|_p=\max\limits_i|\xi_i|$

证明上述极限式：$x=\theta$ 时等式成立. $x\ne\theta$ 时，设 $|\xi_{i_0}|=\max\limits_i|\xi_i|$，则有

$$\|x\|_p=|\xi_{i_0}|\left(\sum\left|\frac{\xi_i}{\xi_{i_0}}\right|^p\right)^{\frac{1}{p}},\quad 1\le\left(\sum\left|\frac{\xi_i}{\xi_{i_0}}\right|^p\right)^{\frac{1}{p}}\le n^{\frac{1}{p}}\to 1$$

$$\lim_{p\to+\infty}\|x\|_p=|\xi_{i_0}|=\max_i|\xi_i|$$

容易验证 $\|x\|_\infty=\max\limits_i|\xi_i|$ 满足向量范数的三个条件，从而是向量范数.

例 3　线性空间 V^n 中，给定基 $x_1,\cdots,x_n$，因为

$$x=\xi_1x_1+\cdots+\xi_nx_n\overset{1-1}{\leftrightarrow}\alpha=(\xi_1,\cdots,\xi_n)^{\mathrm{T}}\in\mathbf{C}^n$$

所以 $\|x\|_p\overset{\Delta}{=}\|\alpha\|_p$ 是 V^n 中元素 x 的 p－范数.

例 4　设 $A\in\mathbf{R}^{n\times n}$ 对称正定，线性空间 $\mathbf{R}^n$ 中，$\|\alpha\|_A\overset{\Delta}{=}\left(\alpha^{\mathrm{T}}A\alpha\right)^{\frac{1}{2}}$ 是向量范数.

证　1° 成立．2° 成立．

3° A 实对称 $\Rightarrow \exists$ 正交矩阵 Q 使得 $Q^{\mathrm{T}}AQ=\begin{bmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{bmatrix}$

A 正定 $\Rightarrow \lambda_i>0$

$$\Rightarrow A=Q\begin{bmatrix}\sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n}\end{bmatrix}\begin{bmatrix}\sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n}\end{bmatrix}Q^{\mathrm{T}}\overset{\Delta}{=}B^{\mathrm{T}}B$$

因为 $\|\alpha\|_A=\left(\alpha^{\mathrm{T}}B^{\mathrm{T}}B\alpha\right)^{\frac{1}{2}}=\left[(B\alpha)^{\mathrm{T}}(B\alpha)\right]^{\frac{1}{2}}=\|B\alpha\|_2$　（$\mathbf{R}^n$ 中）

所以 $\|\alpha+\beta\|_A=\|B(\alpha+\beta)\|_2=\|(B\alpha)+(B\beta)\|_2$

$$\le\|B\alpha\|_2+\|B\beta\|_2=\|\alpha\|_A+\|\beta\|_A$$

三、范数等价

对于线性空间 V^n 的向量范数 $\|x\|_\alpha$ 与 $\|x\|_\beta$，若有正常数 c_1 和 c_2，使得

$$c_1\|x\|_\beta\le\|x\|_\alpha\le c_2\|x\|_\beta\qquad(\forall x\in V^n)$$

成立，称 $\|x\|_\alpha$ 与 $\|x\|_\beta$ 等价．

(1) 自反性：$1\cdot\|x\|_\alpha\le\|x\|_\alpha\le1\cdot\|x\|_\alpha,\forall x\in V^n$

(2) 对称性：$\dfrac{1}{c_2}\|x\|_\alpha\le\|x\|_\beta\le\dfrac{1}{c_1}\|x\|_\alpha,\forall x\in V^n$

(3) 传递性：$\left.\begin{aligned}c_1\|x\|_\beta\le\|x\|_\alpha\le c_2\|x\|_\beta\\ c_1'\|x\|_\gamma\le\|x\|_\beta\le c_2'\|x\|_\gamma\end{aligned}\right\}\Rightarrow c_1''\|x\|_\gamma\le\|x\|_\alpha\le c_2''\|x\|_\gamma,\forall x\in V^n$

例 5　向量空间 $\mathbf{C}^n$ 中，对任意向量 $x=(\xi_1,\cdots,\xi_n)^{\mathrm{T}}$，有

(1) $\|x\|_1=\sum|\xi_i|\le n\cdot\max\limits_i|\xi_i|=n\|x\|_\infty$

$\|x\|_1\ge\max\limits_i|\xi_i|=1\cdot\|x\|_\infty$，故 $1\cdot\|x\|_\infty\le\|x\|_1\le n\cdot\|x\|_\infty$．

(2) $\|x\|_2=\left(\sum|\xi_i|^2\right)^{\frac{1}{2}}\le\left(n\cdot\max\limits_i|\xi_i|^2\right)^{\frac{1}{2}}=\sqrt{n}\cdot\|x\|_\infty$

$$\|x\|_2 \ge \left(\max_i |\xi_i|^2\right)^{\frac{1}{2}} = 1\cdot\|x\|_\infty，故\ 1\cdot\|x\|_\infty \le \|x\|_2 \le \sqrt{n}\cdot\|x\|_\infty .$$

(3) $\dfrac{1}{\sqrt{n}}\cdot\|x\|_2 \le \|x\|_1 \le n\cdot\|x\|_2$

Th1　线性空间V^n中，任意两种向量范数等价.

证　只需证明任一向量范数$\|x\|_\alpha$与向量范数$\|x\|_2$等价即可.

给定V^n的基$x_1,\cdots,x_n$，对任意$x\in V^n$，有

$$x=\xi_1 x_1+\cdots+\xi_n x_n \text{唯一}$$

$$\|x\|_2=\left(\sum_{i=1}^{n}|\xi_i|^2\right)^{\frac{1}{2}} \qquad \text{（见例 3）}$$

$\mathbf{C}^n$的子集$S \overset{\Delta}{=} \left\{(\xi_1,\cdots,\xi_n)\,\middle|\, |\xi_1|^2+\cdots+|\xi_n|^2=1\right\}$是闭区域，实值函数

$$f(\xi_1,\cdots,\xi_n) \overset{\Delta}{=} \|x\|_\alpha = \|\xi_1 x_1+\cdots+\xi_n x_n\|_\alpha$$

在S上连续，故f在S上取得最值 $\min f=c_1$，$\max f=c_2$.

$$\forall(\xi_1,\cdots,\xi_n)\in S \Rightarrow x=\xi_1 x_1+\cdots+\xi_n x_n \ne \theta$$

$$\Rightarrow f(\xi_1,\cdots,\xi_n)=\|x\|_\alpha>0 \Rightarrow c_1>0$$

对$\forall x\in V^n$，当$x\ne\theta$时，有

$$y=\frac{x}{\|x\|_2}=\sum_{i=1}^{n}\frac{\xi_i}{\|x\|_2}x_i=\sum_{i=1}^{n}\eta_i x_i \quad \left(\eta_i=\frac{\xi_i}{\|x\|_2}\right)$$

$$|\eta_1|^2+\cdots+|\eta_n|^2=1 \Rightarrow (\eta_1,\cdots,\eta_n)\in S$$

故　$0<c_1\le f(\eta_1,\cdots,\eta_n)\le c_2 \Rightarrow\ 0<c_1\le\|y\|_\alpha\le c_2$

$$\Rightarrow\ c_1\|x\|_2\le\|x\|_\alpha\le c_2\|x\|_2$$

当$x=\theta$时，上式显然成立.

Th2　向量空间$\mathbf{C}^n$中，$\lim\limits_{k\to\infty}x^{(k)}=x \Leftrightarrow \forall\|x\|,\ \lim\limits_{k\to\infty}\|x^{(k)}-x\|=0$.

证　只需对$\|x\|=\|x\|_1$证明即可.

$$x^{(k)} \to x \Leftrightarrow \xi_i^{(k)} \to \xi_i \ (i=1,2,\cdots,n) \Leftrightarrow \left|\xi_i^{(k)} - \xi_i\right| \to 0 \ (i=1,2,\cdots,n)$$

$$\Leftrightarrow \sum_{i=1}^{n}\left|\xi_i^{(k)} - \xi_i\right| \to 0 \Leftrightarrow \left\|x^{(k)} - x\right\|_1 \to 0$$

§2.2 矩阵范数

一、矩阵范数

集合$\mathbf{C}^{m\times n}$中，$\forall A \in \mathbf{C}^{m\times n}$，定义实数$\|A\|$，且满足

(1) $\|A\| \ge 0$；$\|A\| = 0 \Leftrightarrow A = O_{m\times n}$

(2) $\|kA\| = |k|\cdot\|A\|, \forall k \in \mathbf{C}$

(3) $\|A+B\| \le \|A\| + \|B\|, \forall B \in \mathbf{C}^{m\times n}$

(4) AB 有意义：$\|AB\| \le \|A\|\cdot\|B\|, \forall B \in \mathbf{C}^{n\times l}$

称$\|A\|$为矩阵 A 的范数.

二、矩阵范数与向量范数相容

设$\mathbf{C}^{m\times n}$的矩阵范数$\|A\|_M$，$\mathbf{C}^m$与$\mathbf{C}^n$中的“同类向量范数”　$\|x\|_V$，若

$$\|Ax\|_V \le \|A\|_M \cdot \|x\|_V \quad (\forall A \in \mathbf{C}^{m\times n}, \forall x \in \mathbf{C}^n)$$

称矩阵范数$\|A\|_M$与向量范数$\|x\|_V$相容.

预备：$\alpha = (|a_1|,\cdots,|a_n|)$，$\beta = (|b_1|,\cdots,|b_n|)$

$$(\alpha,\beta) \le \sqrt{(\alpha,\alpha)}\cdot\sqrt{(\beta,\beta)} \Rightarrow \sum_{i=1}^{n}|a_i||b_i| \le \sqrt{\sum_{i=1}^{n}|a_i|^2}\cdot\sqrt{\sum_{i=1}^{n}|b_i|^2}$$

例 6　$A = (a_{ij})_{m\times n}, x = (\xi_1,\xi_2,\cdots,\xi_n)^{\mathrm{T}}$.

(1) $\|A\|_{m_1} = \sum_{i,j}|a_{ij}|$是矩阵范数，且与$\|x\|_1$相容.

(2) $\|A\|_{m_\infty} = n\cdot\max_{i,j}|a_{ij}|$是矩阵范数，且与$\|x\|_\infty$相容.

(3) $\|A\|_{m_2} = \left(\sum_{i,j}|a_{ij}|^2\right)^{\frac{1}{2}}$是矩阵范数，且与$\|x\|_2$相容.

证 (1) $1^0 \sim 3^0$成立.

$$\|Ax\|_1=\sum_{i=1}^{m}|a_{i1}\xi_1+\cdots+a_{in}\xi_n|\le\sum_{i=1}^{m}\left(|a_{i1}||\xi_1|+\cdots+|a_{in}||\xi_n|\right)$$

$$\le\sum_{i=1}^{m}\left[\left(|a_{i1}|+\cdots+|a_{in}|\right)\cdot\left(|\xi_1|+\cdots+|\xi_n|\right)\right]$$

$$=\left[\sum_{i=1}^{m}\left(|a_{i1}|+\cdots+|a_{in}|\right)\right]\cdot\left(|\xi_1|+\cdots+|\xi_n|\right)=\|A\|_{m_1}\cdot\|x\|_1$$

4^0 划分 $B_{n\times l}=(b_1,\cdots,b_l)$，则 $AB=(Ab_1,\cdots,Ab_l)$，且有

$$\|AB\|_{m_1}=\|Ab_1\|_1+\cdots+\|Ab_l\|_1\le\|A\|_{m_1}\|b_1\|_1+\cdots+\|A\|_{m_1}\|b_l\|_1$$

$$=\|A\|_{m_1}\cdot\left(\|b_1\|_1+\cdots+\|b_l\|_1\right)=\|A\|_{m_1}\cdot\|B\|_{m_1}$$

(2) $1^0 \sim 3^0$成立. 4^0 设 $B=(b_{ij})_{n\times l}$，则 $AB=\left(\sum_{k=1}^{n}a_{ik}b_{kj}\right)_{m\times l}$，且有

$$\|AB\|_{m_\infty}=l\cdot\max_{i,j}\left|\sum_{k=1}^{n}a_{ik}b_{kj}\right|\le l\cdot\max_{i,j}\left(\sum_{k=1}^{n}|a_{ik}||b_{kj}|\right)$$

$$\le\left(n\cdot\max_{i,j}|a_{ij}|\right)\cdot\left(l\cdot\max_{i,j}|b_{ij}|\right)=\|A\|_{m_\infty}\cdot\|B\|_{m_\infty}$$

$$\|Ax\|_\infty=\max_i|a_{i1}\xi_1+\cdots+a_{in}\xi_n|\le\max_i\left(|a_{i1}||\xi_1|+\cdots+|a_{in}||\xi_n|\right)$$

$$\le\max_i\left(|a_{i1}|+\cdots+|a_{in}|\right)\cdot\max_i|\xi_i|\le\left(n\cdot\max_i|a_{ij}|\right)\cdot\max_i|\xi_i|=\|A\|_{m_\infty}\cdot\|x\|_\infty$$

(3) 1^0 成立. 2^0 成立.

3^0 设 $B_{m\times n}$，划分 $A=(a_1,\cdots,a_n)$， $B=(b_1,\cdots,b_n)$，则有

$$\|A+B\|_{m_2}^2=\|a_1+b_1\|_2^2+\cdots+\|a_n+b_n\|_2^2$$

$$\le\left(\|a_1\|_2+\|b_1\|_2\right)^2+\cdots+\left(\|a_n\|_2+\|b_n\|_2\right)^2$$

$$\le\|A\|_{m_2}^2+2\left(\|a_1\|_2\|b_1\|_2+\cdots+\|a_n\|_2\|b_n\|_2\right)+\|B\|_{m_2}^2$$

$$\le\|A\|_{m_2}^2+2\left(\sum\|a_i\|_2^2\right)^{\frac{1}{2}}\cdot\left(\sum\|b_i\|_2^2\right)^{\frac{1}{2}}+\|B\|_{m_2}^2=\left(\|A\|_{m_2}+\|B\|_{m_2}\right)^2$$

4^0　设 $B_{n\times l}$, $AB=\left(\sum_{k=1}^{n}a_{ik}b_{kj}\right)_{m\times l}$，则有

$$\|AB\|_{m_2}^2=\sum_{i,j}\left|\sum_k a_{ik}b_{kj}\right|^2\le\sum_{i,j}\left(\sum_k|a_{ik}|\cdot|b_{kj}|\right)^2$$

$$\le\sum_{i,j}\left[\left(\sum_k|a_{ik}|^2\right)\cdot\left(\sum_k|b_{kj}|^2\right)\right]$$

$$=\sum_i\left\{\left(\sum_k|a_{ik}|^2\right)\cdot\sum_j\left(\sum_k|b_{kj}|^2\right)\right\}$$

$$=\left(\sum_{i,k}|a_{ik}|^2\right)\cdot\left(\sum_{k,j}|b_{kj}|^2\right)=\|A\|_{m_2}^2\cdot\|B\|_{m_2}^2$$

特别的，取 $B=x\in\mathbf{C}^{n\times 1}$，则有

$$\|Ax\|_2=\|AB\|_{m_2}\le\|A\|_{m_2}\cdot\|B\|_{m_2}=\|A\|_{m_2}\cdot\|x\|_2$$

[注] ① $\|A\|_{m_2}=\left[\mathrm{tr}\left(A^{\mathrm{H}}A\right)\right]^{\frac{1}{2}}=\left[\mathrm{tr}\left(AA^{\mathrm{H}}\right)\right]^{\frac{1}{2}}$，记作 $\|A\|_{\mathrm{F}}$.

② $\mathbf{C}^{m\times n}$ 中的矩阵范数等价：对任意的两种矩阵范数 $\|A\|_\alpha$ 与 $\|A\|_\beta$，存在 $0<c_1\le c_2$，使得 $c_1\|A\|_\beta\le\|A\|_\alpha\le c_2\|A\|_\beta\quad(\forall A_{m\times n})$.

③ $\mathbf{C}^{m\times n}$ 中，$\lim\limits_{k\to\infty}A^{(k)}=A\Leftrightarrow\forall\|A\|$, $\lim\limits_{k\to\infty}\left\|A^{(k)}-A\right\|=0$.

Th3 对于 $A_{m\times n}$ 及酉矩阵 $P_{m\times m}$ 和 $Q_{n\times n}$，有

$$\|PA\|_{\mathrm{F}}=\|A\|_{\mathrm{F}},\quad\|AQ\|_{\mathrm{F}}=\|A\|_{\mathrm{F}}$$

证　$\|PA\|_{\mathrm{F}}^2=\mathrm{tr}\left((PA)^{\mathrm{H}}(PA)\right)=\mathrm{tr}\left(A^{\mathrm{H}}P^{\mathrm{H}}PA\right)=\mathrm{tr}\left(A^{\mathrm{H}}A\right)=\|A\|_{\mathrm{F}}^2$

$\|AQ\|_{\mathrm{F}}^2=\mathrm{tr}\left((AQ)(AQ)^{\mathrm{H}}\right)=\mathrm{tr}\left(AQQ^{\mathrm{H}}A^{\mathrm{H}}\right)=\mathrm{tr}\left(AA^{\mathrm{H}}\right)=\|A\|_{\mathrm{F}}^2$

引理　对 $\mathbf{C}^{m\times n}$ 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 $\mathbf{C}^{n\times n}$ 的矩阵范数 $\|A\|$，任取 $0\ne y\in\mathbf{C}^n$ (列向量)，则

(1) $\|x\|_V\overset{\Delta}{=}\|xy^{\mathrm{H}}\|$ 是 $\mathbf{C}^n$ 的向量范数；(2) $\|A\|$ 与 $\|x\|_V$ 相容.

证　(1) 略. (2) $\|Ax\|_V=\left\|(Ax)y^{\mathrm{H}}\right\|=\left\|A\left(xy^{\mathrm{H}}\right)\right\|\le\|A\|\cdot\left\|xy^{\mathrm{H}}\right\|=\|A\|\cdot\|x\|_V$.

练习　给定非零列向量 $\boldsymbol{y}\in\mathbf{C}^m$，对 $\mathbf{C}^{m\times n}$ 的矩阵范数 $\|A\|_{\mathrm{F}}$，定义 $\|\boldsymbol{x}\|_V=\|\boldsymbol{y}\boldsymbol{x}^{\mathrm{T}}\|_{\mathrm{F}}$（列向量 $\boldsymbol{x}\in\mathbf{C}^n$），则 $\|\boldsymbol{x}\|_V$ 是 $\mathbf{C}^n$ 的向量范数，且有 $\|A\boldsymbol{x}\|_V\le\|A\|_{\mathrm{F}}\cdot\|\boldsymbol{x}\|_V$.

三、从属范数

定理 4　对 $\mathbf{C}^m$ 与 $\mathbf{C}^n$ 中的“同类向量范数” $\|\boldsymbol{x}\|_V$，定义实数

$$\|A\|=\max_{\|x\|_V=1}\|A\boldsymbol{x}\|_V\quad\left(\forall A_{m\times n},\ \boldsymbol{x}\in\mathbf{C}^n\right)$$

则 $\|A\|$ 是 $\mathbf{C}^{m\times n}$ 中矩阵 A 的范数，且 $\|A\|$ 与 $\|\boldsymbol{x}\|_V$ 相容.

[注]　等价定义 $\max\limits_{\|x\|_V=1}\|A\boldsymbol{x}\|_V=\max\limits_{x\ne\theta}\dfrac{\|A\boldsymbol{x}\|_V}{\|\boldsymbol{x}\|_V}$

证　1^0　$A\ne O:\exists\boldsymbol{x}_0$ 满足 $\|\boldsymbol{x}_0\|_V=1$, st $A\boldsymbol{x}_0\ne\theta\Rightarrow\|A\|\ge\|A\boldsymbol{x}_0\|_V>0$

$A=O:\|A\|=\max\limits_{\|x\|_V=1}\|A\boldsymbol{x}\|_V=\max\limits_{\|x\|_V=1}\|0\|_V=0$

2^0　略

3^0　对 $A+B:\exists\boldsymbol{x}_1$ 满足 $\|\boldsymbol{x}_1\|_V=1$, st $\max\limits_{\|x\|_V=1}\|(A+B)\boldsymbol{x}\|_V=\|(A+B)\boldsymbol{x}_1\|_V$.

$\|A+B\|=\|A\boldsymbol{x}_1+B\boldsymbol{x}_1\|_V\le\|A\boldsymbol{x}_1\|_V+\|B\boldsymbol{x}_1\|_V\le\|A\|+\|B\|$

4^0　先证 $\|A\boldsymbol{y}\|_V\le\|A\|\cdot\|\boldsymbol{y}\|_V\quad\left(\boldsymbol{y}\in\mathbf{C}^n\right)$

$\boldsymbol{y}=\theta$：显然成立.

$\boldsymbol{y}\ne\theta$：$\boldsymbol{y}_0\overset{\Delta}{=}\dfrac{\boldsymbol{y}}{\|\boldsymbol{y}\|_v}$ 满足 $\|\boldsymbol{y}_0\|_V=1\Rightarrow\|A\boldsymbol{y}_0\|_V\le\max\limits_{\|x\|_V=1}\|A\boldsymbol{x}\|_V=\|A\|$

故 $\|A\boldsymbol{y}\|_V=\left\|A\left(\|\boldsymbol{y}\|_V\boldsymbol{y}_0\right)\right\|_V=\|A\boldsymbol{y}_0\|_V\cdot\|\boldsymbol{y}\|_V\le\|A\|\cdot\|\boldsymbol{y}\|_V$

对 AB，$\exists\boldsymbol{x}_2$ 满足 $\|\boldsymbol{x}_2\|_V=1$,使得 $\max\limits_{\|x\|_V=1}\|(AB)\boldsymbol{x}\|_V=\|(AB)\boldsymbol{x}_2\|_V$，于是

$$\|AB\|=\|(AB)\boldsymbol{x}_2\|_V=\|A(B\boldsymbol{x}_2)\|_V\le\|A\|\cdot\|B\boldsymbol{x}_2\|_V\le\|A\|\cdot\|B\|$$

故定理成立.

[注] ① 一般的矩阵范数：$I=I\cdot I\Rightarrow\|I\|\le\|I\|\cdot\|I\|\Rightarrow\|I\|\ge1$.

例如 $\|I\|_{m_1}=n$，$\|I\|_{\mathrm{F}}=\sqrt{n}$.

② 矩阵的从属范数：$\|I\|=\max\limits_{\|x\|_V=1}\|Ix\|_V=1$.

③ 常用从属范数：　$\|x\|_V$　$\|x\|_1$　$\|x\|_2$　$\|x\|_\infty$

$\|A\|_M$　$\|A\|_1$　$\|A\|_2$　$\|A\|_\infty$

Th5　设 $A=(a_{ij})_{m\times n}$，则

(1) 列和范数　$\|A\|_1=\max\limits_j\left\{\sum\limits_{i=1}^m|a_{ij}|\right\}$

(2) 谱范数　$\|A\|_2=\sqrt{\lambda_1}$，$\lambda_1=\max\{\lambda(A^{\mathrm H}A)\}$

(3) 行和范数　$\|A\|_\infty=\max\limits_i\left\{\sum\limits_{j=1}^n|a_{ij}|\right\}$

证　(1) 记 $t=\max\limits_j\left(\sum\limits_{i=1}^m|a_{ij}|\right)$. 左≤右：若 $x\in\mathbf{C}^n$ 满足 $\|x\|_1=1$，则

$$\|Ax\|_1=\sum_{i=1}^m\left|\sum_{j=1}^n a_{ij}\xi_j\right|\le\sum_i\sum_j|a_{ij}|\cdot|\xi_j|=\sum_j\left[|\xi_j|\cdot\left(\sum_i|a_{ij}|\right)\right]$$

$$\le\sum_j\left[|\xi_j|t\right]=t\cdot\|x\|_1=\text{右}$$

故 $\|A\|_1=\max\limits_{\|x\|_1=1}\|Ax\|_1\le$ 右.

左≥右：选取 k 使得 $t=\max\limits_j\left(\sum\limits_{i=1}^m|a_{ij}|\right)=\sum\limits_{i=1}^m|a_{ik}|$，令

$e_k=(0,\cdots,0,1,0,\cdots,0)^{\mathrm T}$，则 $\|e_k\|_1=1$，且有

$$\|A\|_1\ge\|Ae_k\|_1=\left\|\begin{bmatrix}a_{1k}\\\vdots\\a_{mk}\end{bmatrix}\right\|_1=\sum_{i=1}^m|a_{ik}|=\text{右}$$

(2) $A^{\mathrm H}A$ 是 Hermite 矩阵 $\Rightarrow\lambda(A^{\mathrm H}A)\in\mathbf{R}$ 且 $\lambda(A^{\mathrm H}A)\ge0$

$A^{\mathrm H}A$ 的特征值：$\lambda_1\ge\lambda_2\ge\cdots\ge\lambda_n$

$A^{\mathrm H}A$ 的特征向量：$x_1,x_2,\cdots,x_n$ 两两正交且满足 $\|x_i\|_2=1$

左≤右：若 $x\in\mathbf{C}^n$ 满足 $\|x\|_2=1$，则 $x=\xi_1x_1+\cdots+\xi_nx_n$，且有

$$|\xi_1|^2+\cdots+|\xi_n|^2=1\text{，}\ (A^{\mathrm H}A)x=\lambda_1\xi_1x_1+\cdots+\lambda_n\xi_nx_n$$

$$\begin{aligned}\|Ax\|_2^2&=(Ax)^{\mathrm H}(Ax)=x^{\mathrm H}(A^{\mathrm H}A)x\\&=(\bar\xi_1x_1^{\mathrm H}+\cdots+\bar\xi_nx_n^{\mathrm H})(\lambda_1\xi_1x_1+\cdots+\lambda_n\xi_nx_n)\\&=\lambda_1|\xi_1|^2+\cdots+\lambda_n|\xi_n|^2\le\lambda_1\end{aligned}$$

故 $\|A\|_2=\max\limits_{\|x\|_2=1}\|Ax\|_2\le\sqrt{\lambda_1}$

左≥右：$\|Ax_1\|_2^2=(Ax_1)^{\mathrm H}(Ax_1)=x_1^{\mathrm H}(A^{\mathrm H}A)x_1=\lambda_1x_1^{\mathrm H}x_1=\lambda_1$

故 $\|A\|_2\ge\|Ax_1\|_2=\sqrt{\lambda_1}$

(3) 记 $t=\max\limits_i\left(\sum\limits_{j=1}^n|a_{ij}|\right)$．左≤右：若 $x\in\mathbf{C}^n$ 满足 $\|x\|_\infty=1$，则

$$\|Ax\|_\infty=\max_i\left|\sum_{j=1}^na_{ij}\xi_j\right|\le\max_i\sum_j|a_{ij}|\cdot|\xi_j|$$

$$\le\left(\max_i\sum_j|a_{ij}|\right)\cdot\max_j|\xi_j|=t\,\|x\|_\infty=t$$

$$\|A\|_\infty=\max_{\|x\|_\infty=1}\|Ax\|_\infty\le\text{右}$$

左≥右：选取 k 使得 $t=\max\limits_i\left(\sum\limits_{j=1}^n|a_{ij}|\right)=\sum\limits_{j=1}^n|a_{kj}|$，令

$$y_0=\begin{bmatrix}\eta_1\\ \vdots\\ \eta_n\end{bmatrix},\ \eta_j=\begin{cases}1 & (a_{kj}=0)\\ \dfrac{|a_{kj}|}{a_{kj}} & (a_{kj}\ne0)\end{cases}\Rightarrow|\eta_j|=1$$

则 $\|y_0\|_\infty=1$，$Ay_0=\begin{bmatrix}\sum\limits_{j=1}^na_{1j}\eta_j\\ \cdots\\ \sum\limits_{j=1}^n|a_{kj}|\\ \cdots\\ \sum\limits_{j=1}^na_{mj}\eta_j\end{bmatrix}\Rightarrow\|Ay_0\|_\infty=t$

由此可得 $\|A\|_\infty=\max\limits_{\|x\|_\infty=1}\|Ax\|_\infty\ge\|Ay_0\|_\infty=t=\text{右}$

§2.3 范数的应用

Th6　$\|A_{n\times n}\|<1 \Rightarrow (I-A)$可逆, 且$\left\|(I-A)^{-1}\right\| \le \dfrac{\|I\|}{1-\|A\|}$.

证　选取向量范数$\|x\|_V$，使得$\|A\|$与$\|x\|_V$相容(例 7). 若 det$(I-A)=0$,

则$(I-A)x=0$有非零解x_0, 即$(I-A)x_0=0$. 于是有

$$x_0 = Ax_0 \Rightarrow \|x_0\|_V = \|Ax_0\|_V \le \|A\|\cdot\|x_0\|_V < \|x_0\|_V$$

产生矛盾，故$(I-A)$可逆.

$$(I-A)^{-1}(I-A)=I \Rightarrow (I-A)^{-1} = I+(I-A)^{-1}A$$

$$\Rightarrow \left\|(I-A)^{-1}\right\| \le \|I\| + \left\|(I-A)^{-1}\right\|\cdot\|A\|$$

$$\Rightarrow \left\|(I-A)^{-1}\right\| \le \frac{\|I\|}{1-\|A\|}$$

Th7　$\|A_{n\times n}\|<1 \Rightarrow \left\|I-(I-A)^{-1}\right\| \le \dfrac{\|A\|}{1-\|A\|}$.

证　恒等式　　$(I-A)-I=-A$

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

左乘A:　　$A-A(I-A)^{-1} = -A^2(I-A)^{-1}$

$$\Rightarrow A(I-A)^{-1} = A + A\cdot A(I-A)^{-1}$$

$$\Rightarrow \left\|A(I-A)^{-1}\right\| \le \|A\| + \|A\|\cdot\left\|A(I-A)^{-1}\right\|$$

$$\Rightarrow \left\|A(I-A)^{-1}\right\| \le \frac{\|A\|}{1-\|A\|}$$

$$\Rightarrow \left\|I-(I-A)^{-1}\right\| = \left\|-A(I-A)^{-1}\right\| \le \frac{\|A\|}{1-\|A\|}$$

Th8　设$A_{n\times n}$可逆，$B_{n\times n}$，且满足$\left\|A^{-1}B\right\|<1$，则

(1) $(A+B)$可逆；

(2) $F \overset{\Delta}{=} I-\left(I+A^{-1}B\right)^{-1}$：$\|F\| \le \dfrac{\|A^{-1}B\|}{1-\|A^{-1}B\|}$；

(3) $\dfrac{\left\|A^{-1}-(A+B)^{-1}\right\|}{\|A^{-1}\|} \le \dfrac{\|A^{-1}B\|}{1-\|A^{-1}B\|}$.

证　利用定理 6 和定理 7 可得.

谱半径：$A_{n\times n}$，$\rho(A) \overset{\Delta}{=} \max\limits_{i}|\lambda_i(A)|$.

Th9　对 $\forall\, A_{n\times n}, \forall\, \|\bullet\|_M$，有 $\rho(A) \le \|A\|_M$.

证　对矩阵范数 $\|\bullet\|_M$，存在向量范数 $\|\bullet\|_V$，使得 $\|Ax\|_V \le \|A\|_M \cdot \|x\|_V$.

设 $Ax_i=\lambda_i x_i\ (x_i \ne \theta)$，则有

$$|\lambda_i|\cdot\|x_i\|_V=\|\lambda_i x_i\|_V=\|Ax_i\|_V \le \|A\|_M\cdot\|x_i\|_V$$

$$|\lambda_i| \le \|A\|_M \Rightarrow \rho(A) \le \|A\|_M$$

Th10　给定 $A_{n\times n}$，对 $\forall \varepsilon>0$, $\exists$ 矩阵范数 $\|\cdot\|_M$, st $\|A\|_M \le \rho(A)+\varepsilon$.

证　根据矩阵的 Jordan 标准形理论：对于矩阵 A，存在可逆矩阵 $P_{n\times n}$，

使得 $P^{-1}AP=J$.　记

$$\Lambda=\begin{bmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{bmatrix},\quad \tilde{I}=\begin{bmatrix}0 & \delta_1 & & & \\ & 0 & \delta_2 & & \\ & & \ddots & \ddots & \\ & & & 0 & \delta_{n-1} \\ & & & & 0\end{bmatrix}$$

则　$J=\Lambda+\tilde{I}\quad (\delta_i=1 \text{ 或 } 0)$，于是有

$$D=\begin{bmatrix}1 & & & \\ & \varepsilon & & \\ & & \ddots & \\ & & & \varepsilon^{n-1}\end{bmatrix},\quad (PD)^{-1}A(PD)=D^{-1}JD=\Lambda+\varepsilon\tilde{I}$$

$S \overset{\Delta}{=} PD$ 可逆：$\left\|S^{-1}AS\right\|_1 = \left\|\Lambda + \varepsilon\,\tilde{I}\right\|_1 \le \rho(A) + \varepsilon$

可证 $\|B\|_M = \left\|S^{-1}BS\right\|_1$ $(\forall B \in \mathbf{C}^{n\times n})$ 是 $\mathbf{C}^{n\times n}$ 中的矩阵范数，于是有

$\|A\|_M = \left\|S^{-1}AS\right\|_1 \le \rho(A) + \varepsilon$.

〔注〕因为 $\|\bullet\|_M$ 与给定的矩阵 A 有关，所以当 $B_{n\times n} \ne A$ 时，针对 A 构造的矩阵范数 $\|\bullet\|_M$，不等式 $\|B\|_M \le \rho(B) + \varepsilon$ 不一定成立！

讨论：① $\|A\|_{M_1}, \|A\|_{M_2}$ 可与同一种 $\|x\|_V$ 相容 ？

② $\|A\|_M$ 可与不同的 $\|x\|_{V_1}, \|x\|_{V_2}$ 相容？

③ $\forall\|A\|_M$ 与 $\forall\|x\|_V$ 不一定相容？

分析：① $\|A\|_{m_1}, \|A\|_1$ 与 $\|x\|_1$ 相容.

② $\|A\|_{m_1}$ 与 $\|x\|_p$ $(p \ge 1)$ 相容.　（例如 $p=1,\ p=2$）

$$x = (\xi_1, \cdots, \xi_n)^{\mathrm{T}},\quad E_{ij}x = (0, \cdots, 0, \xi_j, 0, \cdots, 0)^{\mathrm{T}} \Rightarrow \|E_{ij}x\|_p \le \|x\|_p$$

$$Ax = \sum_{i,j} a_{ij}E_{ij}x$$

$$\|Ax\|_p \le \sum_{i,j}|a_{ij}| \cdot \|E_{ij}x\|_p \le \left(\sum_{i,j}|a_{ij}|\right) \cdot \|x\|_p = \|A\|_{m_1} \cdot \|x\|_p$$

③ $\|A\|_1 = \max\limits_j\left(\sum\limits_{i=1}^m |a_{ij}|\right)$ 与 $\|x\|_\infty = \max\limits_i|\xi_i|$ 不相容.

$$n > 1:\ A_0 = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}_{m\times n},\quad x_0 = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}_{n\times 1},\quad A_0x_0 = \begin{bmatrix} n \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$\|A_0\|_1 = 1,\quad \|x_0\|_\infty = 1,\quad \|A_0x_0\|_\infty = n$$

$$\|A_0x_0\|_\infty = n > 1 = \|A_0\|_1 \cdot \|x_0\|_\infty$$

构造方法

（1）由向量范数构造新的向量范数：

$\boldsymbol{S}_{m\times n}$ 列满秩，$\|\boldsymbol{x}\| = \|\boldsymbol{S}\boldsymbol{x}\|_V$ 是 $\mathbf{C}^n$ 中的向量范数.

（2）由矩阵范数构造向量范数：

非零列向量 $\boldsymbol{y}_0 \in \mathbf{C}^n$，$\|\boldsymbol{x}\| = \|\boldsymbol{x}\boldsymbol{y}_0^{\mathrm{T}}\|_M$ 是 $\mathbf{C}^m$ 中的向量范数.

例如：$\boldsymbol{y}_0 = \boldsymbol{e}_i$ 时，$\boldsymbol{x}\boldsymbol{y}_0^{\mathrm{T}} = \begin{bmatrix} \mathbf{0} & \cdots & \mathbf{0} & \xi_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & \xi_m & \mathbf{0} & \cdots & \mathbf{0} \end{bmatrix}$.

$$\|\boldsymbol{A}\|_M = \|\boldsymbol{A}\|_{m_1} \Rightarrow \|\boldsymbol{x}\| = \|\boldsymbol{x}\|_1;\qquad \|\boldsymbol{A}\|_M = \|\boldsymbol{A}\|_2 \Rightarrow \|\boldsymbol{x}\| = \|\boldsymbol{x}\|_2$$

$$\|\boldsymbol{A}\|_M = \|\boldsymbol{A}\|_{m_2} \Rightarrow \|\boldsymbol{x}\| = \|\boldsymbol{x}\|_2;\qquad \|\boldsymbol{A}\|_M = \|\boldsymbol{A}\|_\infty \Rightarrow \|\boldsymbol{x}\| = \|\boldsymbol{x}\|_\infty$$

（3）由向量范数构造矩阵范数：

$\|\boldsymbol{A}\| = \max\limits_{\|\boldsymbol{x}\|_V=1} \|\boldsymbol{A}\boldsymbol{x}\|_V$ 是 $\mathbf{C}^{m\times n}$ 中的矩阵范数.

（4）由矩阵范数构造新的矩阵范数：

$\boldsymbol{S}_{n\times n}$ 可逆，$\|\boldsymbol{A}\| = \|\boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S}\|_M$ 是 $\mathbf{C}^{n\times n}$ 中的向量范数.

第三章　矩阵分析及其应用

引言：一元多项式　$f(t)=c_0+c_1t+\cdots+c_mt^m$

矩阵多项式　$f(A)=c_0I+c_1A+\cdots+c_mA^m \quad (\forall A_{n\times n})$

易见，$f(A)$是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

§3.1 矩阵序列

一、敛散性：将矩阵序列 $A^{(k)}=\left(a_{ij}^{(k)}\right)_{m\times n} \quad (k=1,2,3,\cdots)$ 记作$\left\{A^{(k)}\right\}$.

若$\lim\limits_{k\to\infty}a_{ij}^{(k)}=a_{ij} \quad (\forall\, i,j)$，称$\left\{A^{(k)}\right\}$收敛于$A=(a_{ij})_{m\times n}$，记作

$$\lim_{k\to\infty}A^{(k)}=A\text{，或者}\quad A^{(k)}\to A \quad (k\to\infty)$$

若数列$\left\{a_{ij}^{(k)}\right\}$之一发散，称$\left\{A^{(k)}\right\}$发散.

性质 1　若$A^{(k)}\to A_{m\times n},B^{(k)}\to B_{m\times n}$，则对$\forall\lambda_1,\lambda_2\in\mathbf{C}$，有

$$\left(\lambda_1A^{(k)}+\lambda_2B^{(k)}\right)\to\left(\lambda_1A+\lambda_2B\right)$$

性质 2　若$A^{(k)}\to A_{m\times n},B^{(k)}\to B_{n\times l}$，则 $\left(A^{(k)}B^{(k)}\right)\to(AB)$.

性质 3　若$A^{(k)}$与A是可逆矩阵，且$A^{(k)}\to A$，则$\left(A^{(k)}\right)^{-1}\to A^{-1}$.

Th1　$A^{(k)}\to O_{m\times n}\Leftrightarrow\forall\|\cdot\|,\ \left\|A^{(k)}\right\|\to 0$；

$A^{(k)}\to A\Leftrightarrow\forall\|\cdot\|,\ \left\|A^{(k)}-A\right\|\to 0$.

证　(1) 考虑矩阵范数$\|\cdot\|_{m_1}$：

$$A^{(k)}\to O_{m\times n}\Leftrightarrow a_{ij}^{(k)}\to 0\ (\text{all } i,j)\Leftrightarrow\left|a_{ij}^{(k)}\right|\to 0\ (\text{all } i,j)$$

$$\Leftrightarrow\sum_i\sum_j\left|a_{ij}^{(k)}\right|\to 0\Leftrightarrow\left\|A^{(k)}\right\|_{m_1}\to 0\ (k\to\infty)$$

(2) $A^{(k)}\to A\Leftrightarrow a_{ij}^{(k)}\to a_{ij}\ (\text{all } i,j)\Leftrightarrow\left(a_{ij}^{(k)}-a_{ij}\right)\to 0$

$$\Leftrightarrow\left(A^{(k)}-A\right)\to O_{m\times n}\Leftrightarrow\forall\|\cdot\|,\ \left\|A^{(k)}-A\right\|\to 0$$

二、收敛矩阵：若$A_{n\times n}$满足$A^k\to O_{n\times n}$，称A为收敛矩阵.

Th2　A 为收敛矩阵 $\Leftrightarrow \rho(A)<1$.

证 充分性．已知 $\rho(A)<1$，对 $\varepsilon=\frac{1}{2}[1-\rho(A)]>0$，存在矩阵范数 $\|\bullet\|_M$，

$$\text{使得}\quad \|A\|_M \le \rho(A)+\varepsilon=\frac{1}{2}[1+\rho(A)]<1$$

于是有 $\|A^k\|_M \le \|A\|_M^k \to 0$，故由定理 1 可得 $A^k \to O$.

必要性．已知 $A^k \to O$，设 $Ax=\lambda x\ (x\ne 0)$，则有

$$\lambda^k x = A^k x \to 0 \Rightarrow \lambda^k \to 0 \Rightarrow |\lambda|<1$$

故 $\rho(A)<1$.

Th3　若矩阵范数 $\|\bullet\|_M$ 使 $\|A\|_M<1$，则 $A^k \to O$.

证　$\rho(A)\le\|A\|_M<1 \Rightarrow A^k \to O$.

例如：$A=\begin{bmatrix}0.1 & 0.3\\ 0.7 & 0.6\end{bmatrix}$，$\|A\|_1=0.9<1\Rightarrow A^k\to O$

§3.2 矩阵级数

$$A^{(k)}=\left(a_{ij}^{(k)}\right)_{m\times n}:\quad A^{(0)}+A^{(1)}+\cdots+A^{(k)}+\cdots\cdots \overset{\Delta}{=} \sum_{k=0}^{\infty}A^{(k)}$$

部分和 $S^{(N)} \overset{\Delta}{=} \sum_{k=0}^{N}A^{(k)}$ 构成矩阵序列 $\{S^{(N)}\}$.

一、敛散性　若 $\lim\limits_{N\to\infty}S^{(N)}=S$，称 $\sum A^{(k)}$ 收敛于 S，记作 $\sum A^{(k)}=S$；

若 $\{S^{(N)}\}$ 发散，称 $\sum A^{(k)}$ 发散.

性质 1　$\sum A^{(k)}=S \Leftrightarrow \sum a_{ij}^{(k)}=s_{ij}$　(all i,j)

证　左 $\Leftrightarrow \lim\limits_{N\to\infty}S^{(N)}=S \Leftrightarrow \lim\limits_{N\to\infty}s_{ij}^{(N)}=s_{ij}$　(all i,j)

$$\Leftrightarrow \lim_{N\to\infty}\sum_{k=0}^{N}a_{ij}^{(k)}=s_{ij}\ \text{(all } i,j\text{)} \Leftrightarrow \text{右}$$

性质 2　若 $\sum\left|a_{ij}^{(k)}\right|$ 收敛 (all 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　$\sum A^{(k)}$ 绝对收敛 $\Leftrightarrow \forall\|\cdot\|, \sum\|A^{(k)}\|$ 收敛.

证　只对矩阵范数 $\|\bullet\|_{m_1}$ 证明即可.

必要性. $\sum A^{(k)}$ 绝对收敛时，正项级数 $\sum\left|a_{ij}^{(k)}\right|$ 收敛，那么正项数列

$p_{ij}^{(N)} \overset{\Delta}{=} \sum_{k=0}^{N}\left|a_{ij}^{(k)}\right|$ 收敛 (all i,j)，于是 $\exists M_{ij}>0$，使得 $p_{ij}^{(N)} \le M_{ij}$

（$\forall N$），从而有 $p_{ij}^{(N)} \le M = \max\limits_{i,j} M_{ij}\quad (\forall i,j,N)$ 及

$$\sum_{k=0}^{N}\left\|A^{(k)}\right\|_{m_1} = \sum_{k=0}^{N}\left(\sum_{i,j}\left|a_{ij}^{(k)}\right|\right) = \sum_{i,j}\left(\sum_{k=0}^{N}\left|a_{ij}^{(k)}\right|\right) \le (mn)M$$

故 $\sum\left\|A^{(k)}\right\|_{m_1}$ 收敛.（正项级数收敛 $\Leftrightarrow$ 部分和有上界）

充分性. $\sum\left\|A^{(k)}\right\|_{m_1}$ 收敛时，因为 $\left|a_{ij}^{(k)}\right| \le \sum\limits_{i,j}\left|a_{ij}^{(k)}\right| = \left\|A^{(k)}\right\|_{m_1}$，所以

$$\sum\left|a_{ij}^{(k)}\right| \text{收敛 (all } i,j) \Rightarrow \sum A^{(k)} \text{ 绝对收敛}$$

性质 4　$\sum A^{(k)} \overset{\text{收}}{=} S \Rightarrow \sum\left(PA^{(k)}Q\right) \overset{\text{收}}{=} PSQ$

$\sum A^{(k)}$ 绝对收敛 $\Rightarrow \sum(PA^{(k)}Q)$ 绝对收敛

证 (1) $S^{(N)} = \sum\limits_{k=0}^{N} A^{(k)} \to S \Rightarrow \sum\limits_{k=0}^{N} PA^{(k)}Q = PS^{(N)}Q \to PSQ$

$$\Rightarrow \sum PA^{(k)}Q = PSQ$$

(2) 指定矩阵范数 $\|\bullet\|$，由性质 3 知 $\sum\left\|A^{(k)}\right\|$ 收敛.

因为 $\left\|PA^{(k)}Q\right\| \le \|P\|\left\|A^{(k)}\right\|\|Q\| = M\left\|A^{(k)}\right\| \quad \left(M \overset{\Delta}{=} \|P\|\|Q\|\right)$

所以 $\sum\limits_{k=0}^{N}\left\|PA^{(k)}Q\right\| \le \sum\limits_{k=0}^{N}\left(M\left\|A^{(k)}\right\|\right) = M \cdot \sum\limits_{k=0}^{N}\left\|A^{(k)}\right\|$ 有界

故 $\sum\left\|PA^{(k)}Q\right\|$ 收敛 $\Rightarrow \sum\left(PA^{(k)}Q\right)$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** 乘积

$$A^{(1)}B^{(1)}+\left[A^{(1)}B^{(2)}+A^{(2)}B^{(1)}\right]+\left[A^{(1)}B^{(3)}+A^{(2)}B^{(2)}+A^{(3)}B^{(1)}\right]+\cdots\cdots$$

$$+\left[A^{(1)}B^{(k)}+A^{(2)}B^{(k-1)}+\cdots+A^{(k)}B^{(1)}\right]+\cdots\cdots$$

绝对收敛于 ST, 记作 $\sum A^{(k)}\cdot\sum B^{(k)}=ST$.

二、**Neumann-级数**: $A_{n\times n}$, $\sum_{k=0}^{\infty} A^{k}$ $\left(A^{0}=I\right)$

Th4 $A_{n\times n}$, $\sum A^{k}$ 收敛 $\Leftrightarrow A^{k}\to O$; $\sum A^{k}$ 收敛时，其和为 $(I-A)^{-1}$.

证 必要性. $\sum A^{k}$ 收敛时, $\sum\left(A^{k}\right)_{ij}$ 收敛 (all i,j) $\Rightarrow\left(A^{k}\right)_{ij}\to 0$,

即 $A^{k}\to O$.

充分性. $A^{k}\to O$: 由定理 2 知 $\rho(A)<1\Rightarrow(I-A)$可逆

$$\left(I+A+A^{2}+\cdots+A^{N}\right)(I-A)=I-A^{N+1}$$

$$\left(I+A+A^{2}+\cdots+A^{N}\right)=(I-A)^{-1}-A^{N+1}(I-A)^{-1}$$

$$\to(I-A)^{-1}\qquad(N\to\infty)$$

即 $\sum A^{k}=(I-A)^{-1}$

Th5 $A_{n\times n}$, $\|A\|<1\Rightarrow\left\|(I-A)^{-1}-\sum_{k=0}^{N} A^{k}\right\|\le\frac{\|A\|^{N+1}}{1-\|A\|}$, $N=0,1,2,\cdots$

证 $\|A\|<1\Rightarrow\rho(A)<1\Rightarrow I-A$ 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: $(I-A)^{-1}-\sum_{k=0}^{N} A^{k}=A^{N+1}(I-A)^{-1}$

恒等式 $A^{N+1}=A^{N+1}(I-A)^{-1}(I-A)$

$$=A^{N+1}(I-A)^{-1}-A^{N+1}(I-A)^{-1}A$$

即　$$A^{N+1}(I-A)^{-1}=A^{N+1}+A^{N+1}(I-A)^{-1}\cdot A$$

$$\left\|A^{N+1}(I-A)^{-1}\right\|\le\left\|A^{N+1}\right\|+\left\|A^{N+1}(I-A)^{-1}\right\|\cdot\|A\|$$

故　$$\left\|A^{N+1}(I-A)^{-1}\right\|\le\frac{\left\|A^{N+1}\right\|}{1-\|A\|}\Rightarrow\left\|(I-A)^{-1}-\sum_{k=0}^{N}A^k\right\|\le\frac{\|A\|^{N+1}}{1-\|A\|}$$

三、幂级数

对于函数 $f(z)=\sum c_k z^k\ \left(|z|<r\right)$，方阵 $A_{n\times n}$，构造矩阵幂级数 $\sum\limits_{k=0}^{\infty}c_kA^k$.

Th6　(1) $\rho(A)<r\Rightarrow\sum c_kA^k$ 绝对收敛；

(2) $\rho(A)>r\Rightarrow\sum c_kA^k$ 发散.

证 (1) 对 A，取 $\varepsilon=\frac{1}{2}[r-\rho(A)]>0$，存在矩阵范数 $\|\cdot\|_\varepsilon$，使得

$$\|A\|_\varepsilon\le\rho(A)+\varepsilon=\frac{1}{2}[r+\rho(A)]<r$$

$$\left\|c_kA^k\right\|_\varepsilon\le|c_k|\|A\|_\varepsilon^k\le|c_k|[\rho(A)+\varepsilon]^k$$

当 $|z|<r$ 时，$\sum c_kz^k$ 绝对收敛，即 $\sum|c_k||z|^k$ 收敛，于是可得

$$\sum|c_k|[\rho(A)+\varepsilon]^k\text{ 收敛}\Rightarrow\sum\left\|c_kA^k\right\|_\varepsilon\text{ 收敛}\Rightarrow\sum c_kA^k\text{ 绝对收敛}$$

(2)　设 A 的特征值为 $\lambda_1,\cdots,\lambda_n$，则存在可逆矩阵 P，使得

$$P^{-1}AP=\begin{bmatrix}\lambda_1 & * & \cdots & * \\ & \lambda_2 & \ddots & \vdots \\ & & \ddots & * \\ & & & \lambda_n\end{bmatrix}\overset{\Delta}{=}S,\quad S^k=\begin{bmatrix}\lambda_1^k & * & \cdots & * \\ & \lambda_2^k & \ddots & \vdots \\ & & \ddots & * \\ & & & \lambda_n^k\end{bmatrix}$$

$\sum c_kS^k$ 的对角元素为 $\sum c_k\lambda_i^k\quad(i=1,2,\cdots,n)$

$\rho(A)>r\Rightarrow\exists\ |\lambda_l|>r\Rightarrow\sum c_k\lambda_l^k$ 发散

$\Rightarrow\sum c_kS^k$ 发散 $\Rightarrow\sum c_kA^k$ 发散

（否则，$\sum c_kA^k$ 收敛 $\Rightarrow\sum\left(P^{-1}(c_kA^k)P\right)$ 收敛 $\Rightarrow\sum c_kS^k$ 收敛）

§3.3　矩阵函数

一、矩阵函数

设一元函数 $f(z)$ 在 $z=0$ 点的幂级数为

$$f(z)=c_0+c_1z+\cdots+c_kz^k+\cdots \quad (|z|<r,\ r>0)$$

构造矩阵函数

$$f(A)\overset{\Delta}{=}c_0I+c_1A+\cdots+c_kA^k+\cdots \quad (\rho(A)<r)$$

例 1　$e^z=1+\dfrac{1}{1!}z+\cdots+\dfrac{1}{k!}z^k+\cdots \quad (r=+\infty)$

$$e^A=I+\frac{1}{1!}A+\cdots+\frac{1}{k!}A^k+\cdots \quad (\forall A_{n\times n})$$

$$\sin z=z-\frac{1}{3!}z^3+\cdots+(-1)^k\frac{1}{(2k+1)!}z^{2k+1}+\cdots \quad (r=+\infty)$$

$$\sin A=A-\frac{1}{3!}A^3+\cdots+(-1)^k\frac{1}{(2k+1)!}A^{2k+1}+\cdots \quad (\forall A_{n\times n})$$

例 2　$f(z)=\dfrac{1}{1-z}=\sum\limits_{k=0}^{\infty}z^k \quad (|z|<1)$，$f(A)\overset{\Delta}{=}\sum\limits_{k=0}^{\infty}A^k \quad (\rho(A)<1)$.

例 3　$\forall A_{n\times n}$，$e^{jA}=\cos A+j\sin A \quad (j=\sqrt{-1})$

$$\cos A=\frac{1}{2}\left(e^{jA}+e^{-jA}\right),\quad \cos(-A)=\cos A$$

$$\sin A=\frac{1}{2j}\left(e^{jA}-e^{-jA}\right),\quad \sin(-A)=-(\sin A)$$

证　在第一式中，视“jA”为整体，并按“奇偶次幂”组项可得

$$e^{jA}=\left[I+\frac{1}{2!}(jA)^2+\frac{1}{4!}(jA)^4+\cdots\right]+\left[\frac{1}{1!}(jA)+\frac{1}{3!}(jA)^3+\cdots\right]$$

$$=\cos A+j\sin A$$

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 $A^2 = A$： $e^A = I + \left(\frac{1}{1!} + \frac{1}{2!} + \cdots\right)A = I + (e-1)A = \begin{bmatrix} e & e-1 \\ 0 & 1 \end{bmatrix}$

$B^2 = B$： $e^B = I + \left(\frac{1}{1!} + \frac{1}{2!} + \cdots\right)B = I + (e-1)B = \begin{bmatrix} e & 1-e \\ 0 & 1 \end{bmatrix}$

$A + B = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} = 2^1\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, $(A+B)^k = 2^k\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$

$e^{A+B} = I + \left(\frac{2^1}{1!} + \frac{2^2}{2!} + \cdots\right)\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = I + (e^2 - 1)\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} e^2 & 0 \\ 0 & 1 \end{bmatrix}$

[注] $e^A e^B = \begin{bmatrix} e^2 & -(e-1)^2 \\ 0 & 1 \end{bmatrix}$, $e^B e^A = \begin{bmatrix} e^2 & (e-1)^2 \\ 0 & 1 \end{bmatrix}$

Th7 $A_{n\times n}, B_{n\times n}, AB = BA \Rightarrow e^A e^B = e^B e^A = e^{A+B}$

证 $e^A e^B = \left[I + A + \frac{1}{2!}A^2 + \frac{1}{3!}A^3 + \cdots\right]\left[I + B + \frac{1}{2!}B^2 + \frac{1}{3!}B^3 + \cdots\right]$

$$= I + (A+B) + \frac{1}{2!}(A^2 + 2AB + B^2) + \frac{1}{3!}(A^3 + 3A^2B + 3AB^2 + B^3) + \cdots$$

$$= I + (A+B) + \frac{1}{2!}(A+B)^2 + \frac{1}{3!}(A+B)^3 + \cdots = e^{A+B}$$

同理 $e^B e^A = e^{B+A} = e^{A+B}$

[注] ① $e^A e^{-A} = e^O = I \Rightarrow (e^A)^{-1} = e^{-A}$ $(\forall A_{n\times n})$

② $(e^A)^m = e^{mA}$, $m = 2,3,\cdots$

例 5 $A_{n\times n}, B_{n\times n}, AB = BA$：

$$\cos(A+B) = \cos A \cdot \cos B - \sin A \cdot \sin B \qquad (1)$$

$$\sin(A+B) = \sin A \cdot \cos B + \cos A \cdot \sin B \qquad (2)$$

证 式(1)中

$$\text{右端} = \frac{1}{2}\left[e^{jA} + e^{-jA}\right] \cdot \frac{1}{2}\left[e^{jB} + e^{-jB}\right] - \frac{1}{2j}\left[e^{jA} - e^{-jA}\right] \cdot \frac{1}{2j}\left[e^{jB} - e^{-jB}\right]$$

$$=\frac{1}{4}\left[e^{j(A+B)}+\cdots+\cdots+e^{-j(A+B)}\right]+\frac{1}{4}\left[e^{j(A+B)}-\cdots-\cdots+e^{-j(A+B)}\right]$$

$$=\frac{1}{2}\left[e^{j(A+B)}+e^{-j(A+B)}\right]=\text{左端}$$

二、矩阵函数值的计算

1． 待定系数法　设 $A_{n\times n}$ 的零化多项式为

$$\psi(\lambda)=\lambda^m+c_1\lambda^{m-1}+\cdots+c_{m-1}\lambda+c_m \quad (1\le m\le n)$$

且满足 $\psi(\lambda)\,|\,\varphi(\lambda)$，分解

$$\psi(\lambda)=(\lambda-\lambda_1)^{m_1}\cdots(\lambda-\lambda_s)^{m_s} \quad \left(\lambda_i\ne\lambda_j,\sum m_i=m\right)$$

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　$$f(z)=\sum c_k z^k=\psi(z)\cdot g(z)+r(z)$$

$$r(z)=b_0+b_1z+\cdots+b_{m-1}z^{m-1}$$

由 $\psi(\lambda_i)=0,\psi'(\lambda_i)=0,\cdots,\psi^{(m_i-1)}(\lambda_i)=0$ 可得

$$\left.\begin{array}{l} r(\lambda_i)=f(\lambda_i)\\ r'(\lambda_i)=f'(\lambda_i)\\ \quad\cdots\cdots\\ r^{(m_i-1)}(\lambda_i)=f^{(m_i-1)}(\lambda_i)\end{array}\right\}\quad (i=1,2,\cdots,s)$$

解此方程组得出 $b_0,b_1,\cdots,b_{m-1}$．因为 $\psi(A)=O$，所以

$$f(A)=\sum c_kA^k=\psi(A)\cdot g(A)+r(A)=r(A)$$

即　$$f(A)=b_0I+b_1A+\cdots+b_{m-1}A^{m-1}$$

例 6　$A=\begin{bmatrix}2&0&0\\1&1&1\\1&-1&3\end{bmatrix}$，求 $e^A,e^{tA}\ (t\in\mathbf{R})$.

解　$\varphi(\lambda)=|\lambda I-A|=(\lambda-2)^3$，$A-2I=\begin{bmatrix}0&0&0\\1&-1&1\\1&-1&1\end{bmatrix},(A-2I)^2=O$

取 $\psi(\lambda)=m(\lambda)=(\lambda-2)^2$

(1) $f(\lambda)=e^{\lambda}=\psi(\lambda)\cdot g(\lambda)+(a+b\lambda)$

$$f'(\lambda)=e^{\lambda}=[\psi(\lambda)\cdot g(\lambda)]'+b$$

$$\left.\begin{array}{l} f(2)=e^2 \ : a+2b=e^2 \\ f'(2)=e^2 : b=e^2 \end{array}\right\} \Rightarrow \begin{cases} a=-e^2 \\ b=e^2 \end{cases}$$

$$e^A=e^2(A-I)=e^2\begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & 2 \end{bmatrix}$$

(2) $f(\lambda)=e^{t\lambda}=\psi(\lambda)\cdot g(\lambda)+(a+b\lambda)$

$$f'(\lambda)=te^{t\lambda}=[\psi(\lambda)\cdot g(\lambda)]'+b$$

$$\left.\begin{array}{l} f(2)=e^{2t} \ \ : a+2b=e^{2t} \\ f'(2)=te^{2t} : b=te^{2t} \end{array}\right\} \Rightarrow \begin{cases} a=(1-2t)e^{2t} \\ b=te^{2t} \end{cases}$$

$$e^{tA}=e^{2t}[(1-2t)I+tA]=e^{2t}\begin{bmatrix} 1 & 0 & 0 \\ t & 1-t & t \\ t & -t & 1+t \end{bmatrix}$$

2．数项级数求和法　设 $A_{n\times n}$ 的零化多项式如上所述，则有

$$\psi(A)=O \Rightarrow \begin{cases} A^m &= k_0 I + k_1 A + \cdots + k_{m-1}A^{m-1} \\ A^{m+1} &= k_0^{(1)} I + k_1^{(1)} A + \cdots + k_{m-1}^{(1)}A^{m-1} \\ & \cdots\cdots \\ A^{m+l} &= k_0^{(l)} I + k_1^{(l)} A + \cdots + k_{m-1}^{(l)}A^{m-1} \end{cases}$$

$$\begin{aligned} f(A)=\sum c_k A^k &= \left(c_0 I + c_1 A + \cdots + c_{m-1}A^{m-1}\right)+ \\ & c_m\left(k_0 I + k_1 A + \cdots k_{m-1}A^{m-1}\right)+\cdots+ \\ & c_{m+l}\left(k_0^{(l)} I + k_1^{(l)} A + \cdots + k_{m-1}^{(l)}A^{m-1}\right)+\cdots \end{aligned}$$

$$=\left[c_0+\sum_{l=0}^{\infty} c_{m+l}k_0^{(l)}\right]I+\left[c_1+\sum c_{m+l}k_1^{(l)}\right]A+\cdots+\left[c_{m-1}+\sum c_{m+l}k_{m-1}^{(l)}\right]A^{m-1}$$

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 $\psi(\lambda)=\varphi(\lambda)$：

$$\psi(A)=O\Rightarrow A^4=\pi^2A^2, A^5=\pi^2A^3, A^7=\pi^4A^3,\cdots$$

$$\sin A=A-\frac{1}{3!}A^3+\frac{1}{5!}A^5-\frac{1}{7!}A^7+\cdots$$

$$=A+\left[-\frac{1}{3!}+\frac{\pi^2}{5!}-\frac{\pi^4}{7!}+\cdots\right]A^3=A+\frac{1}{\pi^3}\left[-\frac{\pi^3}{3!}+\frac{\pi^5}{5!}-\frac{\pi^7}{7!}+\cdots\right]A^3$$

$$=A+\frac{1}{\pi^3}(\sin\pi-\pi)A^3=A-\frac{1}{\pi^2}A^3$$

$$=\begin{bmatrix}0 & & & \\ & 0 & & \\ & & 0 & 1\\ & & & 0\end{bmatrix}; \qquad A^3=\begin{bmatrix}\pi^3 & & & \\ & -\pi^3 & & \\ & & 0 & \\ & & & 0\end{bmatrix}$$

3．对角形法

设 $P^{-1}AP=\begin{bmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{bmatrix}\overset{\Delta}{=}\Lambda$，则 $A^k=P\Lambda^kP^{-1}$，且有

$$\sum_{k=0}^{N}c_kA^k=P\cdot\sum_{k=0}^{N}c_k\Lambda^k\cdot P^{-1}=P\cdot\begin{bmatrix}\sum_{k=0}^{N}c_k\lambda_1^k & & \\ & \ddots & \\ & & \sum_{k=0}^{N}c_k\lambda_n^k\end{bmatrix}\cdot P^{-1}$$

$$f(A)=\sum_{k=0}^{\infty}c_kA^k=P\cdot\begin{bmatrix}f(\lambda_1) & & \\ & \ddots & \\ & & f(\lambda_n)\end{bmatrix}\cdot P^{-1}$$

例 8　$P^{-1}AP=\Lambda$：

$$e^A=P\cdot\operatorname{diag}\left(e^{\lambda_1},\cdots,e^{\lambda_n}\right)\cdot P^{-1}$$

$$e^{tA}=P\cdot\operatorname{diag}\left(e^{\lambda_1 t},\cdots,e^{\lambda_n t}\right)\cdot P^{-1}$$

$$\sin A=P\cdot\operatorname{diag}(\sin\lambda_1,\cdots,\sin\lambda_n)\cdot P^{-1}$$

4．Jordan 标准形法　设

$$P^{-1}AP=J=\begin{bmatrix}J_1 & & \\ & \ddots & \\ & & J_s\end{bmatrix},\ J_i=\begin{bmatrix}\lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1\\ & & & \lambda_i\end{bmatrix}=\lambda_iI+I^{(1)}$$

易证：$I^{(k)}I^{(1)}=I^{(1)}I^{(k)}=I^{(k+1)}$，　$I^{(m_i)}=O$

$k\le m_i-1$：　$J_i^k=\lambda_i^kI+C_k^1\lambda_i^{k-1}I^{(1)}+\cdots+C_k^{k-1}\lambda_iI^{(k-1)}+I^{(k)}$

$k\ge m_i$：　$J_i^k=\lambda_i^kI+C_k^1\lambda_i^{k-1}I^{(1)}+\cdots+C_k^{m_i-1}\lambda_i^{k-(m_i-1)}I^{(m_i-1)}$

$$f(J_i)=\sum c_kJ_i^k=f(\lambda_i)I+\frac{f'(\lambda_i)}{1!}I^{(1)}+\cdots+\frac{f^{(m_i-1)}(\lambda_i)}{(m_i-1)!}I^{(m_i-1)}$$

$$f(A)=\sum c_kA^k=P\cdot\sum c_kJ^k\cdot P^{-1}$$

$$=P\cdot\begin{bmatrix}\sum c_k J_1^k & & \\ & \ddots & \\ & & \sum c_k J_s^k\end{bmatrix}\cdot P^{-1}=P\cdot\begin{bmatrix}f(J_1) & & \\ & \ddots & \\ & & f(J_s)\end{bmatrix}\cdot P^{-1}$$

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 $(k=0,1,2,\cdots)$

② $\lim\limits_{k\to\infty}\dfrac{f^{(k+1)}(\xi)}{(k+1)!}z^{k+1}=0\ \left(|z|<r\right)$

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，拓宽定义如下：设 $f(z)\in C^{m_i-1}[\delta(\lambda_i)]$，令

$$f(J_i)\overset{\Delta}{=}f(\lambda_i)I+\frac{f'(\lambda_i)}{1!}I^{(1)}+\cdots+\frac{f^{(m_i-1)}(\lambda_i)}{(m_i-1)!}I^{(m_i-1)}$$

$$f(A)\overset{\Delta}{=}P\cdot\mathrm{diag}\left(f(J_1),\cdots,f(J_s)\right)\cdot P^{-1}$$

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 $f(z)$ 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 $f(z)$ 能够展为“z”的幂级数时，矩阵函数 $f(A)$ 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　$f(z)=\dfrac{1}{z}, f'(z)=-z^{-2}, f''(z)=2z^{-3}, f'''(z)=-6z^{-4}$

$$f(A)=f(J)=f(2)\cdot I+f'(2)\cdot I^{(1)}+\frac{f''(2)}{2!}\cdot I^{(2)}+\frac{f'''(2)}{3!}\cdot I^{(3)}$$

$$=\begin{bmatrix} \frac{1}{2} & -\frac{1}{4} & \frac{1}{8} & -\frac{1}{16} \\ & \frac{1}{2} & -\frac{1}{4} & \frac{1}{8} \\ & & \frac{1}{2} & -\frac{1}{4} \\ & & & \frac{1}{2} \end{bmatrix}$$

例 10　$A=\begin{bmatrix} 1 & 1 & 0 \\ & 1 & 0 \\ & & 2 \end{bmatrix}$, $f(z)=\sqrt{z}$，求 $f(A)$.

解　$f(z)=\sqrt{z}$，　$f'(z)=\dfrac{1}{2\sqrt{z}}$

$$J_1=\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}:\quad f(J_1)=f(1)\cdot I+f'(1)\cdot I^{(1)}=\begin{bmatrix} 1 & 1/2 \\ 0 & 1 \end{bmatrix}$$

$$J_2=(2):\quad f(J_2)=f(2)\cdot I=(\sqrt{2})$$

$$f(A)=f(J)=\begin{bmatrix} f(J_1) & \\ & f(J_2) \end{bmatrix}=\begin{bmatrix} 1 & 1/2 & 0 \\ & 1 & 0 \\ & & \sqrt{2} \end{bmatrix}$$

四、矩阵函数的性质

由级数定义或者拓宽定义给出的矩阵函数具有下列性质：

(1)　$f(z)=f_1(z)+f_2(z)\Rightarrow f(A)=f_1(A)+f_2(A)$

$$f^{(l)}(\lambda_i)=f_1^{(l)}(\lambda_i)+f_2^{(l)}(\lambda_i)\Rightarrow f(J_i)=f_1(J_i)+f_2(J_i)$$

$$f(A)=P\cdot\left[\begin{pmatrix} f_1(J_1) & & \\ & \ddots & \\ & & f_1(J_s) \end{pmatrix}+\begin{pmatrix} f_2(J_1) & & \\ & \ddots & \\ & & f_2(J_s) \end{pmatrix}\right]\cdot P^{-1}$$

$$=f_1(A)+f_2(A)$$

(2)　$f(z)=f_1(z)\cdot f_2(z)\Rightarrow f(A)=f_1(A)\cdot f_2(A)=f_2(A)\cdot f_1(A)$

$$f_1(J_i)\cdot f_2(J_i)=\left[f_1\cdot I+\frac{f_1'}{1!}\cdot I^{(1)}+\frac{f_1''}{2!}\cdot I^{(2)}+\cdots+\frac{f_1^{(m_i-1)}}{(m_i-1)!}\cdot I^{(m_i-1)}\right]\cdot$$

$$\left[f_2\cdot I+\frac{f_2'}{1!}\cdot I^{(1)}+\frac{f_2''}{2!}\cdot I^{(2)}+\cdots+\frac{f_2^{(m_i-1)}}{(m_i-1)!}\cdot I^{(m_i-1)}\right]$$

$$=(f_1f_2)\cdot I+\frac{1}{1!}(f_1'f_2+f_1f_2')\cdot I^{(1)}+$$

$$\frac{1}{2!}(f_1''f_2+f_1f_2''+2f_1'f_2')\cdot I^{(2)}+\cdots\cdots$$

$$=(f_1f_2)\cdot I+\frac{1}{1!}(f_1f_2)'\cdot I^{(1)}+\frac{1}{2!}(f_1f_2)''\cdot I^{(2)}+\cdots+\frac{(f_1f_2)^{(m_i-1)}}{(m_i-1)!}\cdot I^{(m_i-1)}$$

$$=f(J_i)$$

$$f(A)=P\begin{bmatrix}f(J_1)&&\\&\ddots&\\&&f(J_s)\end{bmatrix}P^{-1}$$

$$=P\begin{bmatrix}f_1(J_1)&&\\&\ddots&\\&&f_1(J_s)\end{bmatrix}P^{-1}\cdot P\begin{bmatrix}f_2(J_1)&&\\&\ddots&\\&&f_2(J_s)\end{bmatrix}P^{-1}$$

$$=f_1(A)\cdot f_2(A)$$

§3.4 矩阵的微分与积分

一、函数矩阵的导数

设 $A(t)=(a_{ij}(t))_{m\times n}$ ，若 $a_{ij}(t)$ 可导，称 $A(t)$ 可导，记作

$$\frac{d}{dt}A(t)=(a_{ij}'(t))_{m\times n}\quad 或者\quad A'(t)=(a_{ij}'(t))_{m\times n}$$

Th8 设 $A(t),B(t)$ 可导，则有

(1) $A_{m\times n},B_{m\times n}$， $\dfrac{d}{dt}[A(t)+B(t)]=A'(t)+B'(t)$

(2) $A_{m\times n},f(t)$ 可导， $\dfrac{d}{dt}[f(t)A(t)]=f'(t)A(t)+f(t)A'(t)$

(3) $A_{m\times n}, B_{n\times l}$, $\frac{d}{dt}[A(t)B(t)] = A'(t)B(t) + A(t)B'(t)$

证 (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

Th9　设 $A_{n\times n}$ 为数量矩阵，则有

(1) $\frac{d}{dt}e^{tA} = A\cdot e^{tA} = e^{tA}\cdot A$

(2) $\frac{d}{dt}\cos(tA) = -A\cdot\sin(tA) = -\sin(tA)\cdot A$

(3) $\frac{d}{dt}\sin(tA) = A\cdot\cos(tA) = \cos(tA)\cdot A$

证 (1) $e^{tA} = I + \frac{1}{1!}(tA) + \frac{1}{2!}(tA)^2 + \cdots + \frac{1}{k!}(tA)^k + \cdots\cdots$　　绝对收敛

$(e^{tA})_{ij} = \delta_{ij} + \frac{t}{1!}(A)_{ij} + \frac{t^2}{2!}(A^2)_{ij} + \cdots + \frac{t^k}{k!}(A^k)_{ij} + \cdots\cdots$　　绝对收敛

$\frac{d}{dt}(e^{tA})_{ij} = 0 + (A)_{ij} + \frac{t}{1!}(A^2)_{ij} + \cdots + \frac{t^{k-1}}{(k-1)!}(A^k)_{ij} + \cdots\cdots$　绝对收敛

$\frac{d}{dt}e^{tA} = A + \frac{t}{1!}A^2 + \cdots + \frac{t^{k-1}}{(k-1)!}A^k + \cdots\cdots$　　绝对收敛

$$= \begin{cases} A\cdot\left[I + \frac{1}{1!}(tA) + \cdots + \frac{1}{(k-1)!}(tA)^{k-1} + \cdots\cdots\right] = A\cdot e^{tA} \\ \left[I + \frac{1}{1!}(tA) + \cdots + \frac{1}{(k-1)!}(tA)^{k-1} + \cdots\cdots\right]\cdot A = e^{tA}\cdot A \end{cases}$$

二、函数矩阵的积分

设 $A(t) = (a_{ij}(t))_{m\times n}$，若 $a_{ij}(t)$ 在区间 $[t_0, t]$ 上可积，称 $A(t)$ 可积，记作

$$\int_{t_0}^{t} A(\tau)d\tau = \left(\int_{t_0}^{t} a_{ij}(\tau)d\tau\right)_{m\times n}$$

(1) $\int_{t_0}^{t}[A(\tau)+B(\tau)]d\tau=\int_{t_0}^{t}A(\tau)d\tau+\int_{t_0}^{t}B(\tau)d\tau$

(2) A 为常数矩阵： $\int_{t_0}^{t}[A\cdot B(\tau)]d\tau=A\cdot\left[\int_{t_0}^{t}B(\tau)d\tau\right]$

B 为常数矩阵： $\int_{t_0}^{t}A(\tau)\cdot Bd\tau=\left[\int_{t_0}^{t}A(\tau)d\tau\right]\cdot B$

(3) 设 $a_{ij}(t)\in C[t_0,t_1]$，且 $a\in[t_0,t_1]$，则 $\dfrac{d}{dt}\int_{a}^{t}A(\tau)d\tau=A(t)$.

(4) 设 $a_{ij}'(t)\in C[t_0,t_1]$，则 $\int_{t_0}^{t_1}A'(t)dt=A(t_1)-A(t_0)$.

$$左=\left(\int_{t_0}^{t_1}a_{ij}'(t)dt\right)_{m\times n}=\left(a_{ij}(t_1)-a_{ij}(t_0)\right)_{m\times n}=\left(a_{ij}(t_1)\right)_{m\times n}-\left(a_{ij}(t_0)\right)_{m\times n}=右$$

三、函数对矩阵的导数

$X=(\xi_{ij})_{m\times n}$： $f(X)\overset{\Delta}{=}f(\xi_{11},\xi_{12},\cdots,\xi_{1n},\cdots,\xi_{mn})$

$$\frac{df}{dX}\overset{\Delta}{=}\left(\frac{\partial f}{\partial \xi_{ij}}\right)_{m\times n}=\begin{bmatrix}\dfrac{\partial f}{\partial \xi_{11}} & \cdots & \dfrac{\partial f}{\partial \xi_{1n}}\\ \vdots & & \vdots\\ \dfrac{\partial f}{\partial \xi_{m1}} & \cdots & \dfrac{\partial f}{\partial \xi_{mn}}\end{bmatrix}$$

例 1　$x=\begin{bmatrix}\xi_1\\ \vdots\\ \xi_n\end{bmatrix}$： $f(x)\overset{\Delta}{=}f(\xi_1,\cdots,\xi_n)$， $\dfrac{df}{dx}=\begin{bmatrix}\dfrac{\partial f}{\partial \xi_1}\\ \vdots\\ \dfrac{\partial f}{\partial \xi_n}\end{bmatrix}$

例 2　$A=(a_{ij})_{n\times n}$，$x=\begin{bmatrix}\xi_1\\ \vdots\\ \xi_n\end{bmatrix}$： $f(x)=x^{\mathrm{T}}Ax$，求 $\dfrac{df}{dx}$.

解　$f(x)=\xi_1\cdot\sum_{j=1}^{n}a_{1j}\xi_j+\cdots+\xi_k\cdot\sum_{j=1}^{n}a_{kj}\xi_j+\cdots+\xi_n\cdot\sum_{j=1}^{n}a_{nj}\xi_j$

$$\frac{\partial f}{\partial \xi_k}=\xi_1\cdot a_{1k}+\cdots+\xi_{k-1}\cdot a_{k-1,k}+$$

$$\left(\sum_{j=1}^{n}a_{kj}\xi_j+\xi_k\cdot a_{kk}\right)+\xi_{k+1}\cdot a_{k+1,k}+\cdots\xi_n\cdot a_{nk}$$

$$=\sum_{j=1}^{n}a_{kj}\xi_j+\sum_{i=1}^{n}a_{ik}\xi_i$$

$$\frac{df}{dx}=\begin{bmatrix}\sum_{j=1}^{n}a_{1j}\xi_j+\sum_{i=1}^{n}a_{i1}\xi_i\\ \vdots\\ \sum_{j=1}^{n}a_{nj}\xi_j+\sum_{i=1}^{n}a_{in}\xi_i\end{bmatrix}=\left(A+A^{\mathrm{T}}\right)x \qquad \text{“}A^{\mathrm{T}}=A\Rightarrow\frac{df}{dx}=2Ax\text{”}$$

例 3　$X=\left(\xi_{ij}\right)_{n\times n}$：$f(X)\overset{\Delta}{=}[\mathrm{tr}(X)]^2$，求 $\left.\dfrac{df}{dX}\right|_{X=I_n}$.

解　$f(X)=\left(\xi_{11}+\xi_{22}+\cdots+\xi_{nn}\right)^2$

$$\frac{df}{dX}=2\left(\xi_{11}+\xi_{22}+\cdots+\xi_{nn}\right)I_n,\quad \left.\frac{df}{dX}\right|_{X=I_n}=2nI_n\ .$$

例 4　$A\in\mathbf{R}^{m\times n},b\in\mathbf{R}^m$，若 $x\in\mathbf{R}^n$ 使得 $\|Ax-b\|_2=\min$，则 $A^{\mathrm{T}}Ax=A^{\mathrm{T}}b$.

证　$f(x)\overset{\Delta}{=}\|Ax-b\|_2^2=(Ax-b)^{\mathrm{T}}(Ax-b)=x^{\mathrm{T}}A^{\mathrm{T}}Ax-2b^{\mathrm{T}}Ax+b^{\mathrm{T}}b$

$$g(x)\overset{\Delta}{=}b^{\mathrm{T}}Ax=b_1\cdot\sum_{j=1}^{n}a_{1j}\xi_j+\cdots+b_m\cdot\sum_{j=1}^{n}a_{mj}\xi_j$$

$$\frac{dg}{dx}=\begin{bmatrix}\dfrac{\partial g}{\partial\xi_1}\\ \vdots\\ \dfrac{\partial g}{\partial\xi_n}\end{bmatrix}=\begin{bmatrix}b_1a_{11}+\cdots+b_ma_{m1}\\ \vdots\\ b_1a_{1n}+\cdots+b_ma_{mn}\end{bmatrix}=A^{\mathrm{T}}b$$

$$\frac{df}{dx}=2A^{\mathrm{T}}Ax-2A^{\mathrm{T}}b=0\Rightarrow A^{\mathrm{T}}Ax=A^{\mathrm{T}}b$$

[注]　$r(A^{\mathrm{T}}A)=r(A)\Rightarrow r(A^{\mathrm{T}}A\mid A^{\mathrm{T}}b)=r(A^{\mathrm{T}}A)\Rightarrow A^{\mathrm{T}}Ax=A^{\mathrm{T}}b$ 有解

四、函数矩阵对矩阵的导数

$$X=\left(\xi_{ij}\right)_{m\times n},\quad f_{kl}(X)\overset{\Delta}{=}f_{kl}\left(\xi_{11},\xi_{12},\cdots,\xi_{1n},\cdots,\xi_{mn}\right)$$

$$F=\begin{bmatrix}f_{11}&\cdots&f_{1s}\\ \vdots&&\vdots\\ f_{r1}&\cdots&f_{rs}\end{bmatrix},\ \frac{\partial F}{\partial\xi_{ij}}\overset{\Delta}{=}\begin{bmatrix}\dfrac{\partial f_{11}}{\partial\xi_{ij}}&\cdots&\dfrac{\partial f_{1s}}{\partial\xi_{ij}}\\ \vdots&&\vdots\\ \dfrac{\partial f_{r1}}{\partial\xi_{ij}}&\cdots&\dfrac{\partial f_{rs}}{\partial\xi_{ij}}\end{bmatrix},\ \frac{dF}{dX}\overset{\Delta}{=}\begin{bmatrix}\dfrac{\partial F}{\partial\xi_{11}}&\cdots&\dfrac{\partial F}{\partial\xi_{1n}}\\ \vdots&&\vdots\\ \dfrac{\partial F}{\partial\xi_{m1}}&\cdots&\dfrac{\partial F}{\partial\xi_{mn}}\end{bmatrix}_{m\times n\text{块}}$$

例 5　$x=\begin{bmatrix}\xi_1\\ \vdots\\ \xi_n\end{bmatrix}, F=\left(f_1(x),\cdots,f_l(x)\right)$：$\dfrac{dF}{dx}=\begin{bmatrix}\dfrac{\partial f_1}{\partial \xi_1} & \cdots & \dfrac{\partial f_l}{\partial \xi_1}\\ \cdots & & \cdots\\ \dfrac{\partial f_1}{\partial \xi_n} & \cdots & \dfrac{\partial f_l}{\partial \xi_n}\end{bmatrix}$

例 6　$A=\left(a_{ij}\right)_{n\times n}$

$$x=\begin{bmatrix}\xi_1\\ \vdots\\ \xi_n\end{bmatrix},\quad Ax=\begin{bmatrix}\sum_{j=1}^{n}a_{1j}\xi_j\\ \vdots\\ \sum_{j=1}^{n}a_{nj}\xi_j\end{bmatrix},\quad \frac{d(Ax)}{dx^{\mathrm{T}}}=\begin{bmatrix}a_{11} & \cdots & a_{1n}\\ \vdots & & \vdots\\ a_{n1} & \cdots & a_{nn}\end{bmatrix}=A$$

§3.5 应用

$$\begin{cases}\xi_1'(t)=a_{11}\xi_1(t)+a_{12}\xi_2(t)+\cdots+a_{1n}\xi_n(t)+b_1(t)\\ \xi_2'(t)=a_{21}\xi_1(t)+a_{22}\xi_2(t)+\cdots+a_{2n}\xi_n(t)+b_2(t)\\ \qquad\cdots\cdots\\ \xi_n'(t)=a_{n1}\xi_1(t)+a_{n2}\xi_2(t)+\cdots+a_{nn}\xi_n(t)+b_n(t)\end{cases}$$

$$A=\begin{bmatrix}a_{11} & \cdots & a_{1n}\\ \vdots & & \vdots\\ a_{n1} & \cdots & a_{nn}\end{bmatrix},\ x(t)=\begin{bmatrix}\xi_1(t)\\ \vdots\\ \xi_n(t)\end{bmatrix},\ b(t)=\begin{bmatrix}b_1(t)\\ \vdots\\ b_n(t)\end{bmatrix},\ c=\begin{bmatrix}c_1\\ \vdots\\ c_n\end{bmatrix}$$

齐次微分方程(1)：　$x'(t)=A\cdot x(t)$

非齐次微分方程(2)：$x'(t)=A\cdot x(t)+b(t)$　（$b(t)\neq 0$）

一、齐次方程 $x'(t)=A\cdot x(t)$ 的解法

Th10　齐次方程 $x'(t)=A\cdot x(t)$ 满足 $x(t_0)=x_0$ 的解存在且唯一.

证　存在性．设 $x(t)=e^{(t-t_0)A}x_0$，则

$$x'(t)=Ae^{(t-t_0)A}x_0=A\cdot x(t),\quad x(t_0)=e^{O}x_0=x_0$$

唯一性．设 $x(t)$ 满足 $x'(t)=A\cdot x(t)$，$x(t_0)=x_0$，则有

$$x'(t)-A\,x(t)=0\ \Rightarrow\ e^{-tA}x'(t)+e^{-tA}(-A)\,x(t)=0$$

$$\left[e^{-tA}x(t)\right]'=0\ \Rightarrow\ e^{-tA}x(t)=c\ \Rightarrow\ x(t)=e^{tA}c$$

因为 $x(t_0)=x_0$，所以 $x_0=e^{t_0A}c\Rightarrow c=e^{-t_0A}x_0$

故 $x(t)=e^{tA}e^{-t_0A}x_0=e^{(t-t_0)A}x_0$

[注] 齐次方程 $x'(t)=A\cdot x(t)$ 满足 $x(0)=x_0$ 的解为 $x(t)=e^{tA}x_0$

例 1 设 $A=\begin{bmatrix}1&1&0\\&1&0\\&&2\end{bmatrix}$，求 $x'(t)=A\cdot x(t)$ 的通解.

解 $e^{tA}=\begin{bmatrix}e^t&te^t&0\\&e^t&0\\&&e^{2t}\end{bmatrix}$，$x(t)=e^{tA}\cdot c=\begin{bmatrix}c_1e^t+c_2te^t\\c_2e^t\\c_3e^{2t}\end{bmatrix}$

例 2 矩阵函数 e^{tA} 的列向量组 $x_1(t),\cdots,x_n(t)$ 构成齐次方程 $x'(t)=A\cdot x(t)$ 的基础解系.

证 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 $x'(t)=A\cdot x(t)$ 的一个解

通解 $x(t)=e^{tA}\cdot c=c_1\cdot x_1(t)+\cdots+c_n\cdot x_n(t)$

二、非齐次方程 $x'(t)=A\cdot x(t)+b(t)$ 的解法

方程(1)：$x'(t)=A\cdot x(t)$

方程(2)：$x'(t)=A\cdot x(t)+b(t)$

$$\left.\begin{array}{l}\tilde{x}(t)\text{是(2)的特解}\\x(t)\text{是(2)的通解}\end{array}\right\}\Rightarrow\begin{cases}\tilde{x}'(t)=A\cdot\tilde{x}(t)+b(t)\\x'(t)=A\cdot x(t)+b(t)\end{cases}$$

$$\Rightarrow[x(t)-\tilde{x}(t)]'=A[x(t)-\tilde{x}(t)]\ \Rightarrow\ x(t)-\tilde{x}(t)\text{ 是(1)的解}$$

$$\overset{\text{例2}}{\Rightarrow}x(t)-\tilde{x}(t)=c_1\cdot x_1(t)+\cdots+c_n\cdot x_n(t)\ \Rightarrow\ x(t)=e^{tA}\cdot c+\tilde{x}(t)$$

采用常向量变异法求 $\tilde{x}(t)$. 设 $\tilde{x}(t)=e^{tA}c(t)$ 满足(2)，则有

$$Ae^{tA}\cdot c(t)+e^{tA}\cdot c'(t)=A\cdot e^{tA}c(t)+b(t)$$

$$c'(t)=e^{-tA}b(t)\Rightarrow c(t)=\int_{t_0}^{t}e^{-\tau A}b(\tau)d\tau\quad\text{（原函数之一）}$$

故(2)的通解为 $x(t)=e^{tA}\cdot\left[c+\int_{t_0}^{t}e^{-\tau A}b(\tau)d\tau\right]$

特解为 $x(t)\big|_{x(t_0)=x_0}=e^{tA}\cdot\left[e^{-t_0A}x_0+\int_{t_0}^{t}e^{-\tau A}b(\tau)d\tau\right]$

[注] 当 $t_0=0$ 时，特解 $x(t)\big|_{x(0)=x_0}=e^{tA}\cdot\left[x_0+\int_{0}^{t}e^{-\tau A}b(\tau)d\tau\right]$

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 $x'(t)=A\cdot x(t)+b(t)$ 满足初始条件 $x(0)$ 的特解.

解 例 1 求得 $e^{tA}=\begin{bmatrix}e^t&te^t&0\\&e^t&0\\&&e^{2t}\end{bmatrix}$

$$e^{-\tau A}\cdot b(\tau)=\begin{bmatrix}e^{-\tau}\\0\\1\end{bmatrix},\quad \int_0^t e^{-\tau A}b(\tau)d\tau=\begin{bmatrix}1-e^{-t}\\0\\t\end{bmatrix}$$

$$x(t)=e^{tA}\cdot\left[\begin{pmatrix}-1\\1\\0\end{pmatrix}+\begin{pmatrix}1-e^{-t}\\0\\t\end{pmatrix}\right]=\begin{bmatrix}te^t-1\\e^t\\te^{2t}\end{bmatrix}$$

例 4 设 $A=\begin{bmatrix}5&4&2\\4&5&-2\\2&-2&8\end{bmatrix}$, $b(t)=\begin{bmatrix}e^{9t}\\e^{9t}\\0\end{bmatrix}$, $x(0)=\begin{bmatrix}1\\0\\2\end{bmatrix}$.

求 $x'(t)=A\cdot x(t)+b(t)$ 满足初始条件 $x(0)$ 的特解.

解 用 3 种方法求 e^{tA}.

(1) 待定法：$|\lambda I-A|=\lambda(\lambda-9)^2$，因为 $A(A-9I)=O$，所以 A 的最小多项式为 $m(\lambda)=\lambda(\lambda-9)$. 设 $f(\lambda)=e^{t\lambda}=m(\lambda)g(\lambda)+(a+b\lambda)$，则

$$\begin{cases} f(0)=1=a \\ f(9)=e^{9t}=a+9b \end{cases} \qquad \begin{cases} a=1 \\ b=(e^{9t}-1)/9 \end{cases}$$

于是 $$e^{tA}=aI+bA=\frac{1}{9}(9I-A)+\frac{e^{9t}}{9}A$$

$$=\frac{1}{9}\begin{bmatrix} 4 & -4 & -2 \\ -4 & 4 & 2 \\ -2 & 2 & 1 \end{bmatrix}+\frac{e^{9t}}{9}\begin{bmatrix} 5 & 4 & 2 \\ 4 & 5 & -2 \\ 2 & -2 & 8 \end{bmatrix}$$

$$=\frac{1}{9}\begin{bmatrix} 4+5y & -4+4y & -2+2y \\ -4+4y & 4+5y & 2-2y \\ -2+2y & 2-2y & 1+8y \end{bmatrix} \quad (y=e^{9t})$$

(2) 对角形法：A 是实对称矩阵，存在可逆矩阵 P 使得 $P^{-1}AP=\Lambda$，其中

$$\Lambda=\begin{bmatrix} 0 & & \\ & 9 & \\ & & 9 \end{bmatrix}, \quad P=\begin{bmatrix} -2 & 1 & 1 \\ 2 & 1 & -1 \\ 1 & 0 & 4 \end{bmatrix}, \quad P^{-1}=\frac{1}{18}\begin{bmatrix} -4 & 4 & 2 \\ 9 & 9 & 0 \\ 1 & -1 & 4 \end{bmatrix}$$

故 $$e^{tA}=P\begin{bmatrix} 1 & & \\ & e^{9t} & \\ & & e^{9t} \end{bmatrix}P^{-1}=P\begin{bmatrix} 1 & & \\ & 0 & \\ & & 0 \end{bmatrix}P^{-1}+e^{9t}P\begin{bmatrix} 0 & & \\ & 1 & \\ & & 1 \end{bmatrix}P^{-1}$$

$$=\frac{1}{9}\begin{bmatrix} 4 & -4 & -2 \\ -4 & 4 & 2 \\ -2 & 2 & 1 \end{bmatrix}+\frac{e^{9t}}{9}\begin{bmatrix} 5 & 4 & 2 \\ 4 & 5 & -2 \\ 2 & -2 & 8 \end{bmatrix}$$

(3) 级数求和法：$|\lambda I-A|=\lambda(\lambda-9)^2$，因为 $A(A-9I)=O$，所以 A 的最小多项式为 $m(\lambda)=\lambda(\lambda-9)$．由 $m(A)=O$ 知 $A^2=9A$，于是可得

$$e^{tA}=I+\frac{(tA)}{1!}+\frac{(tA)^2}{2!}+\frac{(tA)^3}{3!}+\cdots=I+\left(\frac{t}{1!}+\frac{9t^2}{2!}+\frac{9^2t^3}{3!}+\cdots\right)A$$

$$=I+\frac{1}{9}\left(\frac{9t}{1!}+\frac{(9t)^2}{2!}+\frac{(9t)^3}{3!}+\cdots\right)A$$

$$=I+\frac{1}{9}\left(e^{9t}-1\right)A=\frac{1}{9}(9I-A)+\frac{e^{9t}}{9}A$$

$$=\frac{1}{9}\begin{bmatrix} 4 & -4 & -2 \\ -4 & 4 & 2 \\ -2 & 2 & 1 \end{bmatrix}+\frac{e^{9t}}{9}\begin{bmatrix} 5 & 4 & 2 \\ 4 & 5 & -2 \\ 2 & -2 & 8 \end{bmatrix}$$

$$e^{-\tau A}b(\tau)=\left\{\frac{1}{9}\begin{bmatrix}4&-4&-2\\-4&4&2\\-2&2&1\end{bmatrix}+\frac{e^{-9\tau}}{9}\begin{bmatrix}5&4&2\\4&5&-2\\2&-2&8\end{bmatrix}\right\}\begin{bmatrix}e^{9\tau}\\e^{9\tau}\\0\end{bmatrix}=\begin{bmatrix}1\\1\\0\end{bmatrix}$$

$$x(t)=e^{tA}\left\{x(0)+\int_0^t e^{-\tau A}b(\tau)d\tau\right\}$$

$$=\left\{\frac{1}{9}\begin{bmatrix}4&-4&-2\\-4&4&2\\-2&2&1\end{bmatrix}+\frac{e^{9t}}{9}\begin{bmatrix}5&4&2\\4&5&-2\\2&-2&8\end{bmatrix}\right\}\cdot\left\{\begin{bmatrix}1\\0\\2\end{bmatrix}+\begin{bmatrix}t\\t\\0\end{bmatrix}\right\}=e^{9t}\begin{bmatrix}1+t\\t\\2\end{bmatrix}$$

第四章　矩阵分解

§4.1 三角分解

目的：将 $A_{n\times n}$ 分解为下三角矩阵与上三角矩阵的乘积.

一、分解原理：以 $n=4$ 为例.

① $\Delta_1(A)=a_{11}$： $a_{11}\neq 0 \Rightarrow c_{i1}=\dfrac{a_{i1}}{a_{11}}$ $(i=2,3,4)$

$$L_1=\begin{bmatrix}1 & & & \\ c_{21} & 1 & & \\ c_{31} & 0 & 1 & \\ c_{41} & 0 & 0 & 1\end{bmatrix},\quad L_1^{-1}=\begin{bmatrix}1 & & & \\ -c_{21} & 1 & & \\ -c_{31} & 0 & 1 & \\ -c_{41} & 0 & 0 & 1\end{bmatrix}$$

$$L_1^{-1}A=\left[\begin{array}{c|ccc} a_{11} & a_{12} & a_{13} & a_{14} \\ \hline & a_{22}^{(1)} & a_{23}^{(1)} & a_{24}^{(1)} \\ & a_{32}^{(1)} & a_{33}^{(1)} & a_{34}^{(1)} \\ & a_{42}^{(1)} & a_{43}^{(1)} & a_{44}^{(1)}\end{array}\right]\overset{\Delta}{=}A^{(1)}$$

② $\Delta_2(A)=\Delta_2\left(A^{(1)}\right)=a_{11}a_{22}^{(1)}$： $a_{22}^{(1)}\neq 0 \Rightarrow c_{i2}=\dfrac{a_{i2}^{(1)}}{a_{22}^{(1)}}$ $(i=3,4)$

$$L_2=\begin{bmatrix}1 & & & \\ 0 & 1 & & \\ 0 & c_{32} & 1 & \\ 0 & c_{42} & 0 & 1\end{bmatrix},\quad L_2^{-1}=\begin{bmatrix}1 & & & \\ 0 & 1 & & \\ 0 & -c_{32} & 1 & \\ 0 & -c_{42} & 0 & 1\end{bmatrix}$$

$$L_2^{-1}A^{(1)}=\left[\begin{array}{cc|cc} a_{11} & a_{12} & a_{13} & a_{14} \\ & a_{22}^{(1)} & a_{23}^{(1)} & a_{24}^{(1)} \\ \hline & & a_{33}^{(2)} & a_{34}^{(2)} \\ & & a_{43}^{(2)} & a_{44}^{(2)}\end{array}\right]\overset{\Delta}{=}A^{(2)}$$

③ $\Delta_3(A)=\Delta_3\left(A^{(2)}\right)=a_{11}a_{22}^{(1)}a_{33}^{(2)}$： $a_{33}^{(2)}\neq 0 \Rightarrow c_{43}=\dfrac{a_{43}^{(2)}}{a_{33}^{(2)}}$

$$L_3=\begin{bmatrix}1&&&\\0&1&&\\0&0&1&\\0&0&c_{43}&1\end{bmatrix},\quad L_3^{-1}=\begin{bmatrix}1&&&\\0&1&&\\0&0&1&\\0&0&-c_{43}&1\end{bmatrix}$$

$$L_3^{-1}A^{(2)}=\begin{bmatrix}a_{11}&a_{12}&a_{13}&a_{14}\\&a_{22}^{(1)}&a_{23}^{(1)}&a_{24}^{(1)}\\&&a_{33}^{(2)}&a_{34}^{(2)}\\&&&a_{44}^{(3)}\end{bmatrix}\overset{\Delta}{=}A^{(3)}$$

即 $L_3^{-1}L_2^{-1}L_1^{-1}A=A^{(3)}\Rightarrow A=L_1L_2L_3A^{(3)}$

令 $L\overset{\Delta}{=}L_1L_2L_3=\begin{bmatrix}1&&&\\c_{21}&1&&\\c_{31}&c_{32}&1&\\c_{41}&c_{42}&c_{43}&1\end{bmatrix}$，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 $A=LDU$.

Th1 $A_{n\times n},\ \Delta_k(A)\neq 0\ (k=1,2,\cdots,n-1)\Rightarrow A=LDU$ 唯一.

二、紧凑格式算法：$A=LDU\overset{\Delta}{=}\tilde{L}U$ （Crout 分解）

$$\tilde{L}=\begin{bmatrix}l_{11}&&&\\l_{21}&l_{22}&&\\\vdots&\vdots&\ddots&\\l_{n1}&l_{n2}&\cdots&l_{nn}\end{bmatrix},\quad U=\begin{bmatrix}1&u_{12}&\cdots&u_{1n}\\&1&\cdots&u_{2n}\\&&\ddots&\vdots\\&&&1\end{bmatrix}$$

$(i,1)$元：$a_{i1}=l_{i1}\cdot 1\Rightarrow l_{i1}=a_{i1}\quad (i=1,\cdots,n)$

$(1,j)$元：$a_{1j}=l_{11}\cdot u_{1j}\Rightarrow u_{1j}=\dfrac{a_{1j}}{l_{11}}\quad (j=2,\cdots,n)$

(i,k)元：$a_{ik}=l_{i1}\cdot u_{1k}+\cdots+l_{i,k-1}\cdot u_{k-1,k}+l_{ik}\cdot 1\quad (i\geq k)$

$$\Rightarrow l_{ik}=a_{ik}-\left(l_{i1}\cdot u_{1k}+\cdots+l_{i,k-1}\cdot u_{k-1,k}\right)$$

(k,j)元： $a_{kj}=l_{k1}\cdot u_{1j}+\cdots+l_{k,k-1}\cdot u_{k-1,j}+l_{kk}\cdot u_{kj}\quad (j>k)$

$$\Rightarrow u_{kj}=\frac{1}{l_{kk}}\left[a_{kj}-\left(l_{k1}\cdot u_{1j}+\cdots+l_{k,k-1}\cdot u_{k-1,j}\right)\right]$$

计算框图：
$$\begin{array}{|c|c|c|c|c} l_{11} & u_{12} & u_{13} & u_{14} & \cdots \\ l_{21} & l_{22} & u_{23} & u_{24} & \cdots \\ l_{31} & l_{32} & l_{33} & u_{34} & \cdots \\ l_{41} & l_{42} & l_{43} & l_{44} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{array}\qquad \begin{array}{l} 第1框 \\ 第2框 \\ 第3框 \\ 第4框 \\ \vdots \end{array}$$

例 1 $A=\begin{bmatrix} 5 & 2 & -4 & 0 \\ 2 & 1 & -2 & 1 \\ -4 & -2 & 5 & 0 \\ 0 & 1 & 0 & 2 \end{bmatrix}$， 计算框图：
$$\begin{array}{|c|c|c|c} 5 & 2/5 & -4/5 & 0 \\ 2 & 1/5 & -2 & 5 \\ -4 & -2/5 & 1 & 2 \\ 0 & 1 & 2 & -7 \end{array}$$

$$\tilde{L}=\begin{bmatrix} 5 & 0 & 0 & 0 \\ 2 & 1/5 & 0 & 0 \\ -4 & -2/5 & 1 & 0 \\ 0 & 1 & 2 & -7 \end{bmatrix}=\begin{bmatrix} 1 & & & \\ 2/5 & 1 & & \\ -4/5 & -2 & 1 & \\ 0 & 5 & 2 & 1 \end{bmatrix}\begin{bmatrix} 5 & & & \\ & 1/5 & & \\ & & 1 & \\ & & & -7 \end{bmatrix}$$

$$U=\begin{bmatrix} 1 & 2/5 & -4/5 & 0 \\ 0 & 1 & -2 & 5 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix},\quad A=\tilde{L}U=LDU$$

§4.2 QR 分解

目的：将 $A_{n\times n}$ 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

$$T_{ij}(c,s)\overset{\Delta}{=}\begin{bmatrix} I & & & & \\ & c & & s & \\ & & I & & \\ & -s & & c & \\ & & & & I \end{bmatrix}\begin{array}{c} \\ (i) \\ \\ (j) \\ \\ \end{array}\qquad (c^2+s^2=1)$$

1. $T_{ij}^{\mathrm{T}} T_{ij} = I,\quad \left[T_{ij}(c,s)\right]^{-1} = \left[T_{ij}(c,s)\right]^{\mathrm{T}} = T_{ij}(c,-s),\quad \det T_{ij} = 1$.

2. $x = \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix},\quad T_{ij}x \overset{\Delta}{=} \begin{bmatrix} \eta_1 \\ \vdots \\ \eta_n \end{bmatrix} \Rightarrow \begin{cases} \eta_i = c\xi_i + s\xi_j \\ \eta_j = -s\xi_i + c\xi_j \\ \eta_k = \xi_k\,(k \neq i,j) \end{cases}$

若 $\xi_i^2 + \xi_j^2 \neq 0$，取 $c = \dfrac{\xi_i}{\left(\xi_i^2 + \xi_j^2\right)^{\frac{1}{2}}},\quad s = \dfrac{\xi_j}{\left(\xi_i^2 + \xi_j^2\right)^{\frac{1}{2}}}$

则 $\eta_i = \left(\xi_i^2 + \xi_j^2\right)^{\frac{1}{2}} > 0,\ \eta_j = 0$.

Th3　$x \neq 0 \Rightarrow \exists$ 有限个 G-矩阵之积 T, st. $Tx = |x| e_1$.

证 ① $\xi_1 \neq 0$

$T_{12}(c,s)$ 中，$c = \dfrac{\xi_1}{\left(\xi_1^2 + \xi_2^2\right)^{\frac{1}{2}}},\quad s = \dfrac{\xi_2}{\left(\xi_1^2 + \xi_2^2\right)^{\frac{1}{2}}}$.

$$T_{12}x = \begin{bmatrix} \left(\xi_1^2 + \xi_2^2\right)^{\frac{1}{2}} \\ 0 \\ \xi_3 \\ \vdots \\ \xi_n \end{bmatrix} \overset{\Delta}{=} x^{(2)}$$

… …

$T_{1n}(c,s)$ 中，$c = \dfrac{\left(\xi_1^2 + \cdots + \xi_{n-1}^2\right)^{\frac{1}{2}}}{\left(\xi_1^2 + \cdots + \xi_n^2\right)^{\frac{1}{2}}},\ s = \dfrac{\xi_n}{\left(\xi_1^2 + \cdots + \xi_n^2\right)^{\frac{1}{2}}}$.

$$T_{1n}x^{(n-1)} = |x| e_1$$

$$T_{1n}(T_{1,n-1} \cdots T_{12}x) = |x| e_1$$

② $\xi_1 = \cdots = \xi_{k-1} = 0, \xi_k \neq 0\ (1 < k \leq n)$:　$|x| = \left(\xi_k^2 + \cdots + \xi_n^2\right)^{\frac{1}{2}}$

由 T_{1k} 开始即可.

推论　$\mathbf{R}^n$ 中，$\forall x \neq 0, \forall$ 单位向量 $z \Rightarrow \exists$ 有限个 G-矩阵之积 T, st. $Tx = |x| z$.

证　$T^{(1)}x = |x| e_1,\ T^{(2)}z = |z| e_1 \Rightarrow \left[T^{(2)}\right]^{-1} T^{(1)}x = |x| z$

$$T\overset{\Delta}{=}\left[T^{(2)}\right]^{-1}T^{(1)}=\left[T_{1n}^{(2)}\cdots T_{12}^{(2)}\right]^{-1}T^{(1)}=\left[\left(T_{12}^{(2)}\right)^{\mathrm{T}}\cdots\left(T_{1n}^{(2)}\right)^{\mathrm{T}}\right]\left[T_{1n}^{(1)}\cdots T_{12}^{(1)}\right]$$

例 1　$x=\begin{bmatrix}3\\4\\5\end{bmatrix}$，求 G-矩阵之积 T 使得 $Tx=|x|e_1$.

解　$T_{12}(c,s)$ 中，$c=\dfrac{3}{5}, s=\dfrac{4}{5}$.　$T_{12}x=\begin{bmatrix}5\\0\\5\end{bmatrix}$

$T_{13}(c,s)$ 中，$c=\dfrac{1}{\sqrt{2}}, s=\dfrac{1}{\sqrt{2}}$.　$T_{13}(T_{12}x)=\begin{bmatrix}5\sqrt{2}\\0\\0\end{bmatrix}=|x|e_1$

$$T=T_{13}T_{12}=\frac{1}{\sqrt{2}}\begin{bmatrix}1&0&1\\0&\sqrt{2}&0\\-1&0&1\end{bmatrix}\cdot\frac{1}{5}\begin{bmatrix}3&4&0\\-4&3&0\\0&0&5\end{bmatrix}=\frac{1}{5\sqrt{2}}\begin{bmatrix}3&4&5\\-4\sqrt{2}&3\sqrt{2}&0\\-3&-4&5\end{bmatrix}$$

$Tx=5\sqrt{2}e_1$

二、Householder 矩阵

$H_u\overset{\Delta}{=}I_n-2uu^{\mathrm{T}}$　（$u\in\mathbf{R}^n$ 是单位列向量）

(1)　$H^{\mathrm{T}}=H$　对称　　(2)　$H^{\mathrm{T}}H=I$　正交

(3)　$H^2=I$　对合　　(4)　$H^{-1}=H$　自逆　　(5)　$\det H=-1$

验证(5):

$$\begin{bmatrix}I&0\\-u^{\mathrm{T}}&1\end{bmatrix}\begin{bmatrix}I&2u\\0^{\mathrm{T}}&1\end{bmatrix}\begin{bmatrix}I-2uu^{\mathrm{T}}&0\\u^{\mathrm{T}}&1\end{bmatrix}=\begin{bmatrix}I&0\\-u^{\mathrm{T}}&1\end{bmatrix}\begin{bmatrix}I&2u\\u^{\mathrm{T}}&1\end{bmatrix}=\begin{bmatrix}I&2u\\0^{\mathrm{T}}&-1\end{bmatrix}$$

$$\begin{vmatrix}I-2uu^{\mathrm{T}}&0\\u^{\mathrm{T}}&1\end{vmatrix}=\begin{vmatrix}I&2u\\0^{\mathrm{T}}&-1\end{vmatrix}\Rightarrow\det(I-2uu^{\mathrm{T}})=-1$$

Th4　$\mathbf{R}^n$ 中 $(n>1)$, $\forall x\neq 0$, $\forall$ 单位向量 $z\Rightarrow\exists H_u$, st $H_ux=|x|z$.

证　①　$x=|x|z$:　$n>1$ 时，可取单位向量 u 使得　$u\perp x$，于是

$$H_u=I-2uu^{\mathrm{T}}:\quad H_ux=Ix-2uu^{\mathrm{T}}x=x=|x|z$$

② $x \neq |x|z$：取 $u = \dfrac{x-|x|z}{|x-|x|z|}$，有

$$H_u x = \left[I - 2\frac{(x-|x|z)(x-|x|z)^{\mathrm{T}}}{|x-|x|z|^2}\right]x = x - \frac{2(x-|x|z, x)}{|x-|x|z|^2}(x-|x|z)$$

$$= x - 1\cdot(x-|x|z) = |x|z$$

例 2　$x = \begin{bmatrix}1\\2\\2\end{bmatrix}$，求 H-矩阵 H 使得 $Hx = |x|e_1$．

解　$|x| = 3,\ x-|x|e_1 = \begin{bmatrix}-2\\2\\2\end{bmatrix},\ u = \dfrac{1}{\sqrt{3}}\begin{bmatrix}-1\\1\\1\end{bmatrix}$

$$H = I - \frac{2}{3}\begin{bmatrix}-1\\1\\1\end{bmatrix}\begin{pmatrix}-1 & 1 & 1\end{pmatrix} = \frac{1}{3}\begin{bmatrix}1 & 2 & 2\\2 & 1 & -2\\2 & -2 & 1\end{bmatrix}:\quad Hx = 3e_1$$

三、G-矩阵与 H-矩阵的关系

Th5　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　$c^2+s^2 = 1 \Rightarrow$ 取 $\theta = \mathrm{arctg}\dfrac{s}{c}$，则 $\cos\theta = c, \sin\theta = s$．

$$T_{ij}(c,s) = \begin{bmatrix}I & & & & \\ & \cos\theta & & \sin\theta & \\ & & I & & \\ & -\sin\theta & & \cos\theta & \\ & & & & I\end{bmatrix}\begin{array}{l} \\ (i) \\ \\ (j) \\ \\ \end{array}$$

$$v = \begin{pmatrix}0 & \cdots & 0 & \sin\dfrac{\theta}{4} & 0 & \cdots & 0 & \cos\dfrac{\theta}{4} & 0 & \cdots & 0\end{pmatrix}^{\mathrm{T}}$$

$$H_v=\begin{bmatrix} I &&&& \\ & 1 &&& \\ && I && \\ &&& 1 & \\ &&&& I \end{bmatrix}-2\begin{bmatrix} O &&&& \\ & \sin^2\frac{\theta}{4} && \sin\frac{\theta}{4}\cos\frac{\theta}{4} & \\ && O && \\ & \cos\frac{\theta}{4}\sin\frac{\theta}{4} && \cos^2\frac{\theta}{4} & \\ &&&& O \end{bmatrix}$$

$$=\begin{bmatrix} I &&&& \\ & \cos\frac{\theta}{2} && -\sin\frac{\theta}{2} & \\ && I && \\ & -\sin\frac{\theta}{2} && -\cos\frac{\theta}{2} & \\ &&&& I \end{bmatrix}$$

$$u=\begin{pmatrix}0 & \cdots & 0 & \sin\frac{3\theta}{4} & 0 & \cdots & 0 & \cos\frac{3\theta}{4} & 0 & \cdots & 0\end{pmatrix}^{\mathrm{T}}$$

$$H_u=\begin{bmatrix} I &&&& \\ & \cos\frac{3\theta}{2} && -\sin\frac{3\theta}{2} & \\ && I && \\ & -\sin\frac{3\theta}{2} && -\cos\frac{3\theta}{2} & \\ &&&& I \end{bmatrix},\quad T_{ij}(c,s)=H_uH_v$$

[注]　H-矩阵不能由若干个 G-矩阵的乘积表示.

因为 $\det H=-1$，而 $\det T_{ij}=1$.

例 3　G-矩阵 $T_{12}(0,1)=\begin{bmatrix}0 & 1\\ -1 & 0\end{bmatrix}$中，$c=0,\ s=1\Rightarrow\theta=\dfrac{\pi}{2}$

$$H_u=\frac{1}{\sqrt{2}}\begin{bmatrix}-1 & -1\\ -1 & 1\end{bmatrix},H_v=\frac{1}{\sqrt{2}}\begin{bmatrix}1 & -1\\ -1 & -1\end{bmatrix}\Rightarrow H_uH_v=\begin{bmatrix}0 & 1\\ -1 & 0\end{bmatrix}$$

四、QR 分解

1．Schmidt 正交化方法

Th6　$A_{n\times n}$可逆$\Rightarrow\exists$正交矩阵 Q，可逆上三角矩阵 R，使得 $A=QR$.

证　$A=(a_1,a_2,\cdots,a_n)$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　$b_1 = a_1 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$，$b_2 = a_2 - 1 \cdot b_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$，$b_3 = a_3 - \dfrac{1}{3}b_2 - \dfrac{7}{6}b_1 = \begin{bmatrix} 1/2 \\ 0 \\ -1/2 \end{bmatrix}$

$$Q = \begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} \\ \frac{2}{\sqrt{6}} & -\frac{1}{\sqrt{3}} & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} \end{bmatrix}$$

$$R = \begin{bmatrix} \sqrt{6} & & \\ & \sqrt{3} & \\ & & \frac{1}{\sqrt{2}} \end{bmatrix}\begin{bmatrix} 1 & 1 & \frac{7}{6} \\ & 1 & \frac{1}{3} \\ & & 1 \end{bmatrix} = \begin{bmatrix} \sqrt{6} & \sqrt{6} & \frac{7}{\sqrt{6}} \\ & \sqrt{3} & \frac{1}{\sqrt{3}} \\ & & \frac{1}{\sqrt{2}} \end{bmatrix}$$

Th7 $A_{m\times n}$ 列满秩 $\Rightarrow \exists$ 矩阵 $Q_{m\times n}$ 满足 $Q^{\mathrm{H}}Q = I$，可逆上三角矩阵 $R_{n\times n}$，

使得 $A = QR$.（证明过程同 Th6）

2．G-变换方法

Th8 $A_{n\times n}$ 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 以 $n = 4$ 为例.

(1) $|A| \neq 0$： $\beta^{(0)} = \begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \\ a_{41} \end{bmatrix} \neq 0 \Rightarrow \exists$ 有限个 G-矩阵之积 T_0，使得

$$T_0\beta^{(0)} = \begin{bmatrix} |\beta^{(0)}| \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad a_{11}^{(1)} = |\beta^{(0)}| > 0$$

$$T_0A = \left[\begin{array}{c:ccc} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} & a_{14}^{(1)} \\ \hdashline 0 & & & \\ 0 & & A^{(1)} & \\ 0 & & & \end{array}\right], \quad A^{(1)} = \begin{bmatrix} a_{22}^{(1)} & a_{23}^{(1)} & a_{24}^{(1)} \\ a_{32}^{(1)} & a_{33}^{(1)} & a_{34}^{(1)} \\ a_{42}^{(1)} & a_{43}^{(1)} & a_{44}^{(1)} \end{bmatrix}$$

(2) $|A^{(1)}| \neq 0$： $\beta^{(1)} = \begin{bmatrix} a_{22}^{(1)} \\ a_{32}^{(1)} \\ a_{42}^{(1)} \end{bmatrix} \neq 0 \Rightarrow \exists$ 有限个 G-矩阵之积 T_1，使得

$$T_1\beta^{(1)} = \begin{bmatrix} |\beta^{(1)}| \\ 0 \\ 0 \end{bmatrix}, \quad a_{22}^{(2)} = |\beta^{(1)}| > 0$$

$$T_1A^{(1)} = \left[\begin{array}{c:cc} a_{22}^{(2)} & a_{23}^{(2)} & a_{24}^{(2)} \\ \hdashline 0 & a_{33}^{(2)} & a_{34}^{(2)} \\ 0 & a_{43}^{(2)} & a_{44}^{(2)} \end{array}\right], \quad A^{(2)} = \begin{bmatrix} a_{33}^{(2)} & a_{34}^{(2)} \\ a_{43}^{(2)} & a_{44}^{(2)} \end{bmatrix}$$

(3) $\left|A^{(2)}\right| \neq 0$: $\beta^{(2)} = \begin{bmatrix} a_{33}^{(2)} \\ a_{43}^{(2)} \end{bmatrix} \neq 0 \Rightarrow$ G-矩阵 T_2，使得

$$T_2\beta^{(2)} = \begin{bmatrix} \left|\beta^{(2)}\right| \\ 0 \end{bmatrix},\ a_{33}^{(3)} = \left|\beta^{(2)}\right| > 0,\ T_2 A^{(2)} = \begin{bmatrix} a_{33}^{(3)} & a_{34}^{(3)} \\ 0 & a_{44}^{(3)} \end{bmatrix}$$

令 $$T = \begin{bmatrix} I_2 & \\ & T_2 \end{bmatrix} \cdot \begin{bmatrix} 1 & \\ & T_1 \end{bmatrix} \cdot T_0$$

则 T 为有限个 G-矩阵之积, 且有

$$TA = \begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} & a_{14}^{(1)} \\ & a_{22}^{(2)} & a_{23}^{(2)} & a_{24}^{(2)} \\ & & a_{33}^{(3)} & a_{34}^{(3)} \\ & & & a_{44}^{(3)} \end{bmatrix} \stackrel{\Delta}{=} R$$

[注] $\det T = 1 \Rightarrow \det A = a_{11}^{(1)} a_{22}^{(2)} \cdots a_{n-1,n-1}^{(n-1)} a_{nn}^{(n-1)}$，故 $a_{nn}^{(n-1)}$ 与 $\det A$ 同符号.

在 Th8 中，当 $A_{n\times n}$ 不可逆时，仍可得 $TA = R$，但 R 是不可逆矩阵.

例 5 用 G-变换求 $A = \begin{bmatrix} 3 & 5 & 5 \\ 0 & 3 & 4 \\ 4 & 0 & 5 \end{bmatrix}$ 的 QR 分解.

解 (1) $\beta^{(0)} = \begin{bmatrix} 3 \\ 0 \\ 4 \end{bmatrix}$: $T_{13}(c,s)$ 中 $c = \dfrac{3}{5},\ s = \dfrac{4}{5}$. $T_{13}\beta^{(0)} = \begin{bmatrix} 5 \\ 0 \\ 0 \end{bmatrix}$

$$T_0 \stackrel{\Delta}{=} T_{13} = \begin{bmatrix} 3/5 & 0 & 4/5 \\ 0 & 1 & 0 \\ -4/5 & 0 & 3/5 \end{bmatrix},\quad T_0 A = \left[\begin{array}{c:cc} 5 & 3 & 7 \\ \hdashline 0 & 3 & 4 \\ 0 & -4 & -1 \end{array}\right]$$

(2) $A^{(1)} = \begin{bmatrix} 3 & 4 \\ -4 & -1 \end{bmatrix}, \beta^{(1)} = \begin{bmatrix} 3 \\ -4 \end{bmatrix}$

$T_{12}(c,s)$ 中 $c = \dfrac{3}{5},\ s = -\dfrac{4}{5}$. $T_{12}\beta^{(1)} = \begin{bmatrix} 5 \\ 0 \end{bmatrix}$

$$T_1 \stackrel{\Delta}{=} T_{12} = \begin{bmatrix} 3/5 & -4/5 \\ 4/5 & 3/5 \end{bmatrix},\quad T_1 A^{(1)} = \begin{bmatrix} 5 & 16/5 \\ 0 & 13/5 \end{bmatrix}$$

令 $T=\begin{bmatrix}1 & \\ & T_1\end{bmatrix}T_0=\dfrac{1}{5}\left[\begin{array}{c|cc}5 & & \\ \hline & 3 & -4 \\ & 4 & 3\end{array}\right]\cdot\dfrac{1}{5}\begin{bmatrix}3 & 0 & 4 \\ 0 & 5 & 0 \\ -4 & 0 & 3\end{bmatrix}=\dfrac{1}{25}\begin{bmatrix}15 & 0 & 20 \\ 16 & 15 & -12 \\ -12 & 20 & 9\end{bmatrix}$

则 $Q=T^{-1}=T^{\mathrm{T}}$，$R=\begin{bmatrix}5 & 3 & 7 \\ & 5 & 16/5 \\ & & 13/5\end{bmatrix}$：$A=QR$

例 6　用 Givens 变换求 $A=\begin{bmatrix}0 & 0 & 1 & 3 \\ 0 & 4 & 2 & 3 \\ 1 & 1 & 2 & 4 \\ 0 & 0 & 1 & 2\end{bmatrix}$ 的 QR 分解.

解　(1) $\beta^{(0)}=\begin{bmatrix}0 \\ 0 \\ 1 \\ 0\end{bmatrix}$，构造 $T_{13}(c,s), c=0, s=1$，则

$$T_{13}\beta^{(0)}=\begin{bmatrix}1 \\ 0 \\ 0 \\ 0\end{bmatrix},\quad T_0=T_{13}=\begin{bmatrix}0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1\end{bmatrix},\quad T_0A=\left[\begin{array}{cc|cc}1 & 1 & 2 & 4 \\ 0 & 4 & 3 & 2 \\ \hline 0 & 0 & -1 & -3 \\ 0 & 0 & 1 & 2\end{array}\right]$$

(2) $A^{(2)}=\begin{bmatrix}-1 & -3 \\ 1 & 2\end{bmatrix}$，$\beta^{(2)}=\begin{bmatrix}-1 \\ 1\end{bmatrix}$，构造 $T_{12}(c,s), c=\dfrac{-1}{\sqrt{2}}, s=\dfrac{1}{\sqrt{2}}$，则

$$T_{12}\beta^{(2)}=\begin{bmatrix}\sqrt{2} \\ 0\end{bmatrix},\quad T_2=T_{12}=\frac{1}{\sqrt{2}}\begin{bmatrix}-1 & 1 \\ -1 & -1\end{bmatrix},\quad T_2A^{(2)}=\frac{1}{\sqrt{2}}\begin{bmatrix}2 & 5 \\ 0 & 1\end{bmatrix}$$

令 $T=\begin{bmatrix}I_2 & O \\ O & T_2\end{bmatrix}T_0=\dfrac{1}{\sqrt{2}}\left[\begin{array}{cc|cc}\sqrt{2} & 0 & & \\ 0 & \sqrt{2} & & \\ \hline & & -1 & 1 \\ & & -1 & -1\end{array}\right]\cdot\begin{bmatrix}0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1\end{bmatrix}$

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

则 $Q=T^{\mathrm{T}}=\dfrac{1}{\sqrt{2}}\begin{bmatrix}0&0&1&1\\0&\sqrt{2}&0&0\\\sqrt{2}&0&0&0\\0&0&1&-1\end{bmatrix}$, $R=\begin{bmatrix}1&1&2&4\\&4&3&2\\&&\sqrt{2}&5/\sqrt{2}\\&&&1/\sqrt{2}\end{bmatrix}$

$A=QR$

3. H-变换方法:

Th10 $A_{n\times n}$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　以$n=4$为例.

(1) $|A|\neq 0$: $\beta^{(0)}=\begin{bmatrix}a_{11}\\a_{21}\\a_{31}\\a_{41}\end{bmatrix}\neq 0\Rightarrow\exists$H-矩阵$H_0$，使得$H_0\beta^{(0)}=\begin{bmatrix}|\beta^{(0)}|\\0\\0\\0\end{bmatrix}$.

$a_{11}^{(1)}=|\beta^{(0)}|>0$

$$H_0A=\left[\begin{array}{c:ccc}a_{11}^{(1)}&a_{12}^{(1)}&a_{13}^{(1)}&a_{14}^{(1)}\\\hdashline 0&&&\\0&&A^{(1)}&\\0&&&\end{array}\right],\quad A^{(1)}=\begin{bmatrix}a_{22}^{(1)}&a_{23}^{(1)}&a_{24}^{(1)}\\a_{32}^{(1)}&a_{33}^{(1)}&a_{34}^{(1)}\\a_{42}^{(1)}&a_{43}^{(1)}&a_{44}^{(1)}\end{bmatrix}$$

(2) $|A^{(1)}|\neq 0$: $\beta^{(1)}=\begin{bmatrix}a_{22}^{(1)}\\a_{32}^{(1)}\\a_{42}^{(1)}\end{bmatrix}\neq 0\Rightarrow\exists$H-矩阵$H_1$，使得$H_1\beta^{(1)}=\begin{bmatrix}|\beta^{(1)}|\\0\\0\end{bmatrix}$.

$a_{22}^{(2)}=|\beta^{(1)}|>0$

$$H_1A^{(1)}=\left[\begin{array}{c:cc}a_{22}^{(2)}&a_{23}^{(2)}&a_{24}^{(2)}\\\hdashline 0&a_{33}^{(2)}&a_{34}^{(2)}\\0&a_{43}^{(2)}&a_{44}^{(2)}\end{array}\right],\quad A^{(2)}=\begin{bmatrix}a_{33}^{(2)}&a_{34}^{(2)}\\a_{43}^{(2)}&a_{44}^{(2)}\end{bmatrix}$$

(3) $|A^{(2)}|\neq 0$: $\beta^{(2)}=\begin{bmatrix}a_{33}^{(2)}\\a_{43}^{(2)}\end{bmatrix}\neq 0\Rightarrow\exists$H-矩阵$H_2$，使得$H_2\beta^{(2)}=\begin{bmatrix}|\beta^{(2)}|\\0\end{bmatrix}$.

$$a_{33}^{(3)}=\left|\beta^{(2)}\right|>0,\quad H_2A^{(2)}=\begin{bmatrix} a_{33}^{(3)} & a_{34}^{(3)} \\ 0 & a_{44}^{(3)} \end{bmatrix}$$

令 $S=\begin{bmatrix} I_2 & \\ & H_2 \end{bmatrix}\cdot\begin{bmatrix} 1 & \\ & H_1 \end{bmatrix}\cdot H_0$

则 S 为有限个 H-矩阵之积, 且有

$$SA=\begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} & a_{14}^{(1)} \\ & a_{22}^{(2)} & a_{23}^{(2)} & a_{24}^{(2)} \\ & & a_{33}^{(3)} & a_{34}^{(3)} \\ & & & a_{44}^{(3)} \end{bmatrix}\overset{\Delta}{=}R$$

[注] 设 $H_l=I_{n-l}-2u_{n-l}u_{n-l}^{\mathrm{T}}\quad\left(u_{n-l}^{\mathrm{T}}u_{n-l}=1\right)$, 则

$$\begin{bmatrix} I_l & \\ & H_l \end{bmatrix}=\begin{bmatrix} I_l & \\ & I_{n-l} \end{bmatrix}-2\begin{bmatrix} O & O \\ O & u_{n-l}u_{n-l}^{\mathrm{T}} \end{bmatrix}=I_n-2\begin{bmatrix} 0 \\ \vdots \\ 0 \\ u_{n-l} \end{bmatrix}\begin{pmatrix} 0 & \cdots & 0 & u_{n-l}^{\mathrm{T}} \end{pmatrix}$$

$$=I_n-2u_nu_n^{\mathrm{T}}\quad\left(u_n^{\mathrm{T}}u_n=u_{n-l}^{\mathrm{T}}u_{n-l}=1\right)$$

故$\begin{bmatrix} I_l & \\ & H_l \end{bmatrix}$是 H-矩阵；在 Th10 中，当 $A_{n\times n}$ 不可逆时，仍可得 $SA=R$，

但 R 是不可逆矩阵.

例 7　用 H-变换求 $A=\begin{bmatrix} 3 & 14 & 9 \\ 6 & 43 & 3 \\ 6 & 22 & 15 \end{bmatrix}$ 的 QR 分解.

解 (1) $\beta^{(0)}=\begin{bmatrix} 3 \\ 6 \\ 6 \end{bmatrix}$, $\beta^{(0)}-\left|\beta^{(0)}\right|e_1=\begin{bmatrix} -6 \\ 6 \\ 6 \end{bmatrix}=6\begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$, $u=\dfrac{1}{\sqrt{3}}\begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$

$$H_0=I-2uu^{\mathrm{T}}=\frac{1}{3}\begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & -2 \\ 2 & -2 & 1 \end{bmatrix},\quad H_0A=\begin{bmatrix} 9 & 48 & 15 \\ 0 & 9 & -3 \\ 0 & -12 & 9 \end{bmatrix}$$

(2) $A^{(1)}=\begin{bmatrix} 9 & -3 \\ -12 & 9 \end{bmatrix}$, $\beta^{(1)}=\begin{bmatrix} 9 \\ -12 \end{bmatrix}$

$$\beta^{(1)}-\left|\beta^{(1)}\right|e_1=\begin{bmatrix}-6\\-12\end{bmatrix}=(-6)\begin{bmatrix}1\\2\end{bmatrix},\quad u=\frac{1}{\sqrt{5}}\begin{bmatrix}1\\2\end{bmatrix}$$

$$H_1=I-2uu^{\mathrm{T}}=\frac{1}{5}\begin{bmatrix}3&-4\\-4&-3\end{bmatrix},\quad H_1A^{(1)}=\begin{bmatrix}15&-9\\0&-3\end{bmatrix}$$

令 $S=\begin{bmatrix}1&\\&H_1\end{bmatrix}H_0$

则 $Q=S^{-1}=S^{\mathrm{T}}=H_0\begin{bmatrix}1&\\&H_1\end{bmatrix}=\dfrac{1}{15}\begin{bmatrix}5&-2&-14\\10&11&2\\10&-10&5\end{bmatrix}$

$$R=\begin{bmatrix}9&48&15\\&15&-9\\&&-3\end{bmatrix}:\quad A=QR$$

五、化方阵与 Hessenberg 矩阵相似

上 Hessenberg 矩阵：$F_{上}=\begin{bmatrix}a_{11}&a_{12}&a_{13}&\cdots&a_{1n}\\a_{21}&a_{22}&a_{23}&\cdots&a_{2n}\\&\ddots&\ddots&\ddots&\vdots\\&&a_{n-1,n-2}&a_{n-1,n-1}&a_{n-1,n}\\&&&a_{n,n-1}&a_{nn}\end{bmatrix}$

Th11　设 $A\in\mathbf{R}^{n\times n}$，则存在有限个 G-矩阵之积 Q，使得 $QAQ^{\mathrm{T}}=F_{上}$.

证　(1) 对 A：若 $\beta^{(0)}=\begin{bmatrix}a_{21}\\\vdots\\a_{n1}\end{bmatrix}\neq 0$，则存在有限个 G-矩阵之积 T_0，使得

$$T_0\beta^{(0)}=\left|\beta^{(0)}\right|e_1\overset{\Delta}{=}a_{21}^{(1)}e_1$$

$$\begin{bmatrix}1&\\&T_0\end{bmatrix}A\begin{bmatrix}1&\\&T_0\end{bmatrix}^{\mathrm{T}}=\left[\begin{array}{c:cccc}a_{11}&a_{12}^{(1)}&a_{13}^{(1)}&\cdots&a_{1n}^{(1)}\\\hdashline a_{21}^{(1)}&&&&\\0&&&&\\\vdots&&&A^{(1)}&\\0&&&&\end{array}\right]$$

若 $\beta^{(0)}=\theta$，转入(2)；

(2) 对 $A^{(1)}$：若 $\beta^{(1)}=\begin{bmatrix} a_{32}^{(1)} \\ \vdots \\ a_{n2}^{(1)} \end{bmatrix}\neq 0$，则存在有限个 G-矩阵之积 T_1，使得

$$T_1\beta^{(1)}=\left|\beta^{(1)}\right|e_1 \overset{\Delta}{=} a_{32}^{(2)}e_1$$

$$\begin{bmatrix} 1 & \\ & T_1 \end{bmatrix} A^{(1)} \begin{bmatrix} 1 & \\ & T_1 \end{bmatrix}^{\mathrm{T}} = \left[\begin{array}{c|cccc} a_{22}^{(1)} & a_{21}^{(2)} & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} \\ \hline a_{32}^{(2)} & & & & \\ 0 & & & & \\ \vdots & & & A^{(2)} & \\ 0 & & & & \end{array}\right]$$

若 $\beta^{(1)}=\theta$，转入(3)；

(3) 对 $A^{(2)}$：……. 进行 $n-2$ 步结束.

令 $Q=\begin{bmatrix} I_{n-2} & \\ & T_{n-3} \end{bmatrix}\cdots\begin{bmatrix} I_2 & \\ & T_1 \end{bmatrix}\begin{bmatrix} 1 & \\ & T_0 \end{bmatrix}$，则 $QAQ^{\mathrm{T}}=F_{上}$.

Th12　设 $A\in\mathbf{R}^{n\times n}$，则存在有限个 H-矩阵之积 Q，使得 $QAQ^{\mathrm{T}}=F_{上}$.

证　类似于 Th11 的证明.

推论　$A_{n\times n}$ 实对称 $\Rightarrow\exists$ 有限个 H-矩阵（G-矩阵）之积 Q，使得

$$QAQ^{\mathrm{T}}=\text{“实对称三对角矩阵”}$$

证　由 Th12 知，存在 $Q=H_{u_{n-2}}\cdots H_{u_1}$，使得 $QAQ^{\mathrm{T}}=F_{上}$. 于是有

$$A^{\mathrm{T}}=A\Rightarrow QAQ^{\mathrm{T}}=\left(F_{上}\right)^{\mathrm{T}}$$

即 $F_{上}=\left(F_{上}\right)^{\mathrm{T}}$，故 $F_{上}$ 是“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　$\beta^{(0)}=\begin{bmatrix} 0 \\ 1 \end{bmatrix}$：$\beta^{(0)}-\left|\beta^{(0)}\right|e_1=\begin{bmatrix} -1 \\ 1 \end{bmatrix}, u=\dfrac{1}{\sqrt{2}}\begin{bmatrix} -1 \\ 1 \end{bmatrix}$

$$H_0=\begin{bmatrix}0&1\\1&0\end{bmatrix},\quad Q\overset{\Delta}{=}\begin{bmatrix}1&\\&H_0\end{bmatrix}=\begin{bmatrix}1&0&0\\0&0&1\\0&1&0\end{bmatrix}$$

$$QA=\begin{bmatrix}1&0&1\\1&2&1\\0&1&2\end{bmatrix},\quad QAQ^{\mathrm{T}}=\begin{bmatrix}1&1&0\\1&1&2\\0&2&1\end{bmatrix}$$

§4.3 满秩分解

目的：对 $A\in\mathbf{C}_r^{m\times n}\ (r\geq 1)$，求 $F\in\mathbf{C}_r^{m\times r}$ 及 $G\in\mathbf{C}_r^{r\times n}$ 使 $A=FG$.

一、分解原理

$$\mathrm{rank}A=r\Rightarrow A\xrightarrow{\text{行}}\text{阶梯形 } B=\begin{pmatrix}G\\O\end{pmatrix}:\quad G\in\mathbf{C}_r^{r\times n}$$

$$\Rightarrow\exists\text{有限个初等矩阵之积 } P_{m\times m},\mathrm{st.}PA=B$$

$$\Rightarrow A=P^{-1}B=\left(F_{m\times r}\,\middle|\,S_{m\times(m-r)}\right)\begin{pmatrix}G\\O\end{pmatrix}=FG:\quad F\in\mathbf{C}_r^{m\times r}$$

例 1　$A=\begin{bmatrix}-1&0&1&2\\1&2&-1&1\\2&2&-2&-1\end{bmatrix}$，求 $A=FG$.

解(1)　$$(A\mid I)=\left[\begin{array}{cccc|ccc}-1&0&1&2&1&&\\1&2&-1&1&&1&\\2&2&-2&-1&&&1\end{array}\right]\to\left[\begin{array}{cccc|ccc}-1&0&1&2&1&0&0\\0&2&0&3&1&1&0\\0&0&0&0&1&-1&1\end{array}\right]$$

$$P=\begin{bmatrix}1&0&0\\1&1&0\\1&-1&1\end{bmatrix},\quad P^{-1}=\begin{bmatrix}1&0&0\\-1&1&0\\-2&1&1\end{bmatrix},\quad F=\begin{bmatrix}1&0\\-1&1\\-2&1\end{bmatrix}$$

$$G=\begin{bmatrix}-1&0&1&2\\0&2&0&3\end{bmatrix}:\quad\text{满秩分解为 } A=FG\ .$$

解(2)　$$(A\mid I)\to\cdots\to\left[\begin{array}{cccc|ccc}1&0&-1&-2&-1&0&0\\0&1&0&3/2&1/2&1/2&0\\0&0&0&0&1&-1&1\end{array}\right]$$

$$A=P^{-1}B=(F\mid S)\begin{pmatrix}I_2&B_{12}\\O&O\end{pmatrix}=(F\mid FB_{12})$$

故 $F=$ “A 的前 2 列” $=\begin{bmatrix}-1&0\\1&2\\2&2\end{bmatrix}$，$G=\begin{bmatrix}1&0&-1&-2\\0&1&0&3/2\end{bmatrix}$.

二、Hermite 标准形方法

1．Hermite 标准形：若 $B\in C_r^{m\times n}\ (r\geq 1)$满足

(1) B 的后 $m-r$ 行元素均为零；

(2) B 中有 r 列，设为 $c_1,\cdots,c_r$ 列，构成 I_m 的前 r 个列.

称 B 为拟 **Hermite** 标准形.

[注] 使用初等行变换可将任何非零矩阵化为拟 Hermite 标准形.

矩阵的 Hermite 标准形就是矩阵的行最简形（初等变换意义下).

2． 置换矩阵：划分单位矩阵 $I_n=(e_1,e_2,\cdots,e_n)$，称 $P_1=(e_{j_1},e_{j_2},\cdots,e_{j_n})$

为置换矩阵，其中 $j_1j_2\cdots j_n$ 是 $1,2,\cdots,n$ 的一个排列.

特点：划分 $A_{m\times n}=(a_1,a_2,\cdots,a_n)$，那么 $AP_1=(a_{j_1},a_{j_2},\cdots,a_{j_n})$.

Th14　设 $A\in C_r^{m\times n}$ 的拟 **Hermite** 标准形为 B，$F=$“A 的 $c_1,\cdots,c_r$ 列”

$G=$“B 的前 r 行”，则 A 的满秩分解为 $A=FG$.

证　$A\xrightarrow{\text{行}}B$ (拟Hermite标准形) $\Rightarrow PA=B$：　$A=P^{-1}B$

构造置换矩阵　$P_1=(e_{c_1},\cdots,e_{c_r},e_{c_{r+1}},\cdots,e_{c_n})$

$$B=(b_1,\cdots,b_n)\Rightarrow BP_1=(b_{c_1},\cdots,b_{c_r},b_{c_{r+1}},\cdots,b_{c_n})=\begin{pmatrix}I_r&B_{12}\\O&O\end{pmatrix}$$

$$A=(a_1,\cdots,a_n)\Rightarrow AP_1=(a_{c_1},\cdots,a_{c_r},a_{c_{r+1}},\cdots,a_{c_n})$$

$$P^{-1}=(F\mid S)\ ,\quad AP_1=P^{-1}\cdot(BP_1)=(F\mid S)\begin{pmatrix}I_r&B_{12}\\O&O\end{pmatrix}=(F\mid FB_{12})$$

故　$F=$ “ AP_1 的前 r 列” $=$ “A 的 $c_1,\cdots,c_r$ 列”

$G=$ “B 的前 r 行”

由此可得 $A = P^{-1}B = \left(F \mid S\right)\begin{pmatrix} G \\ O \end{pmatrix} = FG$.

例 2　$A = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 2 & 1 & 2 & 1 \\ 2 & 0 & 2 & 2 \\ 4 & 2 & 4 & 2 \end{bmatrix}$，求 $A = FG$.

解　$A \xrightarrow{行} \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & -2 \end{bmatrix} \xrightarrow{行} \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = B$

(1) $c_1 = 1, c_2 = 2$： $F =$ “ A 的第 1, 2 列” $= \begin{bmatrix} 1 & 0 \\ 2 & 1 \\ 2 & 0 \\ 4 & 2 \end{bmatrix}$

$G =$ “ B 的第 1, 2 行” $= \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & -1 \end{bmatrix}$

(2) $c_1 = 3, c_2 = 2$： $F =$ “ A 的第 3, 2 列”　（同前）

$G =$ “ B 的第 1, 2 行”　（同前）

(3) $A \xrightarrow{行} B \xrightarrow{行} \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = C$：　可取　$c_1 = 1, c_2 = 4$

$F =$ “ A 的第 1, 4 列” $= \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 2 & 2 \\ 4 & 2 \end{bmatrix}$

$G =$ “ C 的第 1, 2 行” $= \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{bmatrix}$

例 3　$A_1 \in \mathbf{C}_{r_1}^{m\times n},\ A_2 \in \mathbf{C}_{r_2}^{m\times n} \Rightarrow \mathrm{rank}(A_1 + A_2) \le r_1 + r_2$.

证　$r_1 \cdot r_2 = 0$：　成立.

$r_1 \cdot r_2 \ne 0$：　$A_1 = F_1 G_1,\ A_2 = F_2 G_2$　（$F_1 \in \mathbf{C}_{r_1}^{m\times r_1},\ F_2 \in \mathbf{C}_{r_2}^{m\times r_2}$）

$$A_1 + A_2 = (F_1 \mid F_2)\begin{pmatrix} G_1 \\ G_2 \end{pmatrix} \Rightarrow \text{秩}(A_1 + A_2) \le \text{秩}(F_1 \mid F_2) \le r_1 + r_2$$

§4.4 奇异值分解

一、预备知识

(1) $\forall A_{m\times n},\ (A^{\mathrm{H}} A)_{n\times n}$ 是 Hermite（半）正定矩阵.

$\forall x \ne 0,\ x^{\mathrm{H}} A^{\mathrm{H}} A x = (Ax)^{\mathrm{H}}(Ax) = |Ax|^2 \ge 0$

(2) 齐次方程组 $Ax = 0$ 与 $A^{\mathrm{H}} A x = 0$ 同解.

若 $Ax = 0$，则 $A^{\mathrm{H}} A x = 0$；

反之，$A^{\mathrm{H}} A x = 0 \Rightarrow |Ax|^2 = (Ax)^{\mathrm{H}}(Ax) = x^{\mathrm{H}}(A^{\mathrm{H}} A x) = 0 \Rightarrow Ax = 0$

(3) $\mathrm{rank}\, A = \mathrm{rank}(A^{\mathrm{H}} A)$

$S_1 = \{x \mid Ax = 0\},\quad S_2 = \{x \mid A^{\mathrm{H}} A x = 0\}$

$S_1 = S_2 \Rightarrow \dim S_1 = \dim S_2 \Rightarrow n - r_A = n - r_{A^{\mathrm{H}} A} \Rightarrow r_A = r_{A^{\mathrm{H}} A}$

(4) $A = O_{m\times n} \Leftrightarrow A^{\mathrm{H}} A = O_{n\times n}$

必要性．左乘 A^{H} 即得；充分性．$r_A = r_{A^{\mathrm{H}} A} = 0 \Rightarrow A = O$.

二、正交对角分解

Th15　$A_{n\times n}$ 可逆 $\Rightarrow \exists$ 酉矩阵 $U_{n\times n}$ 及 $V_{n\times n}$，使得

$$U^{\mathrm{H}} A V = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \end{bmatrix} \overset{\Delta}{=} D \quad (\sigma_i > 0)$$

证　$A^{\mathrm{H}} A$ 是 Hermite 正定矩阵，$\exists$ 酉矩阵 $V_{n\times n}$，使得

$$V^{\mathrm{H}}(A^{\mathrm{H}} A)V = \mathrm{diag}(\lambda_1, \cdots, \lambda_n) \overset{\Delta}{=} \Lambda \quad (\lambda_i > 0)$$

改写为　$D^{-1}V^{\mathrm{H}}A^{\mathrm{H}}\cdot AVD^{-1}=I\quad(\sigma_i=\sqrt{\lambda_i})$

令 $U\overset{\Delta}{=}AVD^{-1}$，则有 $U^{\mathrm{H}}U=I$，从而 U 是酉矩阵.

由此可得 $U^{\mathrm{H}}AV=U^{\mathrm{H}}UD=D$.

［注］ 称 $A=UDV^{\mathrm{H}}$ 为 A 的正交对角分解.

三、奇异值分解

$A\in\mathbf{C}_r^{m\times n}\ (r\geq 1)\Rightarrow A^{\mathrm{H}}A\in\mathbf{C}_r^{n\times n}$ 半正定

$A^{\mathrm{H}}A$ 的特征值：$\lambda_1\geq\lambda_2\geq\cdots\geq\lambda_r>\lambda_{r+1}=\cdots=\lambda_n=0$

A 的奇异值：$\sigma_i\overset{\Delta}{=}\sqrt{\lambda_i}$，$i=1,2,\cdots,n$

特点：⑴ A 的奇异值个数等于 A 的列数；

⑵ A 的非零奇异值个数 $=\mathrm{rank}A$.

Th16 $A\in\mathbf{C}_r^{m\times n}\ (r\geq 1)$，$\Sigma_r\overset{\Delta}{=}\begin{bmatrix}\sigma_1 & & \\ & \ddots & \\ & & \sigma_r\end{bmatrix}\Rightarrow$

存在酉矩阵 $U_{m\times m}$ 及 $V_{n\times n}$，使得 $U^{\mathrm{H}}AV=\begin{bmatrix}\Sigma_r & O\\ O & O\end{bmatrix}_{m\times n}\overset{\Delta}{=}D$.

证　对于 Hermite 半正定矩阵 $A^{\mathrm{H}}A$，$\exists$ 酉矩阵 $V_{n\times n}$，使得

$$V^{\mathrm{H}}(A^{\mathrm{H}}A)V=\begin{bmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{bmatrix}=\begin{bmatrix}\Sigma^2 & O\\ O & O\end{bmatrix}_{n\times n}$$

划分 $V=(V_1\,|\,V_2)$：V_1 是 V 的前 r 列，V_2 是 V 的后 $n-r$ 列.

$$A^{\mathrm{H}}AV=V\begin{bmatrix}\Sigma^2 & O\\ O & O\end{bmatrix}_{n\times n}:\quad (A^{\mathrm{H}}AV_1\,|\,A^{\mathrm{H}}AV_2)=(V_1\Sigma^2\,|\,O)$$

(1) $A^{\mathrm{H}}AV_1=V_1\Sigma^2$：

$$V_1^{\mathrm{H}}A^{\mathrm{H}}AV_1=\Sigma^2\Rightarrow(AV_1\Sigma^{-1})^{\mathrm{H}}(AV_1\Sigma^{-1})=I_r$$

(2) $A^{\mathrm{H}}AV_2=O$：

$$V_2^{\mathrm{H}}A^{\mathrm{H}}AV_2 = O \Rightarrow (AV_2)^{\mathrm{H}}(AV_2) = O \Rightarrow AV_2 = O$$

令 $U_1 \overset{\Delta}{=} AV_1\Sigma^{-1}$： $U_1^{\mathrm{H}}U_1 = I_r$

设 U_1 的列为 $u_1,\cdots,u_r$，扩充为 $\mathbf{C}^m$ 的标准正交基 $u_1,\cdots,u_r,u_{r+1},\cdots,u_m$，

则 $U_2 = (u_{r+1},\cdots,u_m)$满足 $U_2^{\mathrm{H}}U_1 = O_{(m-r)\times r}$. 记 $U = (U_1 \mid U_2)$，则有

$$U^{\mathrm{H}}AV = U^{\mathrm{H}}(AV_1 \mid AV_2) = \begin{pmatrix} U_1^{\mathrm{H}} \\ U_2^{\mathrm{H}} \end{pmatrix}(U_1\Sigma \mid O)$$

$$= \begin{bmatrix} U_1^{\mathrm{H}}U_1\Sigma & O \\ U_2^{\mathrm{H}}U_1\Sigma & O \end{bmatrix} = \begin{bmatrix} \Sigma & O \\ O & O \end{bmatrix}_{m\times n}$$

[注] 称 $A = UDV^{\mathrm{H}}$ 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 $A^{\mathrm{H}}A$ 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 1 $A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$，求 $A = UDV^{\mathrm{T}}$.

解 $A^{\mathrm{T}}A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix} = B$, $|\lambda I - B| = \lambda(\lambda-1)(\lambda-3)$

$$\lambda_1 = 3: \quad 3I - B = \begin{bmatrix} 2 & 0 & -1 \\ 0 & 2 & -1 \\ -1 & -1 & 1 \end{bmatrix}, \xi_1 = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$$

$$\lambda_2 = 1: \quad 1I - B = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & -1 \\ -1 & -1 & -1 \end{bmatrix}, \xi_2 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$$

$$\lambda_3 = 0: \quad 0I - B = \begin{bmatrix} -1 & 0 & -1 \\ 0 & -1 & -1 \\ -1 & -1 & -2 \end{bmatrix}, \xi_3 = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$$

$$r_A = 2: \quad \Sigma = \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \end{bmatrix}$$

$$V=\begin{bmatrix}\frac{1}{\sqrt6}&\frac{1}{\sqrt2}&\frac{1}{\sqrt3}\\\frac{1}{\sqrt6}&-\frac{1}{\sqrt2}&\frac{1}{\sqrt3}\\\frac{2}{\sqrt6}&0&-\frac{1}{\sqrt3}\end{bmatrix},\quad V_1=\begin{bmatrix}\frac{1}{\sqrt6}&\frac{1}{\sqrt2}\\\frac{1}{\sqrt6}&-\frac{1}{\sqrt2}\\\frac{2}{\sqrt6}&0\end{bmatrix}$$

$$U_1=AV_1\Sigma^{-1}=\begin{bmatrix}\frac{1}{\sqrt2}&\frac{1}{\sqrt2}\\\frac{1}{\sqrt2}&-\frac{1}{\sqrt2}\\0&0\end{bmatrix}，\text{取}\,U_2=\begin{bmatrix}0\\0\\1\end{bmatrix}，\quad\text{则}\,U=\begin{bmatrix}\frac{1}{\sqrt2}&\frac{1}{\sqrt2}&0\\\frac{1}{\sqrt2}&-\frac{1}{\sqrt2}&0\\0&0&1\end{bmatrix}$$

$$U^{\mathrm T}AV=\begin{bmatrix}\sqrt3&0&0\\0&1&0\\0&0&0\end{bmatrix}=D,\quad A=UDV^{\mathrm T}$$

Th17　在 $A\in\mathbf{C}_r^{m\times n}\ (r>0)$的奇异值分解 $A=U\begin{bmatrix}\Sigma&O\\O&O\end{bmatrix}V^{\mathrm H}$中，划分

$U=(u_1,u_2,\cdots,u_m),V=(v_1,v_2,\cdots,v_n)$，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) $A=\sigma_1u_1v_1^{\mathrm H}+\sigma_2u_2v_2^{\mathrm H}+\cdots+\sigma_ru_rv_r^{\mathrm H}$.

证　$A=(U_1\,|\,U_2)\begin{pmatrix}\Sigma&O\\O&O\end{pmatrix}\begin{pmatrix}V_1^{\mathrm H}\\V_2^{\mathrm H}\end{pmatrix}=U_1\Sigma V_1^{\mathrm H}$

容易验证：$U_1\Sigma V_1^{\mathrm H}x=0\Leftrightarrow V_1^{\mathrm H}x=0$

(1) $N(A)=\{\,x\,|\,Ax=0\,\}=\{\,x\,|\,U_1\Sigma V_1^{\mathrm H}x=0\,\}$

$=\{\,x\,|\,V_1^{\mathrm H}x=0\,\}=N(V_1^{\mathrm H})=R^{\perp}(V_1)$　　**(Th1-35)**

$=R(V_2)=\mathrm{span}\{\,v_{r+1},\cdots,v_n\}$

(2) $R(A)=\{\,y\,|\,y=Ax\,\}=\{\,y\,|\,y=U_1(\Sigma V_1^{\mathrm H}x)\,\}$

$\subset\{\,y\,|\,y=U_1z\,\}=R(U_1)$

$$R(U_1)=\{\,y\mid y=U_1z\,\}=\{\,y\mid y=A(V_1\Sigma^{-1}z)\,\}$$

$$\subset\{\,y\mid y=Ax\,\}=R(A)$$

故　$R(A)=R(U_1)=\mathrm{span}\{\,u_1,\cdots,u_r\,\}$

(3) $$A=(u_1,\cdots,u_r)\begin{bmatrix}\sigma_1 & & \\ & \ddots & \\ & & \sigma_r\end{bmatrix}\begin{bmatrix}v_1^{\mathrm{H}}\\ \vdots \\ v_r^{\mathrm{H}}\end{bmatrix}$$

$$=\sigma_1u_1v_1^{\mathrm{H}}+\cdots+\sigma_ru_rv_r^{\mathrm{H}}$$　（比较矩阵的谱分解）

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 $U^{\mathrm{H}}AV=B$，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 $\Rightarrow A$ 与 C 正交相抵.

Th18　A 与 B 正交相抵 $\Rightarrow \sigma_A=\sigma_B$.

证　$B=U^{\mathrm{H}}AV\Rightarrow B^{\mathrm{H}}B=\cdots=V^{-1}(A^{\mathrm{H}}A)V$

$$\Rightarrow \lambda_{B^{\mathrm{H}}B}=\lambda_{A^{\mathrm{H}}A}\geq 0\Rightarrow \sigma_B=\sigma_A$$

例 2　$A^{\mathrm{H}}=A\Rightarrow \sigma_A=|\lambda_A|$　$\left(\because \lambda_{A^{\mathrm{H}}A}=\lambda_{A^2}=(\lambda_A)^2\right)$

$A^{\mathrm{H}}=-A\Rightarrow \sigma_A=|\lambda_A|$　$\left(\because \lambda_{A^{\mathrm{H}}A}=\lambda_{(jA)^2}=(j\lambda_A)^2\right)$

（$A^{\mathrm{H}}=-A\Rightarrow \lambda_A$ 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) $A=UDV^{\mathrm{H}}\Rightarrow Dy=U^{\mathrm{H}}b\overset{\text{def}}{=}c,\ V^{\mathrm{H}}x=y$

$$D=\begin{bmatrix}\Sigma & O\\ O & O\end{bmatrix}_{m\times n},\quad \Sigma=\begin{bmatrix}\sigma_1 & & \\ & \ddots & \\ & & \sigma_r\end{bmatrix}$$

$$\begin{bmatrix}\Sigma & O\\ O & O\end{bmatrix}\begin{bmatrix}y_1\\ \vdots\\ y_n\end{bmatrix}=\begin{bmatrix}c_1\\ \vdots\\ c_m\end{bmatrix}\quad（隐含\ c_{r+1}=0,\cdots,c_m=0）$$

通解为
$$\begin{bmatrix}y_1\\ \vdots\\ y_r\\ y_{r+1}\\ \vdots\\ y_n\end{bmatrix}=\begin{bmatrix}c_1/\sigma_1\\ \vdots\\ c_r/\sigma_r\\ k_1\\ \vdots\\ k_{n-r}\end{bmatrix}\quad（k_1,\cdots,k_{n-r}\ 是任意常数）$$

$$x=Vy=(\frac{c_1}{\sigma_1}v_1+\cdots+\frac{c_r}{\sigma_r}v_r)+(k_1v_{r+1}+\cdots+k_{n-r}v_n)$$

[注] $k_1v_{r+1}+\cdots+k_{n-r}v_n$ 是 $A_{m\times n}x=0$ 的通解（Th17(1)）

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

第五章　特征值的估计

§5.1　特征值估计

一、特征值的上界

Th1　$A\in \mathbf{R}^{n\times n},\ M=\frac{1}{2}\cdot\max\limits_{i,j}\{|a_{ij}-a_{ji}|\}\Rightarrow |\operatorname{Im}(\lambda_A)|\le M\left[\frac{1}{2}n(n-1)\right]^{\frac{1}{2}}$

证明略去.

例 1　$A_{n\times n}$ 是实对称矩阵 $\Rightarrow \operatorname{Im}(\lambda_A)=0$, 即 $\lambda_A\in\mathbf{R}$.

引理 1　$\forall B=(b_{ij})_{n\times n},\ \forall y=\begin{bmatrix}\eta_1\\ \vdots\\ \eta_n\end{bmatrix}$ 满足 $\|y\|_2=1\Rightarrow |y^{\mathrm{H}}By|\le\|B\|_{m_\infty}$.

证　$|y^{\mathrm{H}}By|=\left|\sum\limits_{i,j}b_{ij}\bar{\eta}_i\eta_j\right|\le\max\limits_{i,j}|b_{ij}|\cdot\sum\limits_{i,j}|\eta_i||\eta_j|$

$$\le\max_{i,j}|b_{ij}|\cdot\frac{1}{2}\sum_{i,j}\left(|\eta_i|^2+|\eta_j|^2\right)=\max_{i,j}|b_{ij}|\cdot\frac{1}{2}(n+n)=\|B\|_{m_\infty}$$

Th2　$A_{n\times n}$, $|\lambda_A|\le\|A\|_{m_\infty}$, $|\operatorname{Re}(\lambda_A)|\le\frac{1}{2}\|A+A^{\mathrm{H}}\|_{m_\infty}$, $|\operatorname{Im}(\lambda_A)|\le\frac{1}{2}\|A-A^{\mathrm{H}}\|_{m_\infty}$.

证　设 A 的特征值为 λ，特征向量 x 满足 $\|x\|_2=1$，则

$$Ax=\lambda x\Rightarrow\lambda=x^{\mathrm{H}}Ax,\quad \bar{\lambda}=x^{\mathrm{H}}A^{\mathrm{H}}x$$

(1)　$|\lambda|=|x^{\mathrm{H}}Ax|\le\|A\|_{m_\infty}$

(2)　$\lambda+\bar{\lambda}=x^{\mathrm{H}}(A+A^{\mathrm{H}})x\Rightarrow|\operatorname{Re}(\lambda)|=\frac{1}{2}|\lambda+\bar{\lambda}|\le\frac{1}{2}\|A+A^{\mathrm{H}}\|_{m_\infty}$

(3)　$\lambda-\bar{\lambda}=x^{\mathrm{H}}(A-A^{\mathrm{H}})x\Rightarrow|\operatorname{Im}(\lambda)|=\frac{1}{2}|\lambda-\bar{\lambda}|\le\frac{1}{2}\|A-A^{\mathrm{H}}\|_{m_\infty}$

例 2　$A^{\mathrm{H}}=A\Rightarrow\operatorname{Im}(\lambda_A)=0$, 即 $\lambda_A\in\mathbf{R}$.

$A^{\mathrm{H}}=-A\Rightarrow\operatorname{Re}(\lambda_A)=0$, 即 λ_A 是纯虚数或零.

Th4　$A_{n\times n}=(a_1,a_2,\cdots,a_n)\Rightarrow|\lambda_1||\lambda_2|\cdots|\lambda_n|\le\|a_1\|_2\|a_2\|_2\cdots\|a_n\|_2$.

证 $\det A=0$ 时，结论成立.

$\det A\neq 0\Rightarrow a_1,a_2,\cdots,a_n$ 线性无关．正交化可得

$$\begin{cases} b_1=a_1\\ b_2=a_2-k_{21}b_1\\ \quad\cdots\cdots\\ b_n=a_n-k_{n,n-1}b_{n-1}-\cdots-k_{n1}b_1\end{cases}\qquad \begin{cases} a_1=b_1\\ a_2=k_{21}b_1+b_2\\ \quad\cdots\cdots\\ a_n=k_{n1}b_1+\cdots+k_{n,n-1}b_{n-1}+b_n\end{cases}$$

$$(a_1,a_2,\cdots,a_n)=(b_1,b_2,\cdots,b_n)\begin{bmatrix}1 & k_{21} & \cdots & k_{n1}\\ & 1 & \cdots & k_{n2}\\ & & \ddots & \vdots\\ & & & 1\end{bmatrix}$$

$B\overset{\Delta}{=}(b_1,b_2,\cdots,b_n)$： $\det A=\det B$， $b_i\perp b_j\ (i\neq j)$

$$\|a_i\|_2^2=\|(k_{i1}b_1+\cdots+k_{i,i-1}b_{i-1})+b_i\|_2^2$$

$$=\|k_{i1}b_1+\cdots+k_{i,i-1}b_{i-1}\|_2^2+\|b_i\|_2^2\geq\|b_i\|_2^2$$

$$|\det B|^2=\overline{\det B}\cdot\det B=\det\overline{B}\cdot\det B=\det B^{\mathrm H}\cdot\det B=\det(B^{\mathrm H}B)$$

$$=\det\begin{bmatrix}\|b_1\|_2^2 & & \\ & \ddots & \\ & & \|b_n\|_2^2\end{bmatrix}=(\|b_1\|_2\cdots\cdots\|b_n\|_2)^2$$

$$|\lambda_1|\cdots|\lambda_n|=|\det A|=|\det B|=\|b_1\|_2\cdots\|b_n\|_2\leq\|a_1\|_2\cdots\|a_n\|_2$$

Th5 设 $A_{n\times n}$ 的特征值为 $\lambda_1,\cdots,\lambda_n$， 则 $|\lambda_1|^2+\cdots+|\lambda_n|^2\leq\|A\|_{\mathrm F}^2$.

证 对 A，存在酉矩阵 $U_{n\times n}$, st. $U^{\mathrm H}AU=\begin{bmatrix}\lambda_1 & \cdots & *\\ & \ddots & \vdots\\ & & \lambda_n\end{bmatrix}\overset{\Delta}{=}T$

由此可得 $|\lambda_1|^2+\cdots+|\lambda_n|^2\leq\|T\|_{\mathrm F}^2=\|A\|_{\mathrm F}^2$ （Th2.3）.

例 3 设 $A_{n\times n}$ 的特征值为 $\lambda_1,\cdots,\lambda_n$， 则

$$A^{\mathrm H}A=AA^{\mathrm H}\Leftrightarrow|\lambda_1|^2+\cdots+|\lambda_n|^2=\|A\|_{\mathrm F}^2$$

证 必要性． $A^{\mathrm H}A=AA^{\mathrm H}$： 存在酉矩阵 P, 使得

$$P^{\mathrm{H}}AP=\begin{bmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{bmatrix}\overset{\Delta}{=}\Lambda$$

$$|\lambda_1|^2+\cdots+|\lambda_n|^2=\|\Lambda\|_{\mathrm{F}}^2=\|A\|_{\mathrm{F}}^2$$

充分性．对 A，存在酉矩阵 U，使得 $U^{\mathrm{H}}AU=\begin{bmatrix}\lambda_1 & \cdots & * \\ & \ddots & \vdots \\ & & \lambda_n\end{bmatrix}\overset{\Delta}{=}T$

$$\|T\|_{\mathrm{F}}^2=\|A\|_{\mathrm{F}}^2=|\lambda_1|^2+\cdots+|\lambda_n|^2\Rightarrow t_{ij}=0\quad(i<j)$$

$$T=\begin{bmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{bmatrix}\Rightarrow T^{\mathrm{H}}T=TT^{\mathrm{H}}\Rightarrow A^{\mathrm{H}}A=AA^{\mathrm{H}}$$

二、特征值的包含域

1. 盖尔圆：设 $A=(a_{ij})_{n\times n}$，记 $R_i\overset{\Delta}{=}\sum\limits_{\substack{j=1\\ j\neq i}}^{n}|a_{ij}|\quad(i=1,2,\cdots,n)$，称

$G_i\overset{\Delta}{=}\{z\mid |z-a_{ii}|\le R_i\}$为 A 的第 i 个盖尔圆（Gerschgorin 圆）.

Th6　矩阵 $A_{n\times n}$ 的特征值 $\lambda_A\in\bigcup\limits_{i=1}^{n}G_i$.

证　设 A 的特征值为 λ，对应的特征向量为 $x=(\xi_1,\xi_2,\cdots,\xi_n)^{\mathrm{T}}$，

选取 i_0 使得 $|\xi_{i_0}|=\max\limits_{i}|\xi_i|$，则 $\xi_{i_0}\neq 0$，且有

$$Ax=\lambda x\Rightarrow\sum_{j=1}^{n}a_{i_0j}\xi_j=\lambda\xi_{i_0}\quad(i_0-\text{分量})$$

$$\Rightarrow(\lambda-a_{i_0i_0})\xi_{i_0}=\sum_{j\neq i_0}a_{i_0j}\xi_j$$

$$\Rightarrow|\lambda-a_{i_0i_0}|=\left|\sum_{j\neq i_0}a_{i_0j}\frac{\xi_j}{\xi_{i_0}}\right|\le\sum_{j\neq i_0}|a_{i_0j}|\frac{|\xi_j|}{|\xi_{i_0}|}\le R_{i_0}$$

$$\Rightarrow\lambda\in G_{i_0}\subset\bigcup_{i=1}^{n}G_i$$

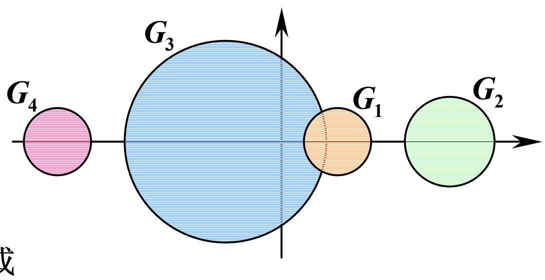
例 4　$A=\begin{bmatrix}10 & 1 & 2 & 3\\ 5 & 30 & 1 & 2\\ 10 & 3 & -10 & 5\\ 2 & -3 & -1 & -40\end{bmatrix}$，估计 λ_A.

$G_1:|z-10|\le 6$　　$G_2:|z-30|\le 8$

$G_3:|z+10|\le 18$

$G_4:|z+40|\le 6$

$\lambda_A\in G_1\cup G_2\cup G_3\cup G_4$



2．连通部分：⑴ 孤立的 G-圆，或

⑵ 相交的 G-圆构成的最大连通区域.

Th7　设 S 为 $A_{n\times n}$ 的 G-圆的一个连通部分，则

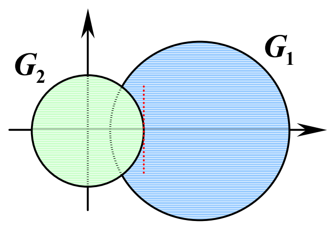
S 由 k 个 G-圆构成 $\Leftrightarrow$ S 中恰好有 A 的 k 个特征值.

其中 G-圆重叠时重复计数，特征值相同是亦重复计数.

例 5　$A=\begin{bmatrix}10 & -8\\ 5 & 0\end{bmatrix}$，估计 λ_A.

$G_1:|z-10|\le 8$

$G_2:|z|\le 5$



$S=G_1\cup G_2$ 是一个连通部分 $\overset{\text{Th7}}{\Rightarrow}\lambda_{1,2}\in S$.

计算得 $\lambda_{1,2}=5\pm j\sqrt{15}$，　$|\lambda_{1,2}|=\sqrt{40}>5\Rightarrow\lambda_{1,2}\notin G_2$.

3．特征值的分离问题

找 n 个孤立的 G-圆，使其中各包含 $A_{n\times n}$ 的一个特征值.

$$D=\begin{bmatrix}d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n\end{bmatrix}\ (d_i>0)$$

$$B \overset{\Delta}{=} DAD^{-1} = \left(a_{ij}\frac{d_i}{d_j}\right)_{n\times n} = \begin{pmatrix} a_{11} & a_{12}\frac{d_1}{d_2} & \cdots & a_{1n}\frac{d_1}{d_n} \\ a_{21}\frac{d_2}{d_1} & a_{22} & \cdots & a_{2n}\frac{d_2}{d_n} \\ \vdots & \vdots & & \vdots \\ a_{n1}\frac{d_n}{d_1} & a_{n2}\frac{d_n}{d_2} & \cdots & a_{nn} \end{pmatrix}$$

B 的 G-圆：$G_i' = \{z \mid |z-a_{ii}| \le r_i\}$，$r_i \overset{\Delta}{=} \sum_{j\ne i} |a_{ij}| \frac{d_i}{d_j}$

A 相似于 $B \Rightarrow \lambda_A \in \bigcup_{i=1}^{n} G_i'$　　　(Th6)

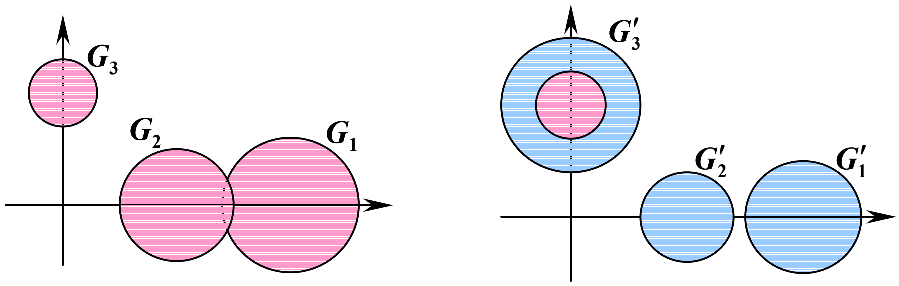
［注］当 G_i' 是 B 的孤立 G-圆时，G_i' 中恰有 A 的一个特征值（Th7）.

例 6　$A = \begin{bmatrix} 20 & 5 & 0.8 \\ 4 & 10 & 1 \\ 1 & 2 & 10j \end{bmatrix}$，分离 λ_A.

解　G_1：$|z-20| \le 5.8$　G_2：$|z-10| \le 5$　G_3：$|z-10j| \le 3$

对 A，取 $D = \begin{bmatrix} 1 & & \\ & 1 & \\ & & 2 \end{bmatrix}$，则 $B = DAD^{-1} = \begin{bmatrix} 20 & 5 & 0.4 \\ 4 & 10 & 0.5 \\ 2 & 4 & 10j \end{bmatrix}$.

G_1'：$|z-20| \le 5.4$　G_2'：$|z-10| \le 4.5$　G_3'：$|z-10j| \le 6$



结论：G_1', G_2' 及 G_3 中各有 A 的一个特征值.

［注］　在例 6 中，G_3 孤立，G_3' 孤立，故可取较小者；

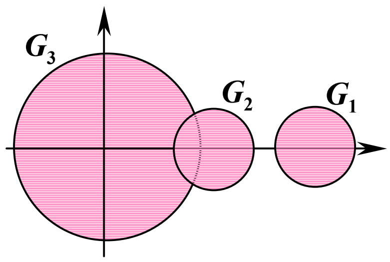
A^{T} 的 G-圆称为 A 的列 G-圆，A^{T} 与 A 的特征值相同.

例 7　$A=\begin{bmatrix}20 & 3 & 1\\ 2 & 10 & 2\\ 8 & 1 & 0\end{bmatrix}$，分离 λ_A.

解 (1) G_1 : $|z-20|\le 4$

G_2 : $|z-10|\le 4$

G_3 : $|z|\le 9$



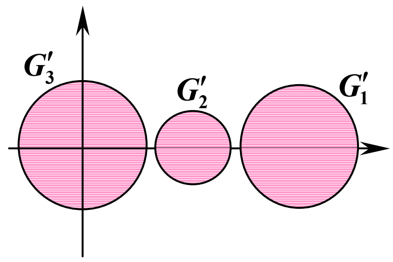
G_1 孤立，G_2 与 G_3 相交.

(2) 对 A，取 $D=\begin{bmatrix}1.5 & & \\ & 1 & \\ & & 1\end{bmatrix}$，则 $B=DAD^{-1}=\begin{bmatrix}20 & 4.5 & 1.5\\ 4/3 & 10 & 2\\ 16/3 & 1 & 0\end{bmatrix}$.

G_1' : $|z-20|\le 6$

G_2' : $|z-10|\le \dfrac{10}{3}$

G_3' : $|z|\le \dfrac{19}{3}$



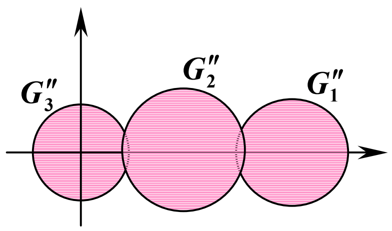
G_1'，G_2'，G_3' 都是孤立的盖尔圆，其中各有 A 的一个特征值.

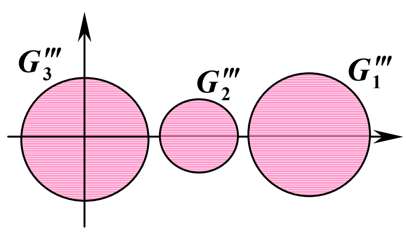
结论：结合(1)可得 G_1 ,G_2' 及 G_3' 中各有 A 的一个特征值.

(3) 对 A，取 $D=\begin{bmatrix}1 & & \\ & 1 & \\ & & 0.5\end{bmatrix}$，则 $B=DAD^{-1}=\begin{bmatrix}20 & 3 & 2\\ 2 & 10 & 4\\ 4 & 0.5 & 0\end{bmatrix}$.

G_1'' : $|z-20|\le 5$　　G_2'' : $|z-10|\le 6$　　G_3'' : $|z|\le 4.5$

G_1'' 与 G_2'' 相交，G_2'' 与 G_3'' 相交.





考虑 B^{T} .　G_1''': $|z-20|\le 6$　G_2''': $|z-10|\le 3.5$　G_3''': $|z|\le 6$

G_1'''，G_2'''，G_3''' 都是孤立的盖尔圆，其中各有 A 的一个特征值.

结论：结合(1)可得 G_1, G_2''' 及 G_3''' 中各有 A 的一个特征值.

例 8　$A_{n\times n}$ 严格对角占优 $\Rightarrow \det A \neq 0$.

证　只证“行优”情形：$|a_{ii}| > \sum_{j\neq i} |a_{ij}| = R_i,\ i = 1,2,\cdots,n$

$\forall \lambda_A, \exists i_0, \text{st.} \lambda_A \in G_{i_0} \Rightarrow |\lambda_A - a_{i_0 i_0}| \le R_{i_0}$

若 $\lambda_A = 0$，则 $|a_{i_0 i_0}| \le R_{i_0}$ 与 A“行优”矛盾. 故 $\lambda_A \neq 0 \Rightarrow \det A \neq 0$.

4．矩阵 $A = (a_{ij})_{n\times n}$ $(n>1)$的 Cassini 卵形

$$\Omega_{ij} \overset{\Delta}{=} \left\{ z \,\middle|\, |z - a_{ii}| \cdot |z - a_{jj}| \le R_i R_j \right\} \quad (i \neq j)$$

Th12 矩阵 $A_{n\times n}$ $(n>1)$ 的特征值 $\lambda_A \in \bigcup_{i\neq j} \Omega_{ij}$.

证 设 A 的特征值为 λ，对应的特征向量为 $x = (\xi_1, \xi_2, \cdots, \xi_n)^{\mathrm{T}}$.

选取 $i_0 \neq j_0$ 满足 $|\xi_{i_0}| \ge |\xi_{j_0}| \ge |\xi_k| \ (k \neq i_0, j_0)$，下证 $\lambda \in \Omega_{i_0 j_0}$.

(1) $\xi_{j_0} = 0$：此时 $\xi_k = 0\ (k \neq i_0, j_0)$, $\xi_{i_0} \neq 0\ (\because\ x \neq 0)$

$$Ax = \lambda x \Rightarrow \lambda \xi_{i_0} = \sum_{k=1}^{n} a_{i_0 k} \xi_k = a_{i_0 i_0} \xi_{i_0} \Rightarrow \lambda = a_{i_0 i_0}$$

$$\Rightarrow |\lambda - a_{i_0 i_0}| \cdot |\lambda - a_{j_0 j_0}| = 0 \le R_{i_0} R_{j_0}$$

(2) $\xi_{j_0} \neq 0$：此时 $|\xi_{i_0}| \ge |\xi_{j_0}| > 0$

$$Ax = \lambda x \Rightarrow \sum_{k=1}^{n} a_{lk} \xi_k = \lambda \xi_l \quad (l = 1,2,\cdots,n)$$

$$\Rightarrow (\lambda - a_{ll}) \xi_l = \sum_{k\neq l} a_{lk} \xi_k \quad (l = 1,2,\cdots,n)$$

$$l = i_0:\quad |\lambda - a_{i_0 i_0}| \cdot |\xi_{i_0}| \le \sum_{k\neq i_0} |a_{i_0 k}| \cdot |\xi_k| \le |\xi_{j_0}| \cdot R_{i_0}$$

$$l = j_0:\quad |\lambda - a_{j_0 j_0}| \cdot |\xi_{j_0}| \le \sum_{k\neq j_0} |a_{j_0 k}| \cdot |\xi_k| \le |\xi_{i_0}| \cdot R_{j_0}$$

故 $\left|\lambda-a_{i_0i_0}\right|\cdot\left|\lambda-a_{j_0j_0}\right|\le R_{i_0}R_{j_0}$

因此 $\lambda\in\Omega_{i_0j_0}\subset\bigcup\limits_{i\ne j}\Omega_{ij}$.

例 9 $A_{n\times n}(n>1)$满足$\left|a_{ii}\right|\cdot\left|a_{jj}\right|>R_iR_j\ (i\ne j)\Rightarrow\det A\ne 0$.

证　因为$\forall\ i\ne j,\ \left|a_{ii}\right|\cdot\left|a_{jj}\right|>R_iR_j$, 所以

$$0\notin\Omega_{ij}\ \Rightarrow\ 0\text{ 不是 }A\text{ 的特征值}\ \Rightarrow\ \det A\ne 0$$

例 10 $A=\begin{bmatrix}20 & 11 & 10\\ -8 & 30 & 20\\ 13 & 11 & 30\end{bmatrix}$， $R_1=21, R_2=28, R_3=24$

$\left|a_{11}\right|\cdot\left|a_{22}\right|=600>588=R_1R_2$， $\left|a_{11}\right|\cdot\left|a_{33}\right|=600>504=R_1R_3$

$\left|a_{22}\right|\cdot\left|a_{33}\right|=900>672=R_2R_3$. 由例 9 可得$\det A\ne 0$.

例 11 $\bigcup\limits_{i\ne j}\Omega_{ij}\subset\bigcup\limits_{k=1}^{n}G_k$

证 $\forall z\in$左，$\exists\ i\ne j$, st. $z\in\Omega_{ij}\Rightarrow\left|z-a_{ii}\right|\cdot\left|z-a_{jj}\right|\le R_iR_j$

$\Rightarrow\left|z-a_{ii}\right|>R_i$与$\left|z-a_{jj}\right|>R_j$至少之一不成立

$\Rightarrow\left|z-a_{ii}\right|\le R_i$与$\left|z-a_{jj}\right|\le R_j$至少之一成立

$\Rightarrow z\in G_i$或者$z\in G_j$， 即$z\in$右.

§5.2　广义特征值问题

$A_{n\times n}$ 实对称， $B_{n\times n}$ 实对称正定，确定数 λ 及非零列向量 x 使 $Ax=\lambda Bx$.

直接解法： $Ax=\lambda Bx \Leftrightarrow (\lambda B-A)x=0$

$$|\lambda B-A|=0 \Rightarrow \lambda_1,\cdots,\lambda_n$$

$$(\lambda_i B-A)x=0 \Rightarrow \text{基础解系}$$

一、等价形式

(1)　$Ax=\lambda Bx \Leftrightarrow (B^{-1}A)x=\lambda x$ ： $B^{-1}A$ 不一定对称

(2)　B 对称正定 $\Rightarrow B=GG^{\mathrm{T}}$ ： G 可逆

$$Ax=\lambda Bx \Leftrightarrow G^{-1}A(G^{-1})^{\mathrm{T}}G^{\mathrm{T}}x=\lambda G^{\mathrm{T}}x$$

$$\Leftrightarrow Sy=\lambda y:\quad S\overset{\Delta}{=}G^{-1}A(G^{-1})^{\mathrm{T}}\text{对称，}\quad y\overset{\Delta}{=}G^{\mathrm{T}}x\,.$$

二、正交性

1．按 B-正交：设 B 为实对称正定矩阵, 列向量 $x_1,\cdots,x_m\in\mathbf{R}^n$.

若 $x_i{}^{\mathrm{T}}Bx_j=0\ (i\neq j)$，称 $x_1,\cdots,x_m$ 按 B-正交.

性质　非零向量 $x_1,\cdots,x_m$ 按 B-正交 $\Rightarrow x_1,\cdots,x_m$ 线性无关.

证　$k_1x_1+\cdots+k_ix_i+\cdots+k_mx_m=0$

左乘 $x_i^{\mathrm{T}}B$ ： $k_ix_i^{\mathrm{T}}Bx_i=0\Rightarrow k_i=0$　$(\because B\text{是正定矩阵})$

故 $x_1,\cdots,x_m$ 线性无关.

2．按 B-标准正交：设 $x_1,\cdots,x_m$ 按 B-正交，若 $\|x_i\|_B=1\ (i=1,\cdots,m)$，

称 $x_1,\cdots,x_m$ 按 B-标准正交.（$\|x\|_B=\sqrt{x^{\mathrm{T}}Bx}$）

三、广义特征向量的正交性

S 实对称 $\Rightarrow$　特征值： $\lambda_1,\cdots,\lambda_n\in\mathbf{R}$

特征向量： $y_1,\cdots,y_n\in\mathbf{R}^n$ 标准正交

$$x_i \overset{\Delta}{=} \left(G^{-1}\right)^{\mathrm{T}} y_i \Rightarrow x_i^{\mathrm{T}} B x_j = y_i^{\mathrm{T}} G^{-1} \cdot G G^{\mathrm{T}} \cdot (G^{-1})^{\mathrm{T}} y_j = y_i^{\mathrm{T}} y_j = \delta_{ij}$$

$\Rightarrow x_1,\cdots,x_n$ 按 B-标准正交

$\Rightarrow x_1,\cdots,x_n$ 线性无关

§5.3　对称矩阵特征值的极性

问题：若 x 是实对称矩阵 A 的特征向量，则有 $Ax=\lambda x \Rightarrow \lambda = \dfrac{x^{\mathrm{T}}Ax}{x^{\mathrm{T}}x}$.

下面讨论多元函数 $\dfrac{x^{\mathrm{T}}Ax}{x^{\mathrm{T}}x}$ $(x\neq 0)$ 的极值问题.

一、常义 Rayleigh 商：

$A_{n\times n}$ 实对称，$x\in \mathbf{R}^n$，$R(x) \overset{\Delta}{=} \dfrac{x^{\mathrm{T}}Ax}{x^{\mathrm{T}}x}$ $(x\neq 0)$.

(1) $R(x)$ 连续；

(2) $\forall$ 实数 $k\neq 0$, $R(kx)=R(x)$；

(3) $\forall x_0\neq 0$，当 $0\neq x\in L(x_0)$ 时，$R(x)=R(x_0)$；

$\because x\in L(x_0) \Rightarrow x = kx_0 \Rightarrow R(x)=R(kx_0)=R(x_0)$

(4) $\min\limits_{x\neq 0} R(x)$ 与 $\max\limits_{x\neq 0} R(x)$ 存在，且在 $S=\left\{x \,\middle|\, x\in \mathbf{R}^n, \|x\|_2=1\right\}$ 上达到.

证　因为子集 S 是闭集，且 $R(x)$ 在 S 上连续，所以

$$\min_{x\in s} R(x) = m_1,\ \max_{x\in s} R(x) = m_2$$

$$\forall 0\neq y\in \mathbf{R}^n, x=\frac{1}{\|y\|_2} y\in S \Rightarrow m_1 \le R(x) \le m_2$$

由(2)：$R(x)=R(y) \Rightarrow m_1 \le R(y) \le m_2$

约定：实对称矩阵 $A_{n\times n}$ 的特征值排序为 $\lambda_1\le\lambda_2\le\cdots\le\lambda_n$.

Th16　$A_{n\times n}$ 实对称 $\Rightarrow \min\limits_{x\neq 0} R(x)=\lambda_1, \max\limits_{x\neq 0} R(x)=\lambda_n$.

证　设 A 的特征值为 $\lambda_1,\cdots,\lambda_n$，特征向量 $p_1,\cdots,p_n$ 标准正交.

$$\forall x \neq 0 \Rightarrow x = c_1 p_1 + \cdots + c_n p_n \quad (c_1^2 + \cdots + c_n^2 \neq 0)$$

$$Ax = c_1\lambda_1 p_1 + \cdots + c_n\lambda_n p_n\text{，}\quad x^{\mathrm{T}}Ax = c_1^2\lambda_1 + \cdots + c_n^2\lambda_n$$

$$x^{\mathrm{T}}x = c_1^2 + \cdots + c_n^2\text{：}\quad k_i \overset{\Delta}{=} \frac{c_i^2}{c_1^2 + \cdots + c_n^2} \Rightarrow k_1 + \cdots + k_n = 1$$

$$\left.\begin{array}{l} R(x) = k_1\lambda_1 + \cdots + k_n\lambda_n \Rightarrow \lambda_1 \leq R(x) \leq \lambda_n \\ R(p_1) = \lambda_1,\quad R(p_n) = \lambda_n \end{array}\right\} \Rightarrow \begin{cases} \min\limits_{x\neq 0} R(x) = \lambda_1 \\ \max\limits_{x\neq 0} R(x) = \lambda_n \end{cases}$$

[注] ① 在闭集 S 上，p_1 是 $R(x)$ 的极小点，p_n 是 $R(x)$ 的极大点.

② 若 $\lambda_1 = \cdots = \lambda_k < \lambda_{k+1}\ (1 \leq k \leq n)$，则在闭集 S 上，$R(x)$ 的全部极小点为 $l_1 p_1 + \cdots + l_k p_k\ (l_1^2 + \cdots + l_k^2 = 1)$.

③ $p_1, \cdots, p_n$ 构成 $\mathbf{R}^n$ 的标准正交基:

$$\mathbf{R}^n = L(p_1, p_n) + L(p_2, \cdots, p_{n-1}),\ L(p_2, \cdots, p_{n-1}) = [L(p_1, p_n)]^{\perp}$$

$$[L(p_1, p_n)]^{\perp} = \{x \mid x \perp p_1, x \perp p_n\} = \left\{ x \,\middle|\, \begin{pmatrix} p_1^{\mathrm{T}} \\ p_n^{\mathrm{T}} \end{pmatrix} x = 0 \right\}$$

Th16 的进一步讨论：设 $0 \neq x \in L(p_2, \cdots, p_{n-1})$，则有

$$x = c_2 p_2 + \cdots + c_{n-1} p_{n-1} \quad (c_2^2 + \cdots + c_{n-1}^2 \neq 0)$$

$$x^{\mathrm{T}}Ax = c_2^2\lambda_2 + \cdots + c_{n-1}^2\lambda_{n-1}$$

$$x^{\mathrm{T}}x = c_2^2 + \cdots + c_{n-1}^2\text{：}\quad k_i \overset{\Delta}{=} \frac{c_i^2}{c_2^2 + \cdots + c_{n-1}^2} \Rightarrow k_2 + \cdots + k_{n-1} = 1$$

$$\left.\begin{array}{l} R(x) = k_2\lambda_2 + \cdots + k_{n-1}\lambda_{n-1} \Rightarrow \lambda_2 \leq R(x) \leq \lambda_{n-1} \\ R(p_2) = \lambda_2,\quad R(p_{n-1}) = \lambda_{n-1} \end{array}\right\} \Rightarrow \begin{cases} \min\limits_{x\neq 0} R(x) = \lambda_2 \\ \max\limits_{x\neq 0} R(x) = \lambda_{n-1} \end{cases}$$

Th17　设 $A_{n\times n}$ 实对称，特征向量 $p_1, \cdots, p_n$ 标准正交，对于 $1 \leq r \leq s \leq n$，当 $x \in L(p_r, p_{r+1}, \cdots, p_s)$ 时，有 $\min\limits_{x\neq 0} R(x) = \lambda_r$，$\max\limits_{x\neq 0} R(x) = \lambda_s$.

Th18　设 $A_{n\times n}$ 实对称，$\forall V_k \subset \mathbf{R}^n$ 且 $\dim V_k = k$，则 A 的第 k 个特征值为 $\lambda_k = \min\limits_{V_k}\left[\max\{R(x) \mid 0 \neq x \in V_k\}\right]$.

证　A 的特征向量：$p_1, \cdots, p_n$ 标准正交

(1) $W_k = L(p_k,\cdots,p_n) \subset \mathbf{R}^n \Rightarrow \dim W_k = n-k+1$

$V_k \subset \mathbf{R}^n \Rightarrow V_k + W_k \subset \mathbf{R}^n$　　　（注意 $\dim V_k = k$）

因为 $n \geq \dim(V_k + W_k) = \dim V_k + \dim W_k - \dim(V_k \cap W_k)$

$$= n+1-\dim(V_k \cap W_k)$$

所以 $\dim(V_k \cap W_k) \geq 1 \Rightarrow \exists x_0 \in V_k \cap W_k$ 且 $x_0 \neq 0$. 由 $x_0 \in W_k$ 可得

$x_0 = c_k p_k + \cdots + c_n p_n$　（$c_k^2 + \cdots + c_n^2 \neq 0$）

故 $$R(x_0) = \frac{x_0^{\mathrm{T}} A x_0}{x_0^{\mathrm{T}} x_0} = \frac{c_k^2 \lambda_k + \cdots + c_n^2 \lambda_n}{c_k^2 + \cdots + c_n^2} \geq \lambda_k$$

$\max\{R(x) \mid x \in V_k\} \geq R(x_0) \geq \lambda_k \Rightarrow$ 右 $\geq$ 左

(2) 令 $V_k^0 = L(p_1,\cdots,p_k)$，对 $\forall x \in V_k^0\ (x \neq 0)$，有

$x = l_1 p_1 + \cdots + l_k p_k \quad (l_1^2 + \cdots + l_k^2 \neq 0)$

$$R(x) = \frac{x^{\mathrm{T}} A x}{x^{\mathrm{T}} x} = \frac{l_1^2 \lambda_1 + \cdots + l_k^2 \lambda_k}{l_1^2 + \cdots + l_k^2} \leq \lambda_k$$

$\max\{R(x) \mid x \in V_k^0\} \leq \lambda_k \Rightarrow$ 右 $\leq$ 左

二、广义 **Rayleigh** 商：$A_{n\times n}$ 实对称，$B_{n\times n}$ 实对称正定，$x \in \mathbf{R}^n$

$$R_B(x) \overset{\Delta}{=} \frac{x^{\mathrm{T}} A x}{x^{\mathrm{T}} B x} \quad (x \neq 0)$$

(1) $R_B(x)$ 连续；

(2) $\forall$ 实数 $k \neq 0$, $R_B(kx) = R_B(x)$；

(3) $\forall x_0 \neq 0$，当 $0 \neq x \in L(x_0)$ 时，$R_B(x) = R_B(x_0)$；

(4) $\min\limits_{x\neq 0} R_B(x)$ 与 $\max\limits_{x\neq 0} R_B(x)$ 存在，且在 $S_B = \{x \mid x \in \mathbf{R}^n, \|x\|_B = 1\}$ 上达到.

Th22　$x_0 \neq 0$ 是 $R_B(x)$ 的驻点 $\Leftrightarrow Ax_0 = \lambda B x_0$.

证　$(x^{\mathrm{T}} B x) \cdot R_B(x) = (x^{\mathrm{T}} A x)$，　$2Bx \cdot R_B(x) + (x^{\mathrm{T}} B x) \cdot \dfrac{dR_B(x)}{dx} = 2Ax$

$$\frac{dR_B(x)}{dx}=\frac{2}{x^{\mathrm{T}}Bx}\cdot\left[Ax-R_B(x)\cdot Bx\right]$$

必要性．$x_0\neq 0$ 使 $\left.\frac{dR_B}{dx}\right|_{x=x_0}=0$：　$Ax_0=R_B(x_0)\cdot Bx_0$　$\left(\lambda\overset{\Delta}{=}R_B(x_0)\right)$

充分性．$x_0\neq 0$ 使 $Ax_0=\lambda Bx_0$：　$\lambda=R_B(x_0)$ 且有 $\left.\frac{dR_B}{dx}\right|_{x=x_0}=0$

约定：$Ax=\lambda Bx$ 的特征值排序为 $\lambda_1\leq\lambda_2\leq\cdots\leq\lambda_n$

广义特征向量 $p_1,p_2,\cdots,p_n$ “按 B-标准正交”

Th23　$\forall V_k\subset \mathbf{R}^n$ 且 $\dim V_k=k\Rightarrow \lambda_k=\min\limits_{V_k}\left[\max\left\{R_B(x)\,\middle|\,0\neq x\in V_k\right\}\right]$

证　类似于 Th18．

Th 23′　$\forall V_k\subset \mathbf{R}^n$ 且 $\dim V_k=k\Rightarrow \lambda_{n-k+1}=\max\limits_{V_k}\left[\min\left\{R_B(x)\,\middle|\,0\neq x\in V_k\right\}\right]$

证　由 Th23 可得，$(-A)x=(-\lambda)Bx$ 的第 k 个特征值为

$$\begin{aligned}-\lambda_{n-k+1}&=\min_{V_k}\left[\max\left\{-R_B(x)\,\middle|\,0\neq x\in V_k\right\}\right]\\&=\min_{V_k}\left[(-1)\cdot\min\left\{R_B(x)\,\middle|\,0\neq x\in V_k\right\}\right]\\&=(-1)\max_{V_k}\left[\min\left\{R_B(x)\,\middle|\,0\neq x\in V_k\right\}\right]\end{aligned}$$

特殊情形：

Th23 $\Rightarrow \lambda_n=\max\left\{R_B(x)\,\middle|\,0\neq x\in\mathbf{R}^n\right\}=\max\left\{x^{\mathrm{T}}Ax\,\middle|\,x\in\mathbf{R}^n\text{且}\ \|x\|_B=1\right\}$

Th 23′ $\Rightarrow \lambda_1=\min\left\{R_B(x)\,\middle|\,0\neq x\in\mathbf{R}^n\right\}=\min\left\{x^{\mathrm{T}}Ax\,\middle|\,x\in\mathbf{R}^n\text{且}\ \|x\|_B=1\right\}$

三、矩阵奇异值的极性

$A\in\mathbf{R}_r^{m\times n}$：　$\sigma(A)=\left[\lambda\left(A^{\mathrm{T}}A\right)\right]^{\frac{1}{2}}$

$A^{\mathrm{T}}A$ 的特征值：$0=\lambda_1=\cdots=\lambda_{n-r}<\lambda_{n-r+1}\leq\cdots\leq\lambda_n$

A 的奇异值：$0=\sigma_1=\cdots=\sigma_{n-r}<\sigma_{n-r+1}\leq\cdots\leq\sigma_n$

Th24　$\forall V_k\subset\mathbf{R}^n$ 且 $\dim V_k=k$，则有

$$\sigma_k = \min_{V_k}\left\{\max_{0\neq x\in V_k}\frac{\|\boldsymbol{Ax}\|_2}{\|\boldsymbol{x}\|_2}\right\},\quad \sigma_{n-k+1} = \max_{V_k}\left\{\min_{0\neq x\in V_k}\frac{\|\boldsymbol{Ax}\|_2}{\|\boldsymbol{x}\|_2}\right\}$$

证 对 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 应用 **Th23**（取 $\boldsymbol{B}=\boldsymbol{I}$ ）可得

$$\sigma_k = \lambda_k^{\frac{1}{2}} = \left[\min_{V_k}\left\{\max_{0\neq x\in V_k}\frac{\boldsymbol{x}^{\mathrm{T}}\left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\right)\boldsymbol{x}}{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}}\right\}\right]^{\frac{1}{2}} = \left[\min_{V_k}\left\{\max_{0\neq x\in V_k}\left(\frac{\|\boldsymbol{Ax}\|_2}{\|\boldsymbol{x}\|_2}\right)^2\right\}\right]^{\frac{1}{2}} = \text{右}$$

应用 **Th 23′** 可得

$$\sigma_{n-k+1} = \lambda_{n-k+1}^{\frac{1}{2}} = \left[\max_{V_k}\left\{\min_{0\neq x\in V_k}\frac{\boldsymbol{x}^{\mathrm{T}}\left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\right)\boldsymbol{x}}{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}}\right\}\right]^{\frac{1}{2}}$$

$$= \left[\max_{V_k}\left\{\min_{0\neq x\in V_k}\left(\frac{\|\boldsymbol{Ax}\|_2}{\|\boldsymbol{x}\|_2}\right)^2\right\}\right]^{\frac{1}{2}} = \text{右}$$

§5.4　矩阵的直积及应用

$$A=\left(a_{ij}\right)_{m\times n}，\quad B=\left(b_{ij}\right)_{p\times q}，\quad A\otimes B=\begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \cdots & \cdots & & \cdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{bmatrix}$$

例如　$A=\begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix}$，$B=\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$

$$A\otimes B=\left[\begin{array}{cc:cc:cc} 2 & 1 & 4 & 2 & 6 & 3 \\ 1 & 2 & 2 & 4 & 3 & 6 \\ \hdashline 6 & 3 & 4 & 2 & 2 & 1 \\ 3 & 6 & 2 & 4 & 1 & 2 \end{array}\right]，\quad B\otimes A=\left[\begin{array}{ccc:ccc} 2 & 4 & 6 & 1 & 2 & 3 \\ 6 & 4 & 2 & 3 & 2 & 1 \\ \hdashline 1 & 2 & 3 & 2 & 4 & 6 \\ 3 & 2 & 1 & 6 & 4 & 2 \end{array}\right]$$

易见　$A\otimes B\neq B\otimes A$

一、基本性质

(1)　$k(A\otimes B)=(kA)\otimes B=A\otimes(kB)$

(2)　A_1与A_2同阶：$(A_1+A_2)\otimes B=A_1\otimes B+A_2\otimes B$

$$B\otimes(A_1+A_2)=B\otimes A_1+B\otimes A_2$$

(3)　$(A\otimes B)\otimes C=A\otimes(B\otimes C)$

证　$$A\otimes B=\begin{bmatrix} & \cdots\cdots & \\ \cdots & \begin{bmatrix} a_{ij}b_{11} & \cdots & a_{ij}b_{1q} \\ \vdots & & \vdots \\ a_{ij}b_{p1} & \cdots & a_{ij}b_{pq} \end{bmatrix} & \cdots \\ & \cdots\cdots & \end{bmatrix}$$

$$(A\otimes B)\otimes C=\begin{bmatrix} & \cdots\cdots & \\ \cdots & \begin{bmatrix} a_{ij}b_{11}C & \cdots & a_{ij}b_{1q}C \\ \vdots & & \vdots \\ a_{ij}b_{p1}C & \cdots & a_{ij}b_{pq}C \end{bmatrix} & \cdots \\ & \cdots\cdots & \end{bmatrix}$$

$$=\begin{bmatrix} & \cdots & \\ \cdots & [a_{ij}(B\otimes C)] & \cdots \\ & \cdots & \end{bmatrix}=A\otimes(B\otimes C)$$

(4) 设 $A_1=(a_{ij}^{(1)})_{m\times n}, A_2=(a_{ij}^{(2)})_{n\times l}$， $B_1=(b_{ij}^{(1)})_{p\times q}, B_2=(b_{ij}^{(2)})_{q\times r}$ ，则

$$(A_1\otimes B_1)(A_2\otimes B_2)=(A_1A_2)\otimes(B_1B_2)$$

证 $A_1\otimes B_1=\begin{bmatrix} \cdots & \cdots & & \cdots \\ a_{i1}^{(1)}B_1 & a_{i2}^{(1)}B_1 & \cdots & a_{in}^{(1)}B_1 \\ \cdots & \cdots & & \cdots \end{bmatrix}$， $A_2\otimes B_2=\begin{bmatrix} \cdots & a_{1j}^{(2)}B_2 & \cdots \\ \cdots & a_{2j}^{(2)}B_2 & \cdots \\ & \cdots & \\ \cdots & a_{nj}^{(2)}B_2 & \cdots \end{bmatrix}$

$$[左]_{ij块}=\sum_{k=1}^{n}(a_{ik}^{(1)}B_1)(a_{kj}^{(2)}B_2)=\sum_{k=1}^{n}a_{ik}^{(1)}a_{kj}^{(2)}\cdot(B_1B_2)=(A_1A_2)_{ij}(B_1B_2)=[右]_{ij块}$$

(5) $A_{m\times m}$ 与 $B_{n\times n}$ 都可逆 $\Rightarrow (A\otimes B)^{-1}=A^{-1}\otimes B^{-1}$.

(6) $A_{m\times m}$ 与 $B_{n\times n}$ 都是上（下）三角矩阵 $\Rightarrow A\otimes B$ 也是上（下）三角矩阵.

(7) $(A\otimes B)^{\mathrm{H}}=A^{\mathrm{H}}\otimes B^{\mathrm{H}}$

(8) $A_{m\times m}$ 与 $B_{n\times n}$ 都是（酉）正交矩阵 $\Rightarrow A\otimes B$ 也是（酉）正交矩阵.

(9) $\operatorname{rank}(A\otimes B)=\operatorname{rank}A\cdot\operatorname{rank}B$

证　记 $r_1=\operatorname{rank}A, r_2=\operatorname{rank}B$

存在满秩矩阵 P_1,Q_1， 使得 $P_1AQ_1=\begin{bmatrix} I_{r_1} & O \\ O & O \end{bmatrix}\overset{\Delta}{=}A_1$

存在满秩矩阵 P_2,Q_2， 使得 $P_2BQ_2=\begin{bmatrix} I_{r_2} & O \\ O & O \end{bmatrix}\overset{\Delta}{=}B_1$

$(P_1\otimes P_2)(A\otimes B)(Q_1\otimes Q_2)=\cdots=A_1\otimes B_1$

因为 $P_1\otimes P_2$ 与 $Q_1\otimes Q_2$ 都可逆，所以

$\operatorname{rank}(A\otimes B)=\operatorname{rank}(A_1\otimes B_1)=$“$A_1\otimes B_1$中 1 的个数”$=r_1r_2$

直积矩阵的特征值： 二元多项式 $f(x,y)\overset{\Delta}{=}\sum_{i,j=0}^{l}c_{ij}x^iy^j$，对于 $A_{m\times m}$ 和 $B_{n\times n}$，

$$\text{定义矩阵 } f(A,B)=\sum_{i,j=0}^{l}c_{ij}A^i\otimes B^j .$$

Th27　设 $A_{m\times m}$ 的特征值为 $\lambda_1,\cdots,\lambda_m$， $B_{n\times n}$ 的特征值为 $\mu_1,\cdots,\mu_n$，则 $f(A,B)$ 的特征值为 $f(\lambda_i,\mu_j),\ i=1,\cdots,m,j=1,\cdots,n$.

证　对于 $A_{m\times m}$ 和 $B_{n\times n}$，存在可逆矩阵 $P_{m\times m}$ 与 $Q_{n\times n}$，使得

$$P^{-1}AP=\begin{bmatrix}\lambda_1 & \cdots & * \\ & \ddots & \vdots \\ & & \lambda_m\end{bmatrix}\overset{\Delta}{=}T_1,\quad Q^{-1}BQ=\begin{bmatrix}\mu_1 & \cdots & * \\ & \ddots & \vdots \\ & & \mu_n\end{bmatrix}\overset{\Delta}{=}T_2$$

因为 $P\otimes Q$ 可逆，所以

$$(P\otimes Q)^{-1}(A^i\otimes B^j)(P\otimes Q)=\cdots=T_1^i\otimes T_2^j \quad \text{是上三角矩阵}$$

$$(P\otimes Q)^{-1}\cdot f(A,B)\cdot(P\otimes Q)=\cdots=f(T_1,T_2) \quad \text{也是上三角矩阵}$$

又　$$T_1^i\otimes T_2^j=\begin{bmatrix}\lambda_1^iT_2^j & \cdots & * \\ & \ddots & \vdots \\ & & \lambda_m^iT_2^j\end{bmatrix},\quad \lambda_k^iT_2^j=\begin{bmatrix}\lambda_k^i\mu_1^j & \cdots & * \\ & \ddots & \vdots \\ & & \lambda_k^i\mu_n^j\end{bmatrix}$$

故 $f(T_1,T_2)$ 的主对角线元素为 $f(\lambda_k,\mu_s),\ k=1,\cdots,m,\ s=1,\cdots,n$.

由此可得： $f(A,B)$ 的特征值为 $f(\lambda_i,\mu_j),\ i=1,\cdots,m,j=1,\cdots,n$.

推论 1　$A_{m\times m},B_{n\times n}$， $A\otimes B$ 的特征值为 $\lambda_i\mu_j,i=1,\cdots,m,j=1,\cdots,n$.

取 $f(x,y)=x^1y^1$，则 $f(A,B)=A\otimes B$.

推论 2　$A_{m\times m},B_{n\times n}$：　$|A\otimes B|=|A|^n|B|^m$.

$$\text{左}=\prod_{i=1}^{m}\prod_{j=1}^{n}\lambda_i\mu_j=\prod_{i=1}^{m}\left(\lambda_i^n\cdot\prod_{j=1}^{n}\mu_j\right)=\prod_{i=1}^{m}\lambda_i^n\cdot\left(\prod_{j=1}^{n}\mu_j\right)^m=\text{右}$$

二、线性矩阵方程的可解性

$$A=\left(a_{ij}\right)_{m\times p},\ B_{q\times n},\quad X_{p\times q}=\begin{bmatrix}x_1^{\mathrm{T}}\\ \vdots\\ x_p^{\mathrm{T}}\end{bmatrix},\ \overline{\mathrm{vec}}(X)\overset{\Delta}{=}\begin{bmatrix}x_1\\ \vdots\\ x_p\end{bmatrix}=\overrightarrow{X}$$

$$AXB=\begin{bmatrix}a_{11}&\cdots&a_{1p}\\ \vdots&&\vdots\\ a_{m1}&\cdots&a_{mp}\end{bmatrix}\begin{bmatrix}x_1^{\mathrm{T}}\\ \vdots\\ x_p^{\mathrm{T}}\end{bmatrix}B=\begin{bmatrix}(a_{11}x_1^{\mathrm{T}}+\cdots+a_{1p}x_p^{\mathrm{T}})B\\ \vdots\\ (a_{m1}x_1^{\mathrm{T}}+\cdots+a_{mp}x_p^{\mathrm{T}})B\end{bmatrix}$$

$$\overrightarrow{AXB}=\begin{bmatrix}B^{\mathrm{T}}(a_{11}x_1+\cdots+a_{1p}x_p)\\ \vdots\\ B^{\mathrm{T}}(a_{m1}x_1+\cdots+a_{mp}x_p)\end{bmatrix}=\begin{bmatrix}a_{11}B^{\mathrm{T}}&\cdots&a_{1p}B^{\mathrm{T}}\\ \vdots&&\vdots\\ a_{m1}B^{\mathrm{T}}&\cdots&a_{mp}B^{\mathrm{T}}\end{bmatrix}\begin{bmatrix}x_1\\ \vdots\\ x_p\end{bmatrix}$$

$$=\left(A\otimes B^{\mathrm{T}}\right)\overrightarrow{X}$$

复习：划分 $A=(a_1,a_2,\cdots,a_n)$，则有

(1) $R(A)=\mathrm{span}\{a_1,a_2,\cdots,a_n\}$；

(2) $Ax=b$ 有解 $\Leftrightarrow (a_1,\cdots,a_n)\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix}=b$ 有解 $\Leftrightarrow b\in R(A)$.

Th28 $A_i\in \mathrm{C}^{m\times p},B_i\in \mathrm{C}^{q\times n},F\in \mathrm{C}^{m\times n}$：

$$\sum_{i=1}^{k}A_iXB_i=F\text{ 有解 }X_{p\times q}\Leftrightarrow \overrightarrow{F}\in R\left(\sum_{i=1}^{k}A_i\otimes B_i^{\mathrm{T}}\right)$$

证　左 $\Leftrightarrow \sum\overrightarrow{A_iXB_i}=\overrightarrow{F}$ 有解 $X\Leftrightarrow \sum\left(A_i\otimes B_i^{\mathrm{T}}\right)\overrightarrow{X}=\overrightarrow{F}$ 有解 $\overrightarrow{X}$

$\Leftrightarrow \left(\sum A_i\otimes B_i^{\mathrm{T}}\right)\overrightarrow{X}=\overrightarrow{F}$ 有解 $\overrightarrow{X}$

$\Leftrightarrow \overrightarrow{F}\in R\left(\sum A_i\otimes B_i^{\mathrm{T}}\right)$

Th29 $A_{m\times m},B_{n\times n},F_{m\times n}$，矩阵方程 $AX+XB=F$ 有唯一解 $X_{m\times n}\Leftrightarrow$

$\lambda_i(A)+\mu_j(B)\neq 0\quad (i=1,\cdots,m,\ j=1,\cdots,n)$

证　左 $\Leftrightarrow AXI_n+I_mXB=F$ 有唯一解 $X_{m\times n}$

$$\Leftrightarrow \left(A\otimes I_n + I_m \otimes B^{\mathrm{T}}\right)\vec{X} = \vec{F} \text{ 有唯一解 } \vec{X}$$

$$\Leftrightarrow \det\left(A\otimes I_n + I_m \otimes B^{\mathrm{T}}\right)\neq 0$$

令 $f(x,y)=x^1y^0+x^0y^1$，并注意 $\mu(B^{\mathrm{T}})=\mu(B)$，由定理 27 知

$$f\left(A,B^{\mathrm{T}}\right)=A\otimes I_n + I_m\otimes B^{\mathrm{T}}$$

的特征值为 $f(\lambda_i,\mu_j)=\lambda_i+\mu_j$，因此可得

$$\det\left(A\otimes I_n + I_m\otimes B^{\mathrm{T}}\right)\neq 0 \Leftrightarrow \lambda_i+\mu_j\neq 0 \quad (i=1,\cdots,m,j=1,\cdots,n)$$

推论 $A_{m\times m},B_{n\times n},AX+XB=O_{m\times n}$ 有非零解 $X\Leftrightarrow \exists i_0,j_0,\mathrm{st}.\,\lambda_{i_0}+\mu_{j_0}=0$.

例 1 $A_{m\times m},B_{n\times n},C_{m\times n}$： $\lambda(A)\neq\mu(B)\Rightarrow\begin{bmatrix}A & C\\ O & B\end{bmatrix}$ 相似于 $\begin{bmatrix}A & O\\ O & B\end{bmatrix}$.

证 设 $P=\begin{bmatrix}I_m & X\\ O & I_n\end{bmatrix}$ ($X_{m\times n}$ 待定)，则有

$$P^{-1}=\begin{bmatrix}I_m & -X\\ O & I_n\end{bmatrix},\quad P\begin{bmatrix}A & C\\ O & B\end{bmatrix}P^{-1}=\begin{bmatrix}A & C+XB-AX\\ O & B\end{bmatrix}$$

$$\lambda(A)\neq\mu(B)\Rightarrow\lambda_i(A)+\mu_j(-B)\neq 0$$

$$\Rightarrow AX+X(-B)=C \text{ 有唯一解 } X^*_{m\times n}$$

$$\Rightarrow \text{存在 } P=\begin{bmatrix}I_m & X^*\\ O & I_n\end{bmatrix} \text{使得 } P\begin{bmatrix}A & C\\ O & B\end{bmatrix}P^{-1}=\begin{bmatrix}A & O\\ O & B\end{bmatrix}$$

[注] 上例中，$\begin{bmatrix}A & C\\ O & B\end{bmatrix}$ 相似于 $\begin{bmatrix}A & O\\ O & B\end{bmatrix}\Leftrightarrow\vec{C}\in R\left(A\otimes I_n - I_m\otimes B^{\mathrm{T}}\right)$

Th30 $A_{m\times m}$ 的特征值为 $\lambda_1,\cdots,\lambda_m$，$B_{n\times n}$ 的特征值为 $\mu_1,\cdots,\mu_n$.

(1) $\sum\limits_{k=0}^{l}A^kXB^k=F$ 有唯一解 $\Leftrightarrow 1+\left(\lambda_i\mu_j\right)+\cdots+\left(\lambda_i\mu_j\right)^l\neq 0\ (\forall i,j)$；

(2) $\sum\limits_{k=0}^{l}A^kXB^k=O$ 有非零解 $\Leftrightarrow \exists i_0,j_0,\mathrm{st}.1+\left(\lambda_{i_0}\mu_{j_0}\right)+\cdots+\left(\lambda_{i_0}\mu_{j_0}\right)^l=0$.

证 将 $\sum\limits_{k=0}^{l}A^kXB^k=F$ 按行拉直可得 $\left(\sum\limits_{k=0}^{l}A^k\otimes(B^{\mathrm{T}})^k\right)\vec{X}=\vec{F}$

$$f(x,y)\overset{\Delta}{=}\sum_k x^k y^k \Rightarrow f(A,B^{\mathrm{T}})=\sum_{k=0}^{l} A^k\otimes(B^{\mathrm{T}})^k$$

$f(A,B^{\mathrm{T}})$的特征值为 $f(\lambda_i,\mu_j)=\sum_{k=0}^{l}\lambda_i^k\mu_j^k=\sum_{k=0}^{l}(\lambda_i\mu_j)^k$

故　$\sum_{k=0}^{l}A^kXB^k=F$ 有唯一解 $\Leftrightarrow f(\lambda_i,\mu_j)\neq 0$

$\sum_{k=0}^{l}A^kXB^k=O$ 有非零解 $\Leftrightarrow \exists i_0,j_0,\mathrm{st.}\ f(\lambda_{i_0},\mu_{j_0})=0$

三、线性矩阵方程的矩阵函数解法

引理 3　设 $A_{m\times m},B_{n\times n},F_{m\times n}$，若 $\mathrm{Re}(\lambda_A)<0,\mathrm{Re}(\mu_B)<0$，

则广义积分 $\int_0^{+\infty}e^{At}Fe^{Bt}dt$ 存在.

Th31　$A_{m\times m},B_{n\times n},F_{m\times n}$：$\lambda_A+\mu_B\neq 0$ 且 $\int_0^{+\infty}e^{At}Fe^{Bt}dt$ 存在 $\Rightarrow$

$AX+XB=F$ 存在唯一解 $X=-\int_0^{+\infty}e^{At}Fe^{Bt}dt$.

证　$Y(t)\overset{\Delta}{=}e^{At}Fe^{Bt}$：$Y(t)\big|_{t=0}=F$，$\int_0^{+\infty}Y(t)dt$ 存在 $\Rightarrow \lim\limits_{t\to+\infty}Y(t)=O$

$$\frac{dY}{dt}=A\cdot Y(t)+Y(t)\cdot B \qquad ①$$

积分①：$Y(t)\big|_0^{+\infty}=A\cdot\int_0^{+\infty}Y(t)dt+\int_0^{+\infty}Y(t)dt\cdot B$

$$-F=A(-X)+(-X)B\Rightarrow AX+XB=F$$

推论 1　设 $A_{m\times m},B_{n\times n},F_{m\times n}$，$\mathrm{Re}(\lambda_A)<0,\mathrm{Re}(\mu_B)<0$，则 $AX+XB=F$

存在唯一解 $X=-\int_0^{+\infty}e^{At}Fe^{Bt}dt$.

推论 2　设 $A_{m\times m},F_{m\times m},\mathrm{Re}(\lambda_A)<0$，则 $A^{\mathrm{H}}X+XA=-F$ 存在唯一解

$X=\int_0^{+\infty}e^{A^{\mathrm{H}}t}Fe^{At}dt$；若 F 是正定矩阵，则 X 也是正定矩阵.

（$\lambda(A^{\mathrm{H}})=\overline{\lambda(A)}\Rightarrow \mathrm{Re}(\lambda(A^{\mathrm{H}}))<0$）

证　②　$0\neq x\in\mathrm{C}^n$：　e^{At} 可逆 $\Rightarrow e^{At}x\neq 0$

$$\left(e^{At}x\right)^{\mathrm{H}}F\left(e^{At}x\right)>0\Rightarrow x^{\mathrm{H}}\cdot e^{A^{\mathrm{H}}t}Fe^{At}\cdot x>0$$

$$\Rightarrow x^{\mathrm{H}}Xx=\int_0^{+\infty}x^{\mathrm{H}}\left(e^{A^{\mathrm{H}}t}Fe^{At}\right)xdt>0$$

典型题分析：

例 2　设 $A\in\mathbf{C}^{n\times n}$ 的 n 个线性无关的特征向量为 $x_1,x_2,\cdots,x_n$，证明：$A\otimes A$ 有 n^2 个线性无关的特征向量，并将它们构造出来.

证　设 $Ax_i=\lambda_i x_i\ (i=1,2,\cdots,n)$，$P=\begin{bmatrix}x_1 & x_2 & \cdots & x_n\end{bmatrix}$，则有

$$P^{-1}AP=\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)\overset{\Delta}{=}\Lambda$$

$$(P\otimes P)^{-1}(A\otimes A)(P\otimes P)=(P^{-1}\otimes P^{-1})(A\otimes A)(P\otimes P)$$

$$=(P^{-1}AP)\otimes(P^{-1}AP)=\Lambda\otimes\Lambda$$

因为 $\Lambda\otimes\Lambda$ 是对角矩阵，所以 $P\otimes P$ 的 n^2 个列向量是 $A\otimes A$ 的线性无关的特征向量. 根据矩阵直积的定义，有

$$x_i\otimes P=x_i\otimes\begin{bmatrix}x_1 & \cdots & x_n\end{bmatrix}=\begin{bmatrix}x_i\otimes x_1 & \cdots & x_i\otimes x_n\end{bmatrix}$$

$$P\otimes P=\begin{bmatrix}x_1 & \cdots & x_n\end{bmatrix}\otimes P=\begin{bmatrix}x_1\otimes P & \cdots & x_n\otimes P\end{bmatrix}$$

$$=\begin{bmatrix}x_1\otimes x_1 & \cdots & x_1\otimes x_n & x_2\otimes x_1 & \cdots & x_n\otimes x_n\end{bmatrix}$$

故 $x_i\otimes x_j\ (i,j=1,2,\cdots,n)$ 是 $A\otimes A$ 的 n^2 个线性无关的特征向量.

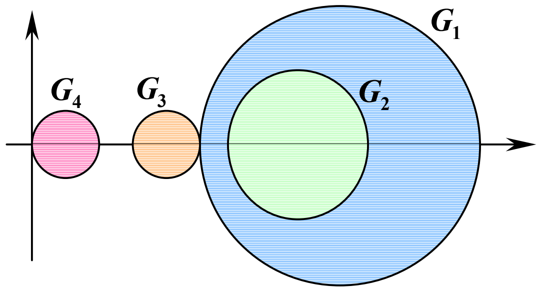
例 3　应用盖尔圆定理说明 $A=\begin{bmatrix}9 & 1 & -2 & 1\\ 0 & 8 & 1 & 1\\ -1 & 0 & 4 & 0\\ 1 & 0 & 0 & 1\end{bmatrix}$ 至少有两个实特征值.

解　$G_1:|z-9|\le 4$

$G_2:|z-8|\le 2$

$G_3:|z-4|\le 1$

$G_4:|z-1|\le 1$



两个连通部分关于实轴对称：

$S_1 = G_4$ 中只有 A 的一个特征值，该特征值为实数；

$S_2 = G_1 \cup G_2 \cup G_3$ 中有 A 的三个特征值，其中至少有一个特征值为实数.

[注] 实矩阵的复特征值一定成对共轭出现.

第六章　广义逆矩阵

§6.1 投影变换

一、投影变换：向量空间$\mathbf{C}^n$中，子空间L与M满足$\mathbf{C}^n = L \oplus M$，对$\forall x \in \mathbf{C}^n$，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的投影变换，称y为x沿着M到L的投影．

性质(1)　$T_{L,M}$是线性变换．

证　$\forall x_1 \in \mathbf{C}^n, x_1 = y_1 + z_1, y_1 \in L, z_1 \in M$唯一 $\Rightarrow T_{L,M}(x_1) = y_1$

$\forall x_2 \in \mathbf{C}^n, x_2 = y_2 + z_2, y_2 \in L, z_2 \in M$唯一 $\Rightarrow T_{L,M}(x_2) = y_2$

因为 $x_1 + x_2 = (y_1 + y_2) + (z_1 + z_2),\ y_1 + y_2 \in L,\ z_1 + z_2 \in M$唯一

所以 $T(x_1 + x_2) = y_1 + y_2 = T(x_1) + T(x_2)$

又 $kx_1 = (ky_1) + (kz_1)$， $ky_1 \in L, kz_1 \in M$唯一

所以 $T(kx_1) = ky_1 = k \cdot T(x_1)$

性质(2)　$R(T_{L,M}) = L,\ N(T_{L,M}) = M$．

证 (1) $\forall x \in L \subset \mathbf{C}^n$： $x = x + \theta$， $x \in L, \theta \in M$唯一

故 $x = T_{L,M}(x) \in R(T_{L,M})$，即$L \subset R(T_{L,M})$；

$\forall x \in \mathbf{C}^n$，由$T_{L,M}(x) \in L$可得$R(T_{L,M}) \subset L$．故第一式成立．

(2) $\forall x \in M \subset \mathbf{C}^n$： $x = \theta + x, \theta \in L, x \in M$唯一

故 $T_{L,M}(x) = \theta \Rightarrow x \in N(T_{L,M})$，即$M \subset N(T_{L,M})$；

$\forall x \in N(T_{L,M})$，若$x \notin M$，则有

$$x = y + z,\quad y \in L, z \in M \text{唯一，且 } y \neq \theta$$

于是$T_{L,M}(x) = y \neq \theta$，矛盾．因此$x \in M$，从而$N(T_{L,M}) \subset M$．

故第二式成立．

性质(3)　$\forall x \in L \Rightarrow T_{L,M}(x) = x$，　$\forall x \in M \Rightarrow T_{L,M}(x) = \theta$.

证　$\forall x \in L \subset \mathbf{C}^n$：　$x = x + \theta$，$x \in L, \theta \in M$ 唯一　$\Rightarrow T_{L,M}(x) = x$

$\forall x \in M \subset \mathbf{C}^n$：　$x = \theta + x$，$\theta \in L, x \in M$ 唯一　$\Rightarrow T_{L,M}(x) = \theta$

[注]　$L = R(T_{L,M})$是$T_{L,M}$的不变子空间$\Rightarrow T_{L,M}$是L中的单位变换

$M = N(T_{L,M})$是$T_{L,M}$的不变子空间$\Rightarrow T_{L,M}$是M中的零变换

二、投影矩阵

取线性空间$\mathbf{C}^n$的基为$e_1, \cdots, e_n$时，元素x与它的坐标“形式一致”.

称$T_{L,M}$在该基下的矩阵为投影矩阵，记作$P_{L,M}$.

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

引理 1　$A_{n\times n}, A^2 = A \Rightarrow N(A) = R(I - A)$.

证　$A^2 = A \Rightarrow A(I - A) = O$.

先证$R(I - A) \subset N(A)$：

$\forall x \in R(I - A) \Rightarrow \exists u \in \mathbf{C}^n, \mathrm{st}. x = (I - A)u$

$Ax = A(I - A)u = \theta \Rightarrow x \in N(A)$

再证$N(A) \subset R(I - A)$：　$\forall \alpha \in N(A) \Rightarrow A\alpha = \theta$

$\alpha = \alpha - A\alpha = (I - A)\alpha \in R(I - A)$

故$N(A) = R(I - A)$.

Th1　方阵$P = P_{L,M} \Leftrightarrow P^2 = P$.

证　必要性.　$\mathbf{C}^n = L \oplus M$

$\forall x \in \mathbf{C}^n$，$x = y + z, y \in L, z \in M$ 唯一 $\Rightarrow P_{L,M}x = y$

$$P_{L,M}^2 x = P_{L,M}(P_{L,M}x) = P_{L,M}y \overset{\text{性质 (4)}}{=} y = P_{L,M}x$$

$$P^2 x = Px \Rightarrow P^2(e_1, \cdots, e_n) = P(e_1, \cdots, e_n) \Rightarrow P^2 = P$$

充分性. $\forall x \in \mathbf{C}^n \Rightarrow x = Px + (I - P)x$

令 $y = Px \in R(P)$, $z = (I - P)x \in R(I - P) \overset{\text{引1}}{=} N(P)$

则 $\mathbf{C}^n = R(P) + N(P)$. 下证 $R(P) \cap N(P) = \{\theta\}$:

对 $\forall \beta \in R(P) \cap N(P)$, $\beta \in R(P) \Rightarrow \exists u \in \mathbf{C}^n$, st. $\beta = Pu$

$$\beta \in N(P) \Rightarrow P\beta = \theta$$

故 $\beta = Pu = P^2u = P \cdot Pu = P\beta = \theta$.

于是可得$\mathbf{C}^n = R(P) \oplus N(P)$, 从而有

$$x = y + z,\ y = Px \in R(P),\ z = (I - P)x \in N(P) \text{ 唯一}$$

因为投影变换$T_{R(P),N(P)}$满足

$$T_{R(P),N(P)}(x) = y \quad \Rightarrow P_{R(P),N(P)}x = y \quad (\forall x \in \mathbf{C}^n)$$

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 $y_1, \cdots, y_{n-r}$: $Y = (y_1, \cdots, y_{n-r})$

$$\left.\begin{array}{l} P_{L,M}x_i = x_i \Rightarrow P_{L,M}X = X \\ P_{L,M}y_j = \theta \Rightarrow P_{L,M}Y = O \end{array}\right\} \Rightarrow P_{L,M}(X \mid Y) = (X \mid O)$$

$$\Rightarrow P_{L,M} = (X \mid O) \cdot (X \mid Y)^{-1}$$

例 1 $\mathbf{R}^2$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 $\mathbf{C}^n$ 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 $P_L=P_{L,M}$.

Th2　方阵 $P=P_L \Leftrightarrow P^2=P,\ P^{\mathrm{H}}=P$.

证　必要性．已知 $P=P_L$：Th1 $\Rightarrow P^2=P$

$$\left.\begin{aligned}&\forall x_1\in\mathbf{C}^n\Rightarrow x_1=y_1+z_1,\ \ y_1\in L,\ z_1\in L^{\perp}\\&\forall x_2\in\mathbf{C}^n\Rightarrow x_2=y_2+z_2,\ \ y_2\in L,\ z_2\in L^{\perp}\end{aligned}\right\}$$

$$\left.\begin{aligned}&P_Lx_1=y_1\in L,\ \ (I-P_L)x_1=x_1-y_1=z_1\in L^{\perp}\\&P_Lx_2=y_2\in L,\ \ (I-P_L)x_2=x_2-y_2=z_2\in L^{\perp}\end{aligned}\right\}$$

$$\left.\begin{aligned}&P_Lx_1\perp(I-P_L)x_2\Rightarrow x_1^{\mathrm{H}}P_L^{\mathrm{H}}(I-P_L)x_2=0\\&(I-P_L)x_1\perp P_Lx_2\Rightarrow x_1^{\mathrm{H}}(I-P_L)^{\mathrm{H}}P_Lx_2=0\end{aligned}\right\}$$

故　$x_1^{\mathrm{H}}(P_L^{\mathrm{H}}-P_L)x_2=0\Rightarrow P_L^{\mathrm{H}}-P_L=O$：$P^{\mathrm{H}}=P$

充分性．已知 $P^2=P$：　Th1 $\Rightarrow P=P_{R(P),N(P)}$

$$P^{\mathrm{H}}=P:\quad N(P)=N(P^{\mathrm{H}})=R^{\perp}(P)\qquad\text{（Th1-35）}$$

故 $P=P_{R(P)}$.

五、正交投影矩阵的确定方法

$$\left.\begin{array}{ll} L\text{ 的基 } x_1,\cdots,x_r: & X=(x_1,\cdots,x_r) \\ L^{\perp}\text{ 的基 } y_1,\cdots,y_{n-r}: & Y=(y_1,\cdots,y_{n-r}) \end{array}\right\} \Rightarrow \begin{cases} X^{\mathrm{H}}Y=O \\ Y^{\mathrm{H}}X=O \end{cases}$$

已求得 $P_L=P_{L,L^{\perp}}=(X\,|\,O)\cdot(X\,|\,Y)^{-1}$

因为 $(X\,|\,Y)^{\mathrm{H}}\cdot(X\,|\,Y)=\begin{pmatrix} X^{\mathrm{H}} \\ Y^{\mathrm{H}} \end{pmatrix}\cdot(X\,|\,Y)=\begin{bmatrix} X^{\mathrm{H}}X & O \\ O & Y^{\mathrm{H}}Y \end{bmatrix}$

所以 $(X\,|\,Y)^{-1}=\begin{bmatrix} (X^{\mathrm{H}}X)^{-1} & O \\ O & (Y^{\mathrm{H}}Y)^{-1} \end{bmatrix}\cdot(X\,|\,Y)^{\mathrm{H}}=\begin{bmatrix} (X^{\mathrm{H}}X)^{-1}X^{\mathrm{H}} \\ (Y^{\mathrm{H}}Y)^{-1}Y^{\mathrm{H}} \end{bmatrix}$

于是 $P_L=(X\,|\,O)\cdot\begin{bmatrix} (X^{\mathrm{H}}X)^{-1}X^{\mathrm{H}} \\ (Y^{\mathrm{H}}Y)^{-1}Y^{\mathrm{H}} \end{bmatrix}=X\cdot(X^{\mathrm{H}}X)^{-1}\cdot X^{\mathrm{H}}$

例 3　向量空间 $\mathbf{R}^3$ 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　$X=\begin{bmatrix}1 & 0\\2 & 1\\0 & 1\end{bmatrix}, X^{\mathrm{T}}X=\begin{bmatrix}5 & 2\\2 & 2\end{bmatrix}, (X^{\mathrm{T}}X)^{-1}=\dfrac{1}{6}\begin{bmatrix}2 & -2\\-2 & 5\end{bmatrix}$

$$P_L=X\cdot(X^{\mathrm{T}}X)^{-1}X^{\mathrm{T}}=\frac{1}{6}\begin{bmatrix}2 & 2 & -2\\2 & 5 & 1\\-2 & 1 & 5\end{bmatrix}$$

[注]　正交投影矩阵 P_L 与子空间 L 的基的选择无关.

§6.2-3　广义逆矩阵

一、定义与算法

1． 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)　$(AX)^{\mathrm{H}} = AX$　　　(4)　$(XA)^{\mathrm{H}} = XA$

称 X 为 A 的 M-P 逆，记作 A^+．(Moore 1920, Penrose1955)

例如　$A_{n\times n}$ 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：$O^+_{m\times n} = O_{n\times m}$

$A = \begin{bmatrix}1\\1\end{bmatrix}, X = \left(\frac{1}{2}, \frac{1}{2}\right)$ 满足 P-方程：$A^+ = \left(\frac{1}{2}, \frac{1}{2}\right)$

例 4　$F \in \mathrm{C}_r^{m\times r}\ (r \geq 1) \Rightarrow F^+ = \left(F^{\mathrm{H}}F\right)^{-1}F^{\mathrm{H}}$，且 $F^+F = I_r$；

$G \in \mathrm{C}_r^{r\times n}\ (r \geq 1) \Rightarrow G^+ = G^{\mathrm{H}}\left(GG^{\mathrm{H}}\right)^{-1}$，且 $GG^+ = I_r$．

证　验证第一式：令 $X = \left(F^{\mathrm{H}}F\right)^{-1}F^{\mathrm{H}}$，则有

$$FXF = F\cdot\left(F^{\mathrm{H}}F\right)^{-1}F^{\mathrm{H}}F = F$$

$$XFX = \left(F^{\mathrm{H}}F\right)^{-1}F^{\mathrm{H}}F\cdot X = X$$

$$(FX)^{\mathrm{H}} = X^{\mathrm{H}}F^{\mathrm{H}} = F\left(F^{\mathrm{H}}F\right)^{-1}F^{\mathrm{H}} = FX$$

$$(XF)^{\mathrm{H}} = I_r^{\mathrm{H}} = I_r = XF$$

Th3　$\forall A_{m\times n}$，A^+ 存在且唯一．

证　存在性．$A = O_{m\times n} \Rightarrow A^+ = O_{n\times m}$

$$A \neq O \Rightarrow \mathrm{rank}A \geq 1:\quad A = FG,\ F \in \mathrm{C}_r^{m\times r},\ G \in \mathrm{C}_r^{r\times n}$$

令 $X = G^+F^+$，则有

$$AXA = FG\cdot G^+F^+\cdot FG = FG = A$$

$$XAX = G^+F^+\cdot FG\cdot G^+F^+ = G^+F^+ = X$$

$$(AX)^{\mathrm{H}}=(FG\cdot G^{+}F^{+})^{\mathrm{H}}=(FF^{+})^{\mathrm{H}}=FF^{+}=F\cdot GG^{+}\cdot F^{+}=AX$$

$$(XA)^{\mathrm{H}}=(G^{+}F^{+}\cdot FG)^{\mathrm{H}}=(G^{+}G)^{\mathrm{H}}=G^{+}G=G^{+}\cdot F^{+}F\cdot G=XA$$

故 $A^{+}=G^{+}F^{+}=G^{\mathrm{H}}(F^{\mathrm{H}}AG^{\mathrm{H}})^{-1}F^{\mathrm{H}}$.

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

$$\begin{aligned}X&=XAX=X\cdot AYA\cdot X=X\cdot(AY)^{\mathrm{H}}\cdot(AX)^{\mathrm{H}}\\&=X\cdot(AXAY)^{\mathrm{H}}=X\cdot(AY)^{\mathrm{H}}=XAY=X\cdot AYA\cdot Y\\&=(XA)^{\mathrm{H}}\cdot(YA)^{\mathrm{H}}\cdot Y=(YAXA)^{\mathrm{H}}\cdot Y=(YA)^{\mathrm{H}}\cdot Y=YAY=Y\end{aligned}$$

例 5　设 $A\in \mathrm{C}_r^{m\times n}$ 的奇异值分解为 $A=U\begin{bmatrix}\Sigma_r & O\\ O & O\end{bmatrix}_{m\times n}V^{\mathrm{H}}$，则

$$A^{+}=V\begin{bmatrix}\Sigma_r^{-1} & O\\ O & O\end{bmatrix}_{n\times m}U^{\mathrm{H}}.$$　（直接验证即得）

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：$A\{1\}, A\{1,2\}, A\{1,3\}, A\{1,4\}$ 及 A^{+}.

2．求 $A^{(1)}$ 与 $A^{(1,2)}$ 的初等变换方法

$A\in \mathrm{C}_r^{m\times n}$，　$A\xrightarrow{\text{行}}B\Rightarrow$　∃可逆矩阵 $Q_{m\times m}$, st. $QA=B$

其中 B 为拟 Hermite 标准形，它的后 $m-r$ 行元素全为零.

$$B\xrightarrow{\text{列对换}}\begin{bmatrix}I_r & K\\ O & O\end{bmatrix}_{m\times n}=C\Rightarrow \exists\text{置换矩阵}P_{n\times n},\ \mathrm{st.}\ BP=C$$

于是 $QAP=C\Rightarrow A=Q^{-1}CP^{-1}$.

Th14　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　$AXA = Q^{-1}CP^{-1}\cdot P\begin{bmatrix} I_r & 0 \\ 0 & L \end{bmatrix}_{n\times m} Q\cdot Q^{-1}CP^{-1}$

$$= Q^{-1}\begin{bmatrix} I_r & K \\ O & O \end{bmatrix}_{m\times n}\begin{bmatrix} I_r & O \\ O & L \end{bmatrix}_{n\times m}\begin{bmatrix} I_r & K \\ O & O \end{bmatrix}_{m\times n} P^{-1}$$

$$= Q^{-1}\begin{bmatrix} I_r & KL \\ O & O \end{bmatrix}_{m\times m}\begin{bmatrix} I_r & K \\ O & O \end{bmatrix}_{m\times n} P^{-1}$$

$$= Q^{-1}\begin{bmatrix} I_r & K \\ O & O \end{bmatrix}_{m\times n} P^{-1} = A$$

故 $X \in A\{1\}$；显然 $AX_0A = A$，且有

$$X_0AX_0 = P\begin{bmatrix} I_r & O \\ O & O \end{bmatrix}Q\cdot Q^{-1}\begin{bmatrix} I_r & K \\ O & O \end{bmatrix}P^{-1}\cdot P\begin{bmatrix} I_r & O \\ O & O \end{bmatrix}Q = \cdots = X_0$$

故 $X_0 \in A\{1,2\}$.

例 6　$A = \begin{bmatrix} 2 & 1 & 0 & 2 \\ 0 & 0 & 1 & 2 \\ 2 & 1 & 1 & 4 \end{bmatrix}$，求 $A^{(1)}, A^{(1,2)}$ 及 A^{+}.

解　$(A|\,I) \to \left[\begin{array}{cccc|ccc} 2 & 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & -1 & -1 & 1 \end{array}\right]$：$c_1 = 2, c_2 = 3$

$$Q = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & -1 & 1 \end{bmatrix},\quad P = (e_2, e_3, e_1, e_4) = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

$$A^{(1)} = P\left[\begin{array}{cc:c} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \hdashline 0 & 0 & a \\ 0 & 0 & b \end{array}\right]Q = \begin{bmatrix} -a & -a & a \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ -b & -b & b \end{bmatrix},\quad A^{(1,2)} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

$A = FG$：

$$F=\begin{bmatrix}1&0\\0&1\\1&1\end{bmatrix},\ F^{\mathrm{T}}F=\begin{bmatrix}2&1\\1&2\end{bmatrix},\ F^{+}=\frac{1}{3}\begin{bmatrix}2&-1&1\\-1&2&1\end{bmatrix}$$

$$G=\begin{bmatrix}2&1&0&2\\0&0&1&2\end{bmatrix},\ GG^{\mathrm{T}}=\begin{bmatrix}9&4\\4&5\end{bmatrix},\ G^{+}=\frac{1}{29}\begin{bmatrix}10&-8\\5&-4\\-4&9\\2&10\end{bmatrix}$$

$$A^{+}=G^{+}F^{+}=\frac{1}{87}\begin{bmatrix}28&-26&2\\14&-13&1\\-17&22&5\\-6&18&12\end{bmatrix}$$

例 7　设 $A_{m\times n}\neq O$，且 A^{+} 已知，记 $B=\begin{pmatrix}A\\A\end{pmatrix}$，求 B^{+}．

解　$\mathrm{rank}A=r\geq 1\Rightarrow A=FG$：　$F\in \mathrm{C}_r^{m\times r}$，$G\in \mathrm{C}_r^{r\times n}$

$$B=\begin{pmatrix}FG\\FG\end{pmatrix}=\begin{pmatrix}F\\F\end{pmatrix}G:\quad \begin{pmatrix}F\\F\end{pmatrix}\in \mathrm{C}_r^{2m\times r},\ G\in \mathrm{C}_r^{r\times n}$$

$$B^{+}=G^{+}\begin{pmatrix}F\\F\end{pmatrix}^{+}=G^{+}\cdot\left[\left(F^{\mathrm{H}}\middle|F^{\mathrm{H}}\right)\begin{pmatrix}F\\F\end{pmatrix}\right]^{-1}\left(F^{\mathrm{H}}\middle|F^{\mathrm{H}}\right)$$

$$=G^{+}\cdot\frac{1}{2}\left(F^{\mathrm{H}}F\right)^{-1}\cdot\left(F^{\mathrm{H}}\middle|F^{\mathrm{H}}\right)=\frac{1}{2}\left(G^{+}F^{+}\middle|G^{+}F^{+}\right)=\frac{1}{2}\left(A^{+}\middle|A^{+}\right)$$

二、性质

Th4　$A_{m\times n}$, $A^{(1)}$唯一$\Leftrightarrow m=n$, A可逆，且 $A^{(1)}=A^{-1}$．

证　设 $X_{n\times m}\in A\{1\}$：

(1)　划分 $X=(x_1,\cdots,x_m)$，$\forall x\in N(A)$，令 $Y=(x_1+x,x_2,\cdots,x_m)$，则

$$AYA=A\cdot[X+(x,\theta,\cdots,\theta)]\cdot A=AXA=A\Rightarrow Y\in A\{1\}$$

(2)　划分 $X=\begin{bmatrix}\alpha_1\\ \vdots\\ \alpha_n\end{bmatrix}$, $\forall\beta\in N(A^{\mathrm H})$, 令 $Z=\begin{bmatrix}\alpha_1+\beta^{\mathrm H}\\ \alpha_2\\ \vdots\\ \alpha_n\end{bmatrix}$, 则

$$AZA=\left[A^{\mathrm H}Z^{\mathrm H}A^{\mathrm H}\right]^{\mathrm H}=\left[A^{\mathrm H}\cdot\left(X^{\mathrm H}+(\beta,\theta,\cdots,\theta)\right)\cdot A^{\mathrm H}\right]^{\mathrm H}$$

$$=\left[A^{\mathrm H}X^{\mathrm H}A^{\mathrm H}\right]^{\mathrm H}=AXA=A\Rightarrow Z\in A\{1\}$$

必要性．已知 $A^{(1)}$唯一，由(1)和(2)可得 $N(A)=\{\theta\}$, $N(A^{\mathrm H})=\{\theta\}$,

故　$n-r_A=0,\ m-r_{A^{\mathrm H}}=0\Rightarrow m=r_{A^{\mathrm H}}=r_A=n\Rightarrow A$ 可逆

$$AA^{(1)}A=A\Rightarrow A^{(1)}=A^{-1}$$

充分性．已知 $m=n$，A 可逆，则有

$$AA^{(1)}A=A\Rightarrow A^{(1)}=A^{-1}\text{唯一}$$

引理 2　$A_{m\times n}, B_{n\times p}, R(AB)\subset R(A), N(B)\subset N(AB)$.

证　$R(AB)=\left\{ABx\,\middle|\,\forall x\in \mathrm{C}^p\right\}=\left\{Ay\,\middle|\,y\overset{\Delta}{=}Bx\in \mathrm{C}^n\right\}$

$$\subset\left\{Ay\mid \forall y\in \mathrm{C}^n\right\}=R(A)$$

$\forall x\in N(B)\Rightarrow Bx=0\Rightarrow ABx=0\Rightarrow x\in N(AB)$

故 $N(B)\subset N(AB)$.

引理 3　$A_{m\times n}, B_{n\times p}, R(AB)=R(A)\Rightarrow\exists C_{p\times n}$, st. $A=ABC$.

证　划分 $A=(a_1,\cdots,a_n)$，则有

$a_i\in R(A)=R(AB)\Rightarrow\exists x_i\in \mathrm{C}^p$, st. $a_i=ABx_i$

$(a_1,\cdots,a_n)=AB(x_1,\cdots,x_n)\Rightarrow A=ABC$：　$C_{p\times n}\overset{\Delta}{=}(x_1,\cdots,x_n)$

Th5　$A_{m\times n}, B_{n\times p}, \lambda\in \mathrm{C}, \lambda^+=\begin{cases}1/\lambda & (\lambda\neq 0)\\ 0 & (\lambda=0)\end{cases}$.

(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) $R(AA^{(1)})=R(A)$： $R(A)=R(AA^{(1)}A)\overset{\text{引2}}{\subset}R(AA^{(1)})\overset{\text{引2}}{\subset}R(A)$

$N(A^{(1)}A)=N(A)$： $N(A)\subset N(A^{(1)}A)\subset N(AA^{(1)}A)=N(A)$

(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”

证 ① 必要性．(5) $\Rightarrow r_A=r_{A^{(1)}A}=n$

充分性．(5) $\Rightarrow r_{A^{(1)}A}=r_A=n\Rightarrow A^{(1)}A$ 可逆

$$(A^{(1)}A)^2=(A^{(1)}A)\Rightarrow A^{(1)}A=I_n$$

(8) ① $(AB)(AB)^{(1)}A=A\Leftrightarrow r_{AB}=r_A$

② $B(AB)^{(1)}(AB)=B\Leftrightarrow r_{AB}=r_B$

证 ① 必要性． $r_A\le r_{AB}\le r_A\Rightarrow r_{AB}=r_A$

充分性．已知 $r_{AB}=r_A$，则

引 2 $\Rightarrow R(AB)\subset R(A)$，故 $R(AB)=R(A)$

引 3 $\Rightarrow \exists C,\ \text{st.}\ A=ABC$

$$\Rightarrow(AB)(AB)^{(1)}A=(AB)(AB)^{(1)}\cdot ABC=(AB)C=A$$

② 充分性． $r_{B^{\mathrm{H}}A^{\mathrm{H}}}=r_{(AB)^{\mathrm{H}}}=r_{AB}=r_B=r_{B^{\mathrm{H}}}$

引 2 $\Rightarrow R(B^{\mathrm{H}}A^{\mathrm{H}})\subset R(B^{\mathrm{H}})\Rightarrow R(B^{\mathrm{H}}A^{\mathrm{H}})=R(B^{\mathrm{H}})$

引 3 $\Rightarrow B^{\mathrm{H}}=B^{\mathrm{H}}A^{\mathrm{H}}\cdot C\Rightarrow B=C^{\mathrm{H}}AB$

$$\Rightarrow B(AB)^{(1)}(AB) = C^{\mathrm{H}}AB\cdot(AB)^{(1)}(AB) = C^{\mathrm{H}}AB = B$$

Th6　$A_{m\times n}, Y\in A\{1\},\ Z\in A\{1\}\Rightarrow X\overset{\Delta}{=}YAZ\in A\{1,2\}$.

证　$AXA = A\cdot YAZ\cdot A = AY(AZA) = AYA = A$

$$XAX = YAZ\cdot A\cdot YAZ = Y(AZA)YAZ = Y\cdot AYA\cdot Z = YAZ = X$$

推论　$A_{m\times n}, Y\in A\{1\}\Rightarrow X\overset{\Delta}{=}YAY\in A\{1,2\}$.

Th7　设 $X\in A\{1\}$，则　$r_X = r_A \Leftrightarrow X\in A\{1,2\}$.

证　必要性．已知 $X\in A\{1\}$ 且 $r_X = r_A$，则

$$\left.\begin{array}{l}\text{Th 5(5)}\Rightarrow r_{XA} = r_A = r_X\\ \text{引理 2}\Rightarrow R(XA)\subset R(X)\end{array}\right\}\Rightarrow R(XA) = R(X)$$

引理 3$\Rightarrow \exists Y_{n\times m}$, st. $X = XAY$

$$\Rightarrow XAX = XA\cdot XAY = XAY = X$$

故 $X\in A\{1,2\}$.

充分性．已知 $X\in A\{1,2\}$，则

$$\left.\begin{array}{l}AXA = A\Rightarrow r_A\le r_X\\ XAX = X\Rightarrow r_X\le r_A\end{array}\right\}\Rightarrow r_X = r_A$$

[注]　$A\{1,2\} = \{\,X \mid X\in A\{1\}\text{ 且 } r(X) = r(A)\,\}$

引理 4　$r_{A^{\mathrm{H}}A} = r_A = r_{A^{\mathrm{H}}} = r_{AA^{\mathrm{H}}}$

Th8　$Y\overset{\Delta}{=}(A^{\mathrm{H}}A)^{(1)}A^{\mathrm{H}}\in A\{1,2,3\}$,　$Z\overset{\Delta}{=}A^{\mathrm{H}}(AA^{\mathrm{H}})^{(1)}\in A\{1,2,4\}$.

证　验证第一式（与教材证明方法不同）

$$\left.\begin{array}{l}\text{引 2：}R(A^{\mathrm{H}}A)\subset R(A^{\mathrm{H}})\\ \text{引 4：}r_{A^{\mathrm{H}}A} = r_{A^{\mathrm{H}}}\end{array}\right\}\Rightarrow R(A^{\mathrm{H}}A) = R(A^{\mathrm{H}})$$

引 3：$\exists B_{n\times m}$, st. $A^{\mathrm{H}} = A^{\mathrm{H}}AB$，即 $A = B^{\mathrm{H}}A^{\mathrm{H}}A$.

$$AYA = B^{\mathrm{H}}A^{\mathrm{H}}A\cdot(A^{\mathrm{H}}A)^{(1)}A^{\mathrm{H}}\cdot A = B^{\mathrm{H}}\cdot A^{\mathrm{H}}A = A$$

$$YAY = (A^{\mathrm{H}}A)^{(1)}A^{\mathrm{H}}\cdot A\cdot(A^{\mathrm{H}}A)^{(1)}A^{\mathrm{H}}$$

$$=\left(A^{\mathrm H}A\right)^{(1)}\cdot\left(A^{\mathrm H}A\right)\left(A^{\mathrm H}A\right)^{(1)}\cdot A^{\mathrm H}AB$$

$$=\left(A^{\mathrm H}A\right)^{(1)}\cdot A^{\mathrm H}A\cdot B=\left(A^{\mathrm H}A\right)^{(1)}A^{\mathrm H}=Y$$

$$AY=A\cdot\left(A^{\mathrm H}A\right)^{(1)}A^{\mathrm H}=B^{\mathrm H}A^{\mathrm H}A\cdot\left(A^{\mathrm H}A\right)^{(1)}\cdot A^{\mathrm H}AB=B^{\mathrm H}\left(A^{\mathrm H}A\right)B$$

由此可得 $(AY)^{\mathrm H}=AY$． 因此 $Y\in A\{1,2,3\}$．

Th9　$A^{+}=A^{(1,4)}AA^{(1,3)}$

证　$X=A^{(1,4)}AA^{(1,3)}$ ： Th6 $\Rightarrow$ X 满足 Penrose 方程(1)和(2)．

$$AX=AA^{(1,4)}AA^{(1,3)}=AA^{(1,3)}\Rightarrow(AX)^{\mathrm H}=\left(AA^{(1,3)}\right)^{\mathrm H}=AA^{(1,3)}=AX$$

$$XA=A^{(1,4)}AA^{(1,3)}A=A^{(1,4)}A\Rightarrow(XA)^{\mathrm H}=\left(A^{(1,4)}A\right)^{\mathrm H}=A^{(1,4)}A=XA$$

故 $A^{+}=X=A^{(1,4)}AA^{(1,3)}$．

Th10　(1) $r_{A^+}=r_A$ ；　　(2) $\left(A^{+}\right)^{+}=A$

(3) $\left(A^{\mathrm H}\right)^{+}=\left(A^{+}\right)^{\mathrm H}$　　对 $A^{\mathrm H}$ 取 $X=\left(A^{+}\right)^{\mathrm H}$

$\left(A^{\mathrm T}\right)^{+}=\left(A^{+}\right)^{\mathrm T}$　　对 $A^{\mathrm T}$ 取 $X=\left(A^{+}\right)^{\mathrm T}$

(4) $\left(A^{\mathrm H}A\right)^{+}=A^{+}\left(A^{\mathrm H}\right)^{+}$　　对 $\left(A^{\mathrm H}A\right)$ 取 $X=A^{+}\left(A^{\mathrm H}\right)^{+}$

$\left(AA^{\mathrm H}\right)^{+}=\left(A^{\mathrm H}\right)^{+}A^{+}$　　对 $\left(AA^{\mathrm H}\right)$ 取 $X=\left(A^{\mathrm H}\right)^{+}A^{+}$

(5) $A^{+}=\left(A^{\mathrm H}A\right)^{+}A^{\mathrm H}=A^{\mathrm H}\left(AA^{\mathrm H}\right)^{+}$

$$A^{+}=A^{+}AA^{+}=\begin{cases}A^{+}\left(AA^{+}\right)^{\mathrm H}=A^{+}\left(A^{+}\right)^{\mathrm H}A^{\mathrm H}\overset{4(1)}{=}\left(A^{\mathrm H}A\right)^{+}A^{\mathrm H}\\\left(A^{+}A\right)^{\mathrm H}A^{+}=A^{\mathrm H}\left(A^{+}\right)^{\mathrm H}A^{+}\overset{4(2)}{=}A^{\mathrm H}\left(AA^{\mathrm H}\right)^{+}\end{cases}$$

(6) $R\left(A^{+}\right)=R\left(A^{\mathrm H}\right)$,　$N\left(A^{+}\right)=N\left(A^{\mathrm H}\right)$.

$$\left.\begin{aligned}&(5)\Rightarrow R\left(A^{+}\right)=R\left(A^{\mathrm H}\left(AA^{\mathrm H}\right)^{+}\right)\subset R\left(A^{\mathrm H}\right)\\&(1)\Rightarrow r_{A^+}=r_A=r_{A^{\mathrm H}}\end{aligned}\right\}\Rightarrow R\left(A^{+}\right)=R\left(A^{\mathrm H}\right)$$

$$\left.\begin{aligned}&(5)\Rightarrow N\left(A^{+}\right)=N\left(\left(A^{\mathrm H}A\right)^{+}A^{\mathrm H}\right)\supset N\left(A^{\mathrm H}\right)\\&(1)\Rightarrow r_{A^+}=r_{A^{\mathrm H}}\Rightarrow m-r_{A^+}=m-r_{A^{\mathrm H}}\end{aligned}\right\}\Rightarrow N\left(A^{+}\right)=N\left(A^{\mathrm H}\right)$$

三、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.（P_L 表示子空间 L 上的正交投影矩阵）

Th11 M-逆与 P-逆等价.

证 (1) 设 X 是 A 的 M-逆：

$$AXA=P_{R(A)}\cdot(a_1,\cdots,a_n)=(a_1,\cdots,a_n)=A$$

$$XAX=P_{R(X)}\cdot(x_1,\cdots,x_m)=(x_1,\cdots,x_m)=X$$

$$(AX)^{\mathrm H}=P_{R(A)}^{\mathrm H}=P_{R(A)}=AX$$

$$(XA)^{\mathrm H}=P_{R(X)}^{\mathrm H}=P_{R(X)}=XA$$

故 X 是 A 的 P-逆.

(2) 设 X 是 A 的 P-逆：

$$\left.\begin{aligned}(AX)^2&=AXAX=AX\\(AX)^{\mathrm H}&=AX\end{aligned}\right\}\overset{\text{Th2}}{\Rightarrow} AX=P_{R(AX)}\overset{\text{Th5}}{=}P_{R(A)}$$

（$R(AA^{(1)})=R(A)$）

$$\left.\begin{aligned}(XA)^2&=XAXA=XA\\(XA)^{\mathrm H}&=XA\end{aligned}\right\}\Rightarrow XA=P_{R(XA)}=P_{R(X)}$$

（将 A 看作 $X^{(1)}$，$R(XX^{(1)})=R(X)$）

故 X 是 A 的 M-逆.

§6.4　应用

$A_{m\times n}, x\in \mathbf{C}^n, b\in \mathbf{C}^m$，线性方程组 $Ax=b$：

有解（相容）时，求极小范数解 x_0 满足 $\|x_0\|_2=\min\limits_{Ax=b}\|x\|_2$；

无解（不相容）时，求极小范数最小二乘解 x_0 满足 $\|x_0\|_2=\min\limits_{\|Ax-b\|_2=\min}\|x\|_2$.

一、矩阵的$\{1\}$-逆

Th26 $A_{m\times n}, B_{p\times q}, C_{m\times q}$.

(1) $AXB=C$ 有解 $\Leftrightarrow AA^{(1)}CB^{(1)}B=C$；

(2) $AXB=C$ 有解 $\Rightarrow$ 通解 $X=A^{(1)}CB^{(1)}+Y-A^{(1)}AYBB^{(1)}$　$(\forall Y_{n\times p})$

证 (1) 充分性．取 $X=A^{(1)}CB^{(1)}$ 即可．

必要性． $AXB=C$： $A=AA^{(1)}A,\ B=BB^{(1)}B$

$$AA^{(1)}AXBB^{(1)}B=C\Rightarrow AA^{(1)}CB^{(1)}B=C$$

(2) $A\cdot\left(A^{(1)}CB^{(1)}+Y-A^{(1)}AYBB^{(1)}\right)\cdot B=\cdots=C$

设 $AXB=C$ 的一个解为 X，则

$$X=X+A^{(1)}CB^{(1)}-A^{(1)}CB^{(1)}$$
$$=A^{(1)}CB^{(1)}+X-A^{(1)}AXBB^{(1)}\quad（取\ Y=X_{n\times p}）$$

[注] $AXB=C$ 的特解为 $A^{(1)}CB^{(1)}$

$AXB=O$ 的通解为 $Y-A^{(1)}AYBB^{(1)}$ $(\forall Y_{n\times p})$

例 1 $A_{m\times n}, AXA=A$ 的通解为 $X=A^{(1)}AA^{(1)}+Y-A^{(1)}AYAA^{(1)}$　$(\forall Y_{n\times m})$.

令 $Y-A^{(1)}\overset{\Delta}{=}Z$

则 $X=A^{(1)}AA^{(1)}+\left(A^{(1)}+Z\right)-\left(A^{(1)}AA^{(1)}AA^{(1)}+A^{(1)}AZAA^{(1)}\right)$

$$=A^{(1)}+Z-A^{(1)}AZAA^{(1)}\quad(\forall Z_{n\times m})$$

故 $A\{1\}=\left\{A^{(1)}+Z-A^{(1)}AZAA^{(1)}\ \middle|\ Z\in\mathbf{C}^{n\times m}\right\}$.

Th27 $A_{m\times n}, b\in \mathrm{C}^m$.　　（**Th4:** B 可逆时， $B^{(1)}=B^{-1}$ 唯一）

(1) $Ax=b$ 有解 $\Leftrightarrow$ $AA^{(1)}b=b$;

(2) $Ax=b$ 有解 $\Rightarrow$ 通解 $x=A^{(1)}b+\left(I-A^{(1)}A\right)y\quad(\forall y\in \mathrm{C}^n)$.

[注] $Ax=b$ 的特解为 $A^{(1)}b$， $Ax=0$ 的通解为 $(I-A^{(1)}A)y\ \ (\forall y\in \mathrm{C}^n)$.

例 2 $A=\begin{bmatrix}1&1&0\\0&0&1\\0&0&0\end{bmatrix}, b=\begin{bmatrix}l\\k\\k\end{bmatrix}$.　（**Hermite 标准形中取** $c_1=1, c_2=3$ ）

(1) 求 $A^{(1)}$;

(2) l 与 k 取何值时，$Ax=b$ 有解，并求通解.

解 置换矩阵 $P=\begin{bmatrix}1&0&0\\0&0&1\\0&1&0\end{bmatrix}, A^{(1)}=P\cdot\begin{bmatrix}1&0&0\\0&1&0\\0&0&0\end{bmatrix}=\begin{bmatrix}1&0&0\\0&0&0\\0&1&0\end{bmatrix}$

$$AA^{(1)}b=b\Rightarrow\begin{bmatrix}l\\k\\0\end{bmatrix}=\begin{bmatrix}l\\k\\k\end{bmatrix}\Rightarrow k=0, l\text{ 任意}$$

$$\text{通解}\quad x=\begin{bmatrix}l\\0\\0\end{bmatrix}+\begin{bmatrix}0&-1&0\\0&1&0\\0&0&0\end{bmatrix}\begin{bmatrix}\eta_1\\\eta_2\\\eta_3\end{bmatrix}=\begin{bmatrix}l\\0\\0\end{bmatrix}+\eta_2\begin{bmatrix}-1\\1\\0\end{bmatrix}\quad(\forall \eta_2\in \mathrm{C})$$

Th28 给定 $A_{m\times n}, X_{n\times m}$. 若对任意的 $b\in R(A), x=Xb$ 都是 $Ax=b$ 的解，则 $X\in A\{1\}$.

证 划分 $A=(a_1,\cdots,a_n)$ ，则有

$$a_i\in R(A)\Rightarrow A(Xa_i)=a_i\ \ (i=1,2,\cdots,n)$$

$$\Rightarrow AX(a_1,\cdots,a_n)=(a_1,\cdots,a_n)$$

$$\Rightarrow AXA=A\Rightarrow X\in A\{1\}$$

二、矩阵的{1,4}-逆

引理 7 $A_{m\times n}, S_4=\left\{X\,\middle|\,XA=A^{(1,4)}A, X\in \mathrm{C}^{n\times m}\right\}\Rightarrow A\{1,4\}=S_4$.

证 (1) 先证 $S_4 \subset A\{1,4\}$：$\forall X \in S_4$

$$\left.\begin{aligned} &AXA = A \cdot A^{(1,4)}A = A \\ &(XA)^{\mathrm{H}} = \left(A^{(1,4)}A\right)^{\mathrm{H}} = A^{(1,4)}A = XA \end{aligned}\right\} \Rightarrow X \in A\{1,4\}$$

(2) 再证 $A\{1,4\} \subset S_4$：$\forall X \in A\{1,4\}$

$$A^{(1,4)}A = A^{(1,4)} \cdot AXA = \left(A^{(1,4)}A\right)^{\mathrm{H}}(XA)^{\mathrm{H}}$$

$$= \left(XAA^{(1,4)}A\right)^{\mathrm{H}} = (XA)^{\mathrm{H}} = XA\,.\quad \text{故 } X \in S_4\,.$$

Th29 $A_{m\times n}$，给定 $A^{(1,4)}$，则 $A\{1,4\} = \left\{A^{(1,4)} + Z\left(I - AA^{(1,4)}\right) \middle| \, Z \in \mathrm{C}^{n\times m}\right\}$.

证 显然，$XA = A^{(1,4)}A$ 有解 $X = A^{(1,4)}$.

通解 $X = \left(A^{(1,4)}A\right)A^{(1,4)} + Y - YAA^{(1,4)}$ $(\forall Y_{n\times m})$

令 $Y - A^{(1,4)} = Z$

则 $X = A^{(1,4)}AA^{(1,4)} + \left(A^{(1,4)} + Z\right) - \left(A^{(1,4)}AA^{(1,4)} + ZAA^{(1,4)}\right)$

$= A^{(1,4)} + Z\left(I - AA^{(1,4)}\right)$ $(\forall Z_{n\times m})$

由引理 7 即得所证.

引理 6 $Ax = b$ 有解 $\Rightarrow$ 极小范数解 x_0 唯一，且 $x_0 \in R\left(A^{\mathrm{H}}\right)$.

证 (1) 先证 $x_0 \in R\left(A^{\mathrm{H}}\right)$：反证法．若 $x_0 \notin R\left(A^{\mathrm{H}}\right)$，由 Th1-35 知

$$\mathrm{C}^n = R\left(A^{\mathrm{H}}\right) \oplus N(A), \quad N(A) = R^{\perp}\left(A^{\mathrm{H}}\right)$$

$x_0 = y_0 + z_0$，$y_0 \in R\left(A^{\mathrm{H}}\right)$，$z_0 \in N(A) = R^{\perp}\left(A^{\mathrm{H}}\right)$ 且 $z_0 \neq 0$

$\|x_0\|_2^2 = \|y_0\|_2^2 + \|z_0\|_2^2 > \|y_0\|_2^2$

$Ax_0 = b,\ Az_0 = 0 \Rightarrow Ay_0 = Ax_0 - Az_0 = Ax_0 = b$

故 x_0 不是 $Ax = b$ 的极小范数解，矛盾！

(2) 再证唯一性：设 $y_0 \in R\left(A^{\mathrm{H}}\right)$，且 $Ay_0 = b$，则

$x_0 - y_0 \in R\left(A^{\mathrm{H}}\right) = N^{\perp}(A)$

$A(x_0 - y_0) = b - b = 0 \Rightarrow x_0 - y_0 \in N(A)$

故 $x_0 - y_0 = 0 \Rightarrow x_0 = y_0$，即 $R(A^{\mathrm{H}})$中只有 $Ax = b$ 的一个解.

Th30 (1) $Ax = b$ 有解 $\Rightarrow x_0 = A^{(1,4)}b$ 是唯一极小范数解；

(2) 给定 $X_{n\times m}$，若对任意的 $b \in R(A)$, $x = Xb$ 都是 $Ax = b$ 的极小范数解，则 $X \in A\{1,4\}$.

证(1) Th27：$Ax = b$ 有解 $\Rightarrow AA^{(1,4)}b = b$，故 $x_0 = A^{(1,4)}b$ 是 $Ax = b$ 的解.

下证 $x_0 \in R(A^{\mathrm{H}})$：

$Ax = b$ 有解 $\Rightarrow b \in R(A) \Rightarrow \exists u \in \mathbf{C}^n, \mathrm{st.}\, b = Au$

$x_0 = A^{(1,4)}b = A^{(1,4)}Au = \left(A^{(1,4)}A\right)^{\mathrm{H}} u = A^{\mathrm{H}}\left[\left(A^{(1,4)}\right)^{\mathrm{H}} u\right] \in R(A^{\mathrm{H}})$

由引理 6 的唯一性证明知，x_0 是 $Ax = b$ 的唯一极小范数解.

(2) 划分 $A = (a_1, \cdots, a_n)$，则有

$a_i \in R(A) \Rightarrow Xa_i$ 是 $Ax = a_i$ 的极小范数解（已知）

$$\overset{(1)}{\Rightarrow} Xa_i = A^{(1,4)}a_i \quad (i = 1, \cdots, n)$$

$$\Rightarrow X(a_1, \cdots, a_n) = A^{(1,4)}(a_1, \cdots, a_n)$$

$$\Rightarrow XA = A^{(1,4)}A \Rightarrow X \in S_4 = A\{1,4\} \quad \text{（引理 7）}$$

三、矩阵的$\{1,3\}$-逆

引理 8 $A_{m\times n}, S_3 \overset{\Delta}{=} \left\{X \mid AX = AA^{(1,3)}, X \in \mathbf{C}^{n\times m}\right\} \Rightarrow A\{1,3\} = S_3$.

证　类似于引理 7 的证明.

Th31 $A_{m\times n}$，给定 $A^{(1,3)}$, 则 $A\{1,3\} = \left\{A^{(1,3)} + \left(I - A^{(1,3)}A\right)Z \mid Z \in \mathbf{C}^{n\times m}\right\}$.

证　利用 Th26, 类似于 Th29 的证明.

Th32 $Ax = b$ 无解 $\Rightarrow \tilde{x}_0 = A^{(1,3)}b$ 是最小二乘解，即

$$\|A\tilde{x}_0 - b\|_2 = \min\|Ax - b\|_2$$

证　$P \overset{\Delta}{=} AA^{(1,3)} \Rightarrow Pb = A \cdot A^{(1,3)}b \in R(A)$

$$b = Pb + (I-P)b \Rightarrow Ax - b = (Ax - Pb) + [-(I-P)b]$$

因为 $(Ax)^{\mathrm{H}} \cdot (I-P)b = x^{\mathrm{H}} A^{\mathrm{H}} \left(I - AA^{(1,3)}\right) b$

$$= x^{\mathrm{H}}\left(A^{\mathrm{H}} - A^{\mathrm{H}}\left(AA^{(1,3)}\right)^{\mathrm{H}}\right)b$$

$$= x^{\mathrm{H}}\left(A^{\mathrm{H}} - \left(AA^{(1,3)}A\right)^{\mathrm{H}}\right)b = 0$$

$$(Pb)^{\mathrm{H}} \cdot (I-P)b = b^{\mathrm{H}} \cdot AA^{(1,3)}\left(I - AA^{(1,3)}\right)b = 0$$

所以 $(Ax - Pb) \perp (I-P)b \Rightarrow \|Ax-b\|_2^2 = \|Ax - Pb\|_2^2 + \|(I-P)b\|_2^2$

故 $\|Ax-b\|_2 = \min \Leftrightarrow \|Ax-Pb\|_2 = \min \Leftrightarrow Ax = Pb \quad (\because Pb \in R(A))$

又 $A\tilde{x}_0 = AA^{(1,3)}b = Pb$，即 $\tilde{x}_0$ 是 $Ax = Pb$ 的解.

所以 $\tilde{x}_0$ 是 $\|Ax-b\|_2 = \min$ 的解.

例 3 (1) $Ax = b$ 有解： $Ax_0 = b \Leftrightarrow A^{\mathrm{H}}Ax_0 = A^{\mathrm{H}}b$；

(2) $Ax = b$ 无解： $\|Ax_0 - b\|_2 = \min\|Ax-b\|_2 \Leftrightarrow A^{\mathrm{H}}Ax_0 = A^{\mathrm{H}}b$.

证 (1) $Ax_0 = b \Rightarrow A^{\mathrm{H}}Ax_0 = A^{\mathrm{H}}b$；

$$A^{\mathrm{H}}Ax_0 = A^{\mathrm{H}}b \Rightarrow A^{\mathrm{H}}(Ax_0 - b) = 0 \Rightarrow Ax_0 - b \in N\left(A^{\mathrm{H}}\right) = R^{\perp}(A)$$

$Ax = b$ 有解 $\Rightarrow b \in R(A) \Rightarrow Ax_0 - b \in R(A)$

故 $Ax_0 - b = 0 \Rightarrow Ax_0 = b$.

(2) Th32： x_0 是最小二乘解 $\Rightarrow Ax_0 = Pb = AA^{(1,3)}b$

$$\Rightarrow A^{\mathrm{H}}(Ax_0 - b) = A^{\mathrm{H}}\left(AA^{(1,3)}b - b\right) = A^{\mathrm{H}}\left(AA^{(1,3)}\right)^{\mathrm{H}}b - A^{\mathrm{H}}b$$

$$= \left(AA^{(1,3)}A\right)^{\mathrm{H}}b - A^{\mathrm{H}}b = 0$$

$$\Rightarrow A^{\mathrm{H}}Ax_0 = A^{\mathrm{H}}b$$

若 $A^{\mathrm{H}}Ax_0 = A^{\mathrm{H}}b$，则有

$$A^{\mathrm{H}}\left(Ax_0 - AA^{(1,3)}b\right) = A^{\mathrm{H}}Ax_0 - A^{\mathrm{H}}\left(AA^{(1,3)}\right)^{\mathrm{H}}b$$

$$= A^{\mathrm{H}}b - \left(AA^{(1,3)}A\right)^{\mathrm{H}}b = 0$$

于是　$Ax_0 - AA^{(1,3)}b \in N(A^{\mathrm{H}}) = R^{\perp}(A)$

又　$Ax_0 - AA^{(1,3)}b = A\left(x_0 - A^{(1,3)}b\right) \in R(A)$

所以　$Ax_0 - AA^{(1,3)}b = 0 \Rightarrow Ax_0 = AA^{(1,3)}b = Pb$

$$\overset{\text{Th32}}{\Rightarrow} \|Ax_0 - b\|_2 = \min\|Ax - b\|_2$$

Th33　$Ax = b$ 无解 $\Rightarrow x_0 = A^{+}b$ 是极小范数最小二乘解，且唯一.

证　**Th32：** $\|Ax - b\|_2 = \min \Leftrightarrow Ax = AA^{(1,3)}b$　　（有解 $x = A^{(1,3)}b$）

Th30： $Ax = AA^{(1,3)}b$ 有解 $\Rightarrow$

极小范数解 $x_0 = A^{(1,4)} \cdot AA^{(1,3)}b \overset{\text{Th9}}{=} A^{+}b$ 唯一

结论：(1) $Ax = b$ 有解 $\overset{\text{Th30}}{\Rightarrow} x_0 = A^{+}b$ 是极小范数解，且唯一.

(2) $Ax = b$ 无解 $\overset{\text{Th33}}{\Rightarrow} x_0 = A^{+}b$ 是极小范数最小二乘解，且唯一.

(3) $Ax = b$ 有解 $\overset{\text{Th27}}{\Leftrightarrow} AA^{+}b = b$.

(4) $Ax = b$ 有解 $\overset{\text{Th27}}{\Rightarrow}$ 通解　$x = A^{+}b + \left(I - A^{+}A\right)y$　$(\forall y \in \mathbf{C}^n)$.

例 4　已知 $A=\begin{bmatrix}1&1&0&2&2\\2&2&3&1&1\\1&1&1&1&1\end{bmatrix}$，$b=\begin{bmatrix}1\\2\\1\end{bmatrix}$.

1. 求 A 的满秩分解；　2. 求 A^+；

3. 用广义逆矩阵方法判断线性方程组 $Ax=b$ 是否有解；

4. 求线性方程组 $Ax=b$ 的极小范数解或极小范数最小二乘解 x_0

（要求指出所求的是那种解）.

解 1. $A\xrightarrow{\text{行}}\begin{bmatrix}1&1&0&2&2\\0&0&1&-1&-1\\0&0&0&0&0\end{bmatrix}$：$c_1=1, c_2=3$

$$F=\begin{bmatrix}1&0\\2&3\\1&1\end{bmatrix},\quad G=\begin{bmatrix}1&1&0&2&2\\0&0&1&-1&-1\end{bmatrix}：\quad A=FG$$

2. $F^+=\left(F^{\mathrm{T}}F\right)^{-1}F^{\mathrm{T}}=\begin{bmatrix}6&7\\7&10\end{bmatrix}^{-1}F^{\mathrm{T}}$

$$=\frac{1}{11}\begin{bmatrix}10&-7\\-7&6\end{bmatrix}F^{\mathrm{T}}=\frac{1}{11}\begin{bmatrix}10&-1&3\\-7&4&-1\end{bmatrix}$$

$$G^+=G^{\mathrm{T}}\left(GG^{\mathrm{T}}\right)^{-1}=G^{\mathrm{T}}\begin{bmatrix}10&-4\\-4&3\end{bmatrix}^{-1}=G^{\mathrm{T}}\cdot\frac{1}{14}\begin{bmatrix}3&4\\4&10\end{bmatrix}=\frac{1}{14}\begin{bmatrix}3&4\\3&4\\4&10\\2&-2\\2&-2\end{bmatrix}$$

$$A^+=G^+F^+=\frac{1}{154}\begin{bmatrix}2&13&5\\2&13&5\\-30&36&2\\34&-10&8\\34&-10&8\end{bmatrix}$$

3/4. $x_0=A^+b=\frac{1}{14}\begin{bmatrix}3&3&4&2&2\end{bmatrix}^{\mathrm{T}}$

$$AA^{+}b = Ax_0 = \frac{1}{14}\begin{bmatrix}14\\28\\14\end{bmatrix} = b \quad \Rightarrow Ax = b \text{有解}$$

故 x_0 是 $Ax=b$ 的极小范数解.

例 5 设非零矩阵 $A_{m\times n}$ 的 M-P 逆为 A^{+}，$B_{n\times n}$ 为酉矩阵，则 $[A \vdots AB]^{+} =$ ___ .

分析：设 A 的满秩分解为 $A=FG$，则 $A^{+}=G^{+}F^{+}$.

满秩分解 $[A \vdots AB] = F[G \vdots GB] \stackrel{\Delta}{=} F\tilde{G}$ （$\tilde{G}$ 为行满秩矩阵）

$$\tilde{G}^{+} = \begin{bmatrix}G^{\mathrm{H}}\\B^{\mathrm{H}}G^{\mathrm{H}}\end{bmatrix}\left([G \vdots GB]\begin{bmatrix}G^{\mathrm{H}}\\B^{\mathrm{H}}G^{\mathrm{H}}\end{bmatrix}\right)^{-1} = \begin{bmatrix}G^{\mathrm{H}}\\B^{\mathrm{H}}G^{\mathrm{H}}\end{bmatrix}\left(GG^{\mathrm{H}} + GBB^{\mathrm{H}}G^{\mathrm{H}}\right)^{-1}$$

$$= \begin{bmatrix}G^{\mathrm{H}}\\B^{\mathrm{H}}G^{\mathrm{H}}\end{bmatrix}\left(2GG^{\mathrm{H}}\right)^{-1} = \begin{bmatrix}G^{\mathrm{H}}\\B^{\mathrm{H}}G^{\mathrm{H}}\end{bmatrix}\cdot\frac{1}{2}\left(GG^{\mathrm{H}}\right)^{-1} = \frac{1}{2}\begin{bmatrix}G^{+}\\B^{\mathrm{H}}G^{+}\end{bmatrix}$$

$$[A \vdots AB]^{+} = \tilde{G}^{+}F^{+} = \frac{1}{2}\begin{bmatrix}G^{+}\\B^{\mathrm{H}}G^{+}\end{bmatrix}F^{+} = \frac{1}{2}\begin{bmatrix}G^{+}F^{+}\\B^{-1}G^{+}F^{+}\end{bmatrix} = \frac{1}{2}\begin{bmatrix}A^{+}\\B^{+}A^{+}\end{bmatrix}$$