北京航空航天大学

矩阵理论 A 笔记

任课教师：赵迪
编辑：　张京蕊

写在前边

编者按：矩阵理论 A 课程是我校一门研究生公共课程，本人特将 2008 年秋季本课程赵迪老师大班的笔记整理成电子版，以供后人学习、参考之用。本笔记包括七大部分，编号从零至六。

众所周知，赵老师上课从不用课件，完全是板书，所以选这门课程的同学每堂课必然要仔仔细细的记笔记，虽然我把赵老师这门课程的笔记整理成了电子版，但仍不鼓励大家拿着打印稿，不记笔记，甚至不去上课。俗话说："好记性不如烂笔头。"勤奋一些，平时认认真真把笔记记清，可以巩固对这门课程知识的记忆，为以后考试和应用打好基础，事半功倍。

同时严正声明：禁止将此笔记用于任何商业用途。虽然这个电子版是我搞出来的，但我仍认为这套笔记的版权应该归赵老师或者北航理学院所有，希望同学不要因贪小利而忘大义。

最后，希望这份电子版的笔记能够给同学们学习这门课程带来方便，祝同学们在北航生活、学习、工作愉快！

目 录

§0 补充公式

令 $A=(a_{ij})_{n\times n}\in C^{n\times n}$，$f(x)=a_0+a_1x+\ \cdots\ +a_mx^m$

定义 $f(A)=a_0I+a_1A+\ \cdots\ +a_mA^m$，其中 $I=I_n=\begin{pmatrix}1 & \cdots & 0\\ \vdots & \ddots & \vdots\\ 0 & \cdots & 1\end{pmatrix}$

若 $g(x)=b_0+b_1x+\ \cdots\ +b_kx^k$，$f(x)\bullet g(x)=g(x)\bullet f(x)$，则 $f(A)\bullet g(A)=g(A)\bullet f(A)$

分块公式

令 $A=\begin{pmatrix}A_1 & 0\\ 0 & A_2\end{pmatrix}$，$A_1$，$A_2$ 为方阵

则：（1）$A^k=\begin{pmatrix}A_1^k & 0\\ 0 & A_2^k\end{pmatrix}$

（2）$f(A)=\begin{pmatrix}f(A_1) & 0\\ 0 & f(A_2)\end{pmatrix}$，$f(x)$为多项式

令 $A=\begin{pmatrix}A_1 & & & (*)\\ & A_2 & & \\ & & \ddots & \\ O & & & A_s\end{pmatrix}$，$A_1,\ldots,A_s$ 为方阵

则：（1）$A^k=\begin{pmatrix}A_1^k & & & (*)\\ & A_2^k & & \\ & & \ddots & \\ O & & & A_s^k\end{pmatrix}$

（2）$f(A)=\begin{pmatrix}f(A_1) & & & (*)\\ & f(A_2) & & \\ & & \ddots & \\ O & & & f(A_s)\end{pmatrix}$

相似关系：$A\backsim B$，（$P^{-1}AP=B$）

则：（1）$(P^{-1}AP)^k=P^{-1}A^kP$，（k=0,1,2,…）

（2）$f(P^{-1}AP)=P^{-1}f(A)P$，$f(x)$为多项式

许尔公式（schur）：每个复方阵，$A=(a_{ij})_{n\times n}$ 都相似于上三角形。

即：$P^{-1}AP=\begin{pmatrix}\lambda_1 & & & (*)\\ & \lambda_2 & & \\ & & \ddots & \\ O & & & \lambda_n\end{pmatrix}$，其中 $\lambda_1,\ldots,\lambda_n$ 的次序可以任意指定

Pf: 用归纳法

$n = 1$ 时成立

可以设为$(n = 1)$阶方阵成立

对于 n 阶方阵 $A = (a_{ij})_{n\times n}$ 设特征值为 $\lambda_1,\ldots,\lambda_n$

取 λ_1 对应的特征向量，记为 $\alpha_1\neq 0$，$A\alpha_1=\lambda_1\alpha_1$

把 α_1 扩展为可逆方阵 $Q = (\alpha_1,\alpha_2,\ldots,\alpha_n)$

$\therefore Q^TQ = I_n = (e_1,e_2,\ldots,e_n)$

又$\because Q^{-1}(\alpha_1,\alpha_2,\ldots,\alpha_n) = (Q^{-1}\alpha_1, Q^{-1}\alpha_2,\ldots, Q^{-1}\alpha_n)$

其中 $Q^{-1}\alpha_1=\begin{pmatrix}1\\0\\ \vdots\\0\end{pmatrix}=e_1$，$Q^{-1}\alpha_2=\begin{pmatrix}0\\1\\ \vdots\\0\end{pmatrix}=e_2$，…，$Q^{-1}\alpha_n=\begin{pmatrix}0\\0\\ \vdots\\1\end{pmatrix}=e_n$

$$\begin{aligned}Q^{-1}AQ&=Q^{-1}A(\alpha_1,\alpha_2,\cdots,\alpha_n)\\&=Q^{-1}(A\alpha_1,A\alpha_2,\cdots,A\alpha_n)\\&=Q^{-1}(\lambda\alpha_1,\cdots,*,*,*)\\&=(\lambda_1Q^{-1}\alpha_1,(*),\cdots,(*))\end{aligned}$$

$\therefore$ $=\left(\begin{pmatrix}\lambda_1\\0\\ \vdots\\0\end{pmatrix}\begin{pmatrix}\vdots\\ \vdots\\ \vdots\\ \vdots\end{pmatrix}\cdots\begin{pmatrix}\vdots\\ \vdots\\ \vdots\\ \vdots\end{pmatrix}\right)$，其中 A_1 为$(n\text{-}1)$阶

$\overset{\text{记}}{\underset{\text{为}}{=}}\begin{pmatrix}\lambda_1 & (*)\\ 0 & \mathrm{A}_1\end{pmatrix}$

$\therefore$由假设，对于 A_1 必有$(n\text{-}1)$阶 P_1，可推出 $P^{-1}AP=\begin{pmatrix}\lambda_2 & & *\\ & \ddots & \\ \mathrm{O} & & \lambda_\mathrm{n}\end{pmatrix}$

$\therefore$得证。

Eg.知 n 阶方阵 A，适合 $A^k = 0$，则$|A+I|=1$

Pf: $A^k=0\Rightarrow$任意特征值$\lambda^k=0\Rightarrow\lambda=0$

即全体特征值为 0,0,…,0

由需要$P^{-1}AP=\begin{pmatrix}0 & & *\\ & \ddots & \\ O & & 0\end{pmatrix}\Rightarrow\left|P^{-1}AP+I\right|=1$

$\because\left|P^{-1}AP+P^{-1}IP\right|=\left|P^{-1}(A+I)P\right|=\left|A+I\right|\Rightarrow\left|A+I\right|=1$

注:（1）若$A\backsim B$（相似），则A、B有相同特征值$\lambda_1,\ldots,\lambda_n$

可引入记号：谱集$\sigma(A)=\{\lambda_1,\lambda_2,\cdots,\lambda_n\}$（全体特征值，含重复）

$\therefore A\backsim B\Rightarrow\sigma(A)=\sigma(B)$

（2）$A\backsim B\Rightarrow\left|\lambda I-A\right|=\left|\lambda I-B\right|=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n)$，特征多项式

$\because P^{-1}AP=B\Rightarrow\left|\lambda I-A\right|=\left|P^{-1}(\lambda I-A)P\right|=\left|\lambda I-B\right|$

引理: 若$A=\begin{pmatrix}A_1 & 0\\ 0 & A_2\end{pmatrix}$，则$|\lambda I-A|=|\lambda I_1-A|=|\lambda I_1-A_1||\lambda I_2-A_2|$

$\Rightarrow\sigma(A)=\sigma(A_1)\cup\sigma(A_2)$

即$\{\lambda_1,\lambda_2,\ldots,\lambda_n\}=\{\lambda_1,\lambda_2,\ldots,\lambda_k\}\cup\{\lambda_{k+1},\lambda_{k+2},\ldots,\lambda_n\}$

设$B=\begin{pmatrix}\lambda_1 & & & *\\ & \lambda_2 & & \\ & & \ddots & \\ O & & & \lambda_n\end{pmatrix}$，$f(x)$为多项式，则$f(B)=\begin{pmatrix}f(\lambda_1) & & (*)\\ & \ddots & \\ O & & f(\lambda_n)\end{pmatrix}$

引理: 若n阶方阵A的谱集$\sigma(A)=\{\lambda_1,\lambda_2,\ldots,\lambda_n\}$，

则$f(A)$的全体特征值为$\{f(\lambda_1),f(\lambda_2),\ldots,f(\lambda_n)\}$，$f(x)$为多项式

Pf: 由许尔定理，$A\backsim B=\begin{pmatrix}\lambda_1 & & *\\ & \ddots & \\ O & & \lambda_n\end{pmatrix}\Rightarrow f(A)\backsim f(B)=\begin{pmatrix}f(\lambda_1) & & *\\ & \ddots & \\ O & & f(\lambda_n)\end{pmatrix}$

$\Rightarrow f(x)$的全体特征值为$\{f(\lambda_1),f(\lambda_2),\ldots,f(\lambda_n)\}$，$f(x)$为多项式

例如：λ为A的特征值$\Rightarrow\lambda^k$为A^k的特征值。（$f(x)=x^k$）

引理：令 $B=\begin{pmatrix}\lambda_1 & & * \\ & \ddots & \\ O & & \lambda_n\end{pmatrix}$，$f(x) = |xI - B| = (x-\lambda_1)(x-\lambda_2)\dots(x-\lambda_n)$

则 $f(B) = (B - \lambda_1 I)(B - \lambda_2 I)\dots(B - \lambda_n I) = 0$

Pf：当 $n = 2$ 时，$B=\begin{pmatrix}\lambda_1 & * \\ 0 & \lambda_2\end{pmatrix}$，$f(x) = (x-\lambda_1)(x-\lambda_2)$

$$\Rightarrow f(B)=(B-\lambda_1 I)(B-\lambda_2 I)=\begin{pmatrix}0 & * \\ 0 & (\lambda_2-\lambda_1)\end{pmatrix}\begin{pmatrix}(\lambda_1-\lambda_2) & * \\ 0 & 0\end{pmatrix}=\begin{pmatrix}0 & 0 \\ 0 & 0\end{pmatrix}$$

∴得证

★Cayley 公式：设 n 阶方阵 A 的特征多项式为 $f(x) = |xI - A| = a_0 + a_1x + \dots + x^n$

则 $f(A) = a_0I + a_1A + \dots + A^n = 0$

Pf：由许尔 $P^{-1}AP=B=\begin{pmatrix}\lambda_1 & & & (*) \\ & \lambda_2 & & \\ & & \ddots & \\ O & & & \lambda_n\end{pmatrix}$

$$\Rightarrow P^{-1}f(A)P=f(P^{-1}AP)=f(B)=0$$（引理）

定义：若多项式 $f(x)$ 使 $f(A) = 0$，则称 $f(x)$ 为 A 的一个零化式

结论：方阵 A 的特征多项式 $f(x) = |xI - A|$ 为 A 的一个零化式

Eg：$A=\begin{pmatrix}0 & -1 \\ 1 & 0\end{pmatrix}$，特征多项式 $f(x) = x^2 + 1$

可知：$f(A)=A^2+I=\begin{pmatrix}-1 & 0 \\ 0 & -1\end{pmatrix}+I=0$

且 $f(x) = |xI - A| = (x - i)(x + i)$，（$i=\sqrt{-1}, i^2=-1$）

$f(A) = (A - iI)(A + iI) = 0$

也可取 $P=\begin{pmatrix}i & -1 \\ 1 & i\end{pmatrix}$，则 $P^{-1}=\frac{1}{2}\begin{pmatrix}-i & -1 \\ 1 & -i\end{pmatrix}$

$P^{-1}AP=\begin{pmatrix}i & 0 \\ 0 & -1\end{pmatrix}$，对角形

Eg：知 $A=\begin{pmatrix}0 & & (*)\\ & \ddots & \\ O & & 0\end{pmatrix}_{n\times n}$，则 $A^n=0$

由 Cayley 特征多项式：$f(x)=x^n \Rightarrow f(A)=A^n=0$

Ex.1. $A=\begin{pmatrix}1 & -1\\ 1 & 1\end{pmatrix}$，求 P 使得 $P^{-1}AP$ 为对角阵，并验证 Cayley 定理。

2. $A=\begin{pmatrix}a & b\\ c & d\end{pmatrix}$，求 $f(x)=|xI-A|$验证 $f(A)=0$

补充知识（schur 公式、Cayley 公式）应用

由 $A^n=-(a_0I+a_1A+\ldots+a_{n-1}A^{n-1})$ ①

$\Rightarrow A^{n+1}=A\bullet A^n=-\left(a_0A+a_1A^2+\cdots+a_{n-1}A^n\right)$ ②

把①代入②$\Rightarrow A^{n+1}=(*)I+(*)A+\cdots+(*)A^{n-1}$

可知： 任何 A^m $(m\geq n)$都可写成 $I,A,\ldots,A^{n-1}$ 的线性组合。

任何多项式 $g(A)$，可写成 $I,A,\ldots,A^{n-1}$ 的组合。

Eg：若$|A|\neq 0$，$f(x)=|xI-A|=a_0+a_1x+\ldots+x^n$，$a_0=|-A|\neq 0$

则 A^{-1} 可用 A 的多项式表示

$\because a_1A+a_2A^2+\ldots+a_{n-1}A^{n-1}+A^n=-a_0I$

$A(a_1I+a_2A+\ldots+a_{n-1}A^{n-2}+A^{n-1})=-a_0I$

$\Rightarrow A^{-1}=-\dfrac{1}{a_0}\left(a_1I+\cdots+a_{n-1}A^{n-2}+A^{n-1}\right)$

零化式定义： 若 $g(x)=b_0+b_1x+\ldots+b_mx^m$，使得 $g(A)=b_0I+b_1A+\ldots+b_mA^m=0$，称 $g(x)$为方阵 A 的零化式

注： 方阵 A 的零化式有无穷多个

$\because$取特征多项式 $f(x)$则 $f(A)=0$

任取式 $h(x)$，$f(A)h(A)=0\Rightarrow f(x)h(x)$也是零化式

极小式定义： 在方阵 A 的零化式集合中，去次数最小的且首项系数为 1 的零化式 $m_A(x)$，称它为 A 的极小式

注： 极小式唯一

性质： ①极小式 $m(x)$必为特征多项式 $f(x) = |xI - A|$的因式。

②特征多项式 $f(x) = |xI - A|$的每个单因子$(x - \lambda)$也是极小式的因子。

③若 $f(x)=|xI-A|=(x-\lambda_1)^{n_1}(x-\lambda_2)^{n_2}\cdots(x-\lambda_s)^{n_s}$，

则极小式 $m(x)=(x-\lambda_1)^{l_1}(x-\lambda_2)^{l_2}\cdots(x-\lambda_s)^{l_s}$，

且 $1 \le l_1 \le n_1, 1 \le l_2 \le n_2, \ldots, 1 \le l_s \le n_s$，$\lambda_1, \lambda_2, \ldots, \lambda_n$ 互不相同。

Eg. $A=\begin{pmatrix}2&1&0\\0&2&0\\0&0&1\end{pmatrix}$，$B=\begin{pmatrix}2&0&0\\0&2&0\\0&0&1\end{pmatrix}$，求极小式 $m_A(x)$，$m_B(x)$

解：（1）$|xI - A| = (x-2)^2(x-1)$

极小式为：$(x-2)^2(x-1)$或$(x-2)(x-1)$

计算：$(A-2I)(A-I)=\begin{pmatrix}0&1&0\\0&0&0\\0&0&-1\end{pmatrix}\begin{pmatrix}1&1&0\\0&1&0\\0&0&0\end{pmatrix}\neq 0$

$\therefore$极小式为 $m_A(x) = (x-2)^2(x-1)$

（2）$|xI - B| = (x-2)^2(x-1)$

计算：$(B-2I)(B-I)=\begin{pmatrix}0&0&0\\0&0&0\\0&0&-1\end{pmatrix}\begin{pmatrix}1&0&0\\0&1&0\\0&0&0\end{pmatrix}=0$

$\therefore$极小式为 $m_B(x) = (x-2)(x-1)$

Eg.求下列极小式 $m(x)$

（1）$A=\begin{pmatrix}4&6&0\\-3&-5&0\\-3&-6&1\end{pmatrix}$，（2）$B=\begin{pmatrix}4&-6&0\\2&-3&0\\-2&3&2\end{pmatrix}$，

（3）$C=\begin{pmatrix}2&1&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix}$，（4）$D=\begin{pmatrix}2&1&0\\-1&0&0\\0&0&2\end{pmatrix}$

解：（1）特征多项式$|xI - A| = (x-1)^2(x+2)$

极小式为：$(x-1)^2(x+2)$或$(x-1)(x+2)$

验证：$(A - I)(A + 2I) = 0$

∴极小式为 $m(x) = (x - 1)(x + 2)$

（3）解法如下

引理： A_1，A_2 的极小式为 $m_1(x)$，$m_2(x)$

则 $\begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}$ 的极小式 $m(x)$ 等于 $m_1(x)$，$m_2(x)$ 的最小公倍式

（此引力可推广到 $A_1,A_2,\ldots,A_s$）

$$C = \begin{pmatrix} 1 & 2 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}，\quad A_1 = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \text{极小式为} (x - 1)^2，\quad A_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \text{极小式为} (x - 1)$$

取最小公倍式 $(x - 1)^2$ 为 C 的极小式。

（5）$F = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}_{6\times6}$，$A_1 = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 1 & 2 \end{pmatrix}$，$A_2 = \begin{pmatrix} 4 & 6 & 0 \\ -3 & -5 & 0 \\ -3 & -6 & 1 \end{pmatrix}$

引理： 设 $D = \begin{pmatrix} 0 & 1 & & O \\ & 0 & \ddots & \\ & & \ddots & 1 \\ O & & & 0 \end{pmatrix}$，则 D 的极小式 $m(x) = x^n$

验证：先证 D 的性质（**右推公式**）

设 $A = (a_{ij})_{n\times n} = (\alpha_1,\alpha_2,\ldots,\alpha_n)$

则有 $AD = (0,\alpha_1,\alpha_2,\ldots, \alpha_{n-1})$

$AD^2 = (0,0,\alpha_1,\ldots, \alpha_{n-2})$

$AD^k = (0, \ldots,0,\alpha_1,\ldots, \alpha_{n-k})$

单位向量技巧： ∵$AI = A(e_1,e_2,\ldots,e_n) = (Ae_1, Ae_2,\ldots, Ae_n) = A =(\alpha_1,\alpha_2,\ldots,\alpha_n)$

∴$Ae_1 =\alpha_1, Ae_2 =\alpha_2, \ldots , Ae_n =\alpha_n$

$\Rightarrow AD = A(0,e_1,e_2,\cdots,e_{n-1}) = (0,\alpha_1,\alpha_2,\cdots,\alpha_{n-1})$

同理 $AD^2 = (AD)D = (0,0,\alpha_1,\ldots, \alpha_{n-2})$

可知：$D^{n-1} = (D)D^{n-2} = (0,0, \ldots,e_1) \neq 0$

$D^n = (D)D^{n-1} = 0$，而特征多项式 $f(x) = | xI - D | = x^n$，极小式为某个 x^k

由计算知：极小式为 $m(x)=x^n$

引理 2：设 $B=\begin{pmatrix} b & 1 & & O \\ & b & \ddots & \\ & & \ddots & 1 \\ O & & & b \end{pmatrix}$，则极小式为 $m(x)=(x-b)^n$

$\because B-bI=D=\begin{pmatrix} 0 & 1 & & O \\ & 0 & \ddots & \\ & & \ddots & 1 \\ O & & & 0 \end{pmatrix}$

$\therefore (B-bI)^{n-1}=D^{n-1}\neq 0$，且特征多项式 $f(x)=|xI-D|=(x-b)^n$

极小式为某个 $(x-b)^k$

$\therefore$ 极小式为 $m(x)=(x-b)^n$

复习：（1）可用“分块形”行变换求逆

例：$\begin{pmatrix} A & C \\ 0 & B \end{pmatrix}$ 的逆

（2）“分块形”倍加变换不改变行列式的值

（3）**换位公式**：若 $A=A_{m\times n}$，$B=B_{n\times m}$

则 $|xI_m-kAB|=x^{m-n}|xI_n-kBA|$，$(m\geq n)$

特征值（谱估计）

盖尔圆方法（Ger）

定义：n 阶方阵 $A=(a_{ij})_{n\times n}$ 的第 p 个 Ger（盖尔）半径为

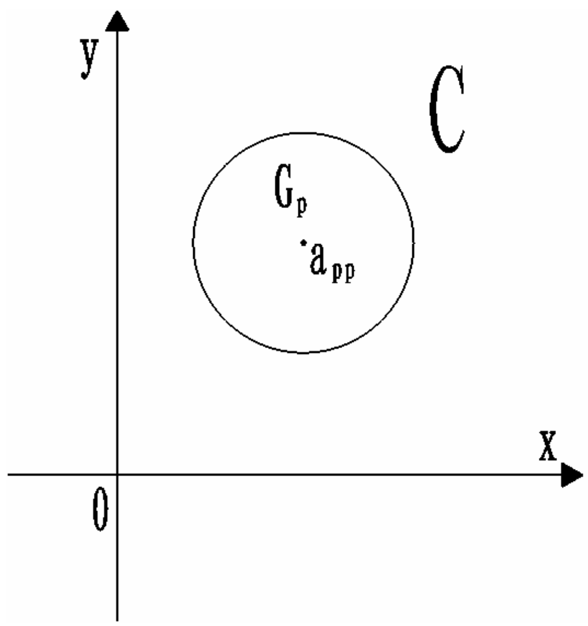
$R_p=|a_{p1}|+|a_{p2}|+\cdots+|\hat{a}_{pp}|+\cdots+|a_{pn}|$，（记号“∧”表示去掉该项）

规定第 p 个 Ger 圆为

$G_p=\{Z \,\big|\, |Z-a_{pp}|\leq R_p\}$，$Z\in C$

第 1 圆盘定理：方阵 $A=(a_{ij})_{n\times n}$ 的全体特征值（谱）都在 A 的 n 个 Ger 圆的并集中。

即：$\sigma(A)=\{\lambda_1,\lambda_2,\cdots,\lambda_n\}\subset G_1\cup G_2\cup\cdots\cup G_n$，（略证）



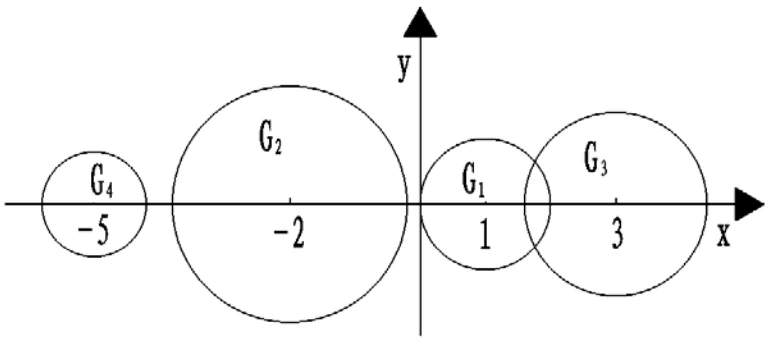
Eg. $A=\begin{pmatrix} 1 & 0.2 & 0.5 & 0.3 \\ 0.6 & -2 & -1 & 0.2 \\ 0.3 & 0.4 & 3 & 0.7 \\ 0.2 & 0.3 & 0.3 & -5 \end{pmatrix}$，估计 $\sigma(A)$的范围。

解：Ger 圆为：$G_1: |Z-a_{11}|=|Z-1|\le R_1=1$

$G_2: |Z-a_{22}|=|Z+2|\le R_2=1.8$

$G_3: |Z-a_{33}|=|Z-3|\le R_3=1.4$

$G_4: |Z-a_{44}|=|Z+5|\le R_4=0.8$



$\sigma(A)\subset G_1\cup G_2\cup G_3\cup G_4$

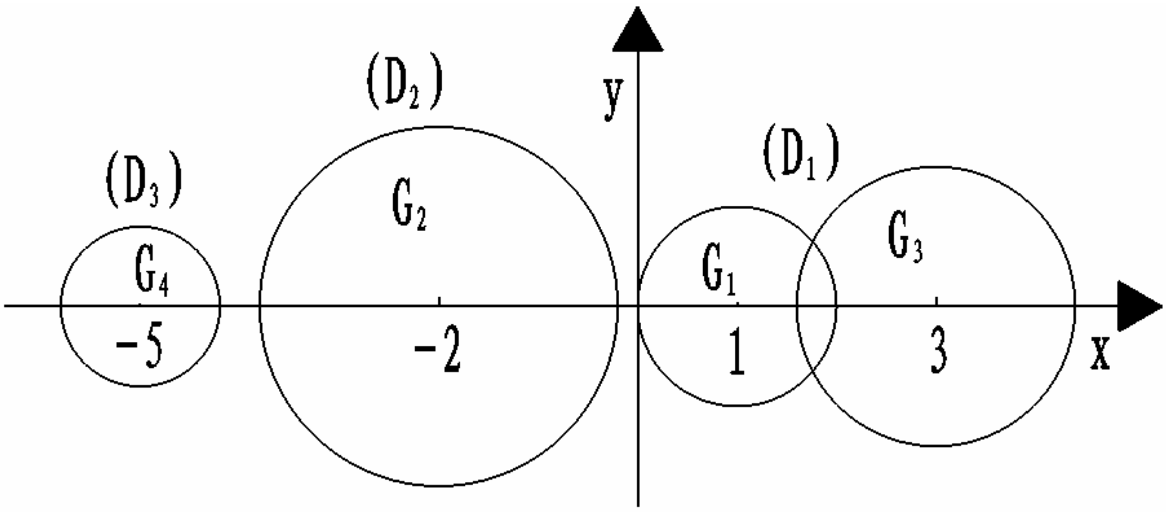
规定：若 A 的 s 个 Ger 圆相连（或相切）在一起，且与其它 n - s 个圆分离，称此 s 个圆的并集为一个连通区域，简称区域。

特别：一个孤立圆也是连通区域。

第 2 圆盘定理：设 D 是 A 的 s 个 Ger 圆构成的区域（分支），则在 D 中恰有 s 个特征

值（含重复）

特别： 一个孤立 Ger 圆中恰有一个特征值（略证）



注： A（指上边例子中）至少有两个实特征值（利用实系数方程的虚根成双出现）

Ex.1. $A=\begin{pmatrix} 9 & 1 & -2 & 1 \\ 0 & 8 & 1 & 1 \\ -1 & 0 & 4 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}$，（1）估计 $\sigma(A)$，（2）说明 A 至少有 2 个实根

Ex.2.估计下列谱 $\sigma(A)$

（1）$A=\begin{pmatrix} 20 & 5 & 0.3 \\ 4 & 10 & 0.5 \\ 2 & 4 & 10i \end{pmatrix}$，（2）$A=\begin{pmatrix} 20 & 5 & 0.6 \\ 4 & 10 & 1 \\ 1 & 2 & 10i \end{pmatrix}$，（3）$A=\begin{pmatrix} 1 & 1 & 2 \\ 2 & 1 & 3 \\ 2 & 3 & 5 \end{pmatrix}$

注： 由于 A 与转置 A^T 有相同的特征值，$\sigma(A)=\sigma(A^T)$，可用 A^T 的 Ger 半径代替 A 的半径。

Ex3.证明 n 阶方阵 $A=\begin{pmatrix} 2 & 2/n & 1/n & \cdots & 1/n \\ 1/n & 4 & 1/n & \cdots & 1/n \\ 1/n & 1/n & 6 & \cdots & 1/n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1/n & 1/n & 1/n & \cdots & 2n \end{pmatrix}$ 恰有 n 个不同实特征值，

且 $|A| > 1 \times 3 \times 5 \times \ldots \times (2n-1)$

§1 Jordan（约当）标准形（简介）

规定： n_k阶上三角阵 $J_k=\begin{pmatrix}\lambda & 1 & & O\\ & \lambda & \ddots & \\ & & \ddots & 1\\ & & & \lambda\end{pmatrix}_{n_k\times n_k}$ 叫做一个 n_k 阶 Jordan 块，λ 是任意复数。

特别： $n_k=1$ 时，对应 1 阶 Jordan 块，$J_1=(\lambda)$是一个数 λ

定义： 称上三角阵 J：$J=\begin{pmatrix}J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_s\end{pmatrix}_{n\times n}$ 为 Jordan 标准形（矩阵），

其中 $J_1,J_2,\ldots,J_s$ 都是 Jordan 块，$(n_1+n_2+\ldots+n_s=n)$

例如：$J=\begin{pmatrix}\begin{pmatrix}2 & 1\\ & 2\end{pmatrix} & \\ & \begin{pmatrix}3 & 1\\ & 3\end{pmatrix}\end{pmatrix}$，$J=\begin{pmatrix}\begin{pmatrix}2 & 1\\ & 2\end{pmatrix} & & \\ & (3) & \\ & & (3)\end{pmatrix}$，$J=(2+i)$

分别为 2 块、3 块、1 块

特别： 对角阵 $A=\begin{pmatrix}(\lambda_1) & & & \\ & (\lambda_2) & & \\ & & \ddots & \\ & & & (\lambda_n)\end{pmatrix}$含有 n 块

注： Jordan 形 J 中的快数是确定的，块的排列次序是任意的。

注： 可证明$\begin{pmatrix}J_1 & 0\\ 0 & J_2\end{pmatrix}\backsim\begin{pmatrix}J_2 & 0\\ 0 & J_1\end{pmatrix}$，相似

注： 全体对角元构成全体特征值 $\sigma(J)$，$\sigma(J)=\sigma(J_1)\cup\sigma(J_2)\cup\ldots\cup\sigma(J_s)$

Jordan 标准形定理： 每个复 n 阶方阵 A 都相似于一个 Jordan 矩阵

即：$P^{-1}AP=J=\begin{pmatrix}J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_s\end{pmatrix}$，$(n_1+n_2+\ldots+n_s=n)$

且除了 Jordan 块次序外 J 由 A 唯一确定，称 J 是 A 的 Jordan 形。

$\sigma(A)=\sigma(J)=\sigma(J_1)\cup\sigma(J_2)\cup\ldots\cup\sigma(J_s)$

利用求秩方法确定 A 的 J

注： 若 $A \backsim J$，则$(A \pm bI)^k \backsim (J \pm bI)^k$，$\mathrm{rank}(A \pm bI)^k = \mathrm{rank}(J \pm bI)^k$

Eg. $A = \begin{pmatrix} 1 & 3 & 0 \\ 0 & 1 & 3 \\ 0 & 0 & 2 \end{pmatrix}$ 求 Jordan 形 J

注： A 的单根对应 1 阶 Jordan

解：先求特征多项式：$|\lambda I - A| = (\lambda - 1)^2(\lambda - 2)$

可设 $A \backsim J = \begin{pmatrix} \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix} & \begin{matrix} 0 \\ 0 \end{matrix} \\ \begin{matrix} 0 & 0 \end{matrix} & (2) \end{pmatrix}$，*是 1 或 0

取 $b = 1$，$(A - I) \backsim (J - I) = \begin{pmatrix} \begin{pmatrix} 0 & * \\ 0 & 0 \end{pmatrix} & \begin{matrix} 0 \\ 0 \end{matrix} \\ \begin{matrix} 0 & 0 \end{matrix} & 1 \end{pmatrix} = \begin{pmatrix} 0 & 3 & 0 \\ 0 & 0 & 3 \\ 0 & 0 & 1 \end{pmatrix}$

$r(J - I) = r(A - I) = 2 \Rightarrow * = 1$，$\quad J = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$

Ex.求下列 Jordan 形

（1）$A = \begin{pmatrix} -2 & -1 & -1 & -1 \\ 2 & 1 & 3 & 2 \\ 1 & 1 & 0 & 1 \\ -1 & -1 & -2 & -2 \end{pmatrix}$

可知：$|\lambda I - A| = \lambda(\lambda - 1)^3$

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

Jordan 形（续）

Jordan 标准形定理： 每个 n 阶复矩阵 A 都相似于一个 Jordan 形

即：$P^{-1}AP = J = \begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_s \end{pmatrix}$，其中 $J_1, J_2, \ldots, J_s$ 为 Jordan 块（可以重复）

且 A 的 Jordan 形 J 由 A 唯一确定（各块次序可任意）

用求秩 $\text{rank}(A \pm bI)^k$ 可确定 J（差分格式）

（1） 求秩：直至有连续两个秩相等为止。

令 $r_0 = n, r_1 = (A - \lambda I), r_2 = (A - \lambda I)^2, \ldots, r_k = (A - \lambda I)^k, \ldots$

（2） 令 $d_0 = r_0 - r_1, d_1 = r_1 - r_2, \ldots, d_k = r_k - r_{k+1}, \ldots$

（3） 令 $l_1 = d_0 - d_1, l_2 = d_1 - d_2, \ldots, l_k = d_{k-1} - d_k, \ldots$

结论：（1）J 中含 λ 的块共有 $d_0 = n - r(A - \lambda I)$ 个

（2）J 中含 λ 的 k 阶块恰有 l_k 个$(k = 1, 2, 3, \ldots)$

Eg. $A = \begin{pmatrix} 3 & -4 & 0 & 1 \\ 4 & -5 & -1 & 0 \\ 0 & 0 & 3 & -2 \\ 0 & 0 & 2 & -1 \end{pmatrix}$，求 Jordan 形 J，（$A \backsim J$）

解：特征多项式：$|xI - A| = \begin{vmatrix} x-3 & 4 \\ -4 & x+5 \end{vmatrix} \bullet \begin{vmatrix} x-3 & 2 \\ -2 & x+1 \end{vmatrix} = (x-1)^2(x+1)^2$

特征值 $\sigma(A) = \{1, 1, -1, -1\}$

求秩数：$\lambda = 1$ 时，$r(A - I) = 3, r(A - I)^2 = 2, r(A - I)^3 = 2$

令 $r_0 = n = 4, r_1 = 3, r_2 = 2, r_3 = 2$

列表：$\begin{array}{ccccc} 4 & & & & \\ & > & 1 & & \\ 3 & & & > & 0 \\ & > & 1 & & \\ 2 & & & > & 1 \\ & > & 0 & & \\ 2 & & & & \end{array}$，可知 J 中含有 $\lambda = 1$ 的块共有 1 个，

且含 $\lambda = 1$ 的 2 阶块有 1 个$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$

同理：$\lambda = -1$ 时，$r(A + I) = 3, r(A + I)^2 = 2, r(A + I)^3 = 2$

列表：$\begin{array}{ccccc} 4 & & & & \\ & > & 1 & & \\ 3 & & & > & 0 \\ & > & 1 & & \\ 2 & & & > & 1 \\ & > & 0 & & \\ 2 & & & & \end{array}$，恰有 1 个 2 阶块$\begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix}$

最后 $J=\begin{pmatrix} J_1 & \\ & J_2 \end{pmatrix}=\begin{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} & \begin{matrix} 0 & 0 \\ 0 & 0 \end{matrix} \\ \begin{matrix} 0 & 0 \\ 0 & 0 \end{matrix} & \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix} \end{pmatrix}$，$A \backsim J$

Eg. $A=\begin{pmatrix} 1 & 0 & 0 & 0 \\ 4 & 2 & 0 & 0 \\ 5 & 3 & 3 & 0 \\ 6 & 3 & 3 & 4 \end{pmatrix}$，有 $n=4$ 个互异特征值 1，2，3，4

必有 $A \backsim J=\begin{pmatrix} (1) & & & \\ & (2) & & \\ & & (3) & \\ & & & (4) \end{pmatrix}$

Eg. $A=\begin{pmatrix} b & & & & O \\ a_1 & b & & & \\ & a_2 & b & & \\ & & \ddots & \ddots & \\ * & & & a_{n-1} & b \end{pmatrix}_{n\times n}$，$(a_i \neq 0)$

A 的 $\sigma(A)=\{b,b,\cdots,b\}$，可设 $A \backsim J=\begin{pmatrix} b & * & & \\ & b & \ddots & \\ & & \ddots & * \\ O & & & b \end{pmatrix}$，*为 1 或 0

$\because A-bI \backsim J-bI \Rightarrow r(J-bI)=r(A-bI)$

$A-bI=\begin{pmatrix} 0 & & & & O \\ a_1 & 0 & & & \\ & a_2 & 0 & & \\ & & \ddots & \ddots & \\ * & & & a_{n-1} & 0 \end{pmatrix}$，$\mathrm{r}(A-bI)=n-1$

$J-bI=\begin{pmatrix} 0 & * & & & \\ & 0 & * & & \\ & & 0 & \ddots & \\ & & & \ddots & * \\ O & & & & 0 \end{pmatrix} \Rightarrow$ 全体*都为 1

$\therefore J=\begin{pmatrix} b & 1 & & \\ & b & \ddots & \\ & & \ddots & 1 \\ O & & & b \end{pmatrix}$

Eg. $A=\begin{pmatrix}2&-1&-1\\2&-1&-2\\-1&1&-2\end{pmatrix}$，$|xI-A|=(x-1)^3$，$\sigma(A)=\{1,1,1\}$

可知 $A\backsim J=\begin{pmatrix}(1)&\\&\begin{pmatrix}1&1\\0&1\end{pmatrix}\end{pmatrix}$

求 P，已知 $P^{-1}AP=J$

解：令 $P=\{X_1,X_2,X_3\}$，由 $AP=PJ$

$$(AX_1,AX_2,AX_3)=(X_1,X_2,X_2+X_3)\Rightarrow\begin{cases}AX_1=X_1\\AX_2=X_2\\AX_3=X_2+X_3\end{cases}\Rightarrow\begin{cases}(A-I)X_1=0\\(A-I)X_2=0\\(A-I)X_3=X_2\end{cases}$$

由$(A-I)X_1=0$ 可得基础解：$\alpha=\begin{pmatrix}1\\1\\0\end{pmatrix}$，$\beta=\begin{pmatrix}1\\0\\1\end{pmatrix}$

通解：$X=k_1\alpha+k_2\beta=\begin{pmatrix}k_1+k_2\\k_1\\k_2\end{pmatrix}=X_2$

求解：$(A-I)X_3=X_2$

增广阵：$(A-I|X_2)=\left(\begin{array}{ccc|c}1&-1&-1&k_1+k_2\\2&-2&-2&k_1\\-1&1&1&k_2\end{array}\right)$

$\Rightarrow\left(\begin{array}{ccc|c}1&-1&-1&k_1+k_2\\0&0&0&k_1+2k_2\\0&0&0&k_1+2k_2\end{array}\right)\Rightarrow k_1+2k_2=0$，可取 $k_1=2, k_2=-1$

$\Rightarrow\left(\begin{array}{ccc|c}1&-1&-1&1\\0&0&0&0\\0&0&0&0\end{array}\right)\Rightarrow x_1-x_2-x_3=1$

取 $X_3=\begin{pmatrix}1\\0\\0\end{pmatrix}$，令 $X_1=\alpha=\begin{pmatrix}1\\1\\0\end{pmatrix}$，$X_2=k_1\alpha+k_2\beta=\begin{pmatrix}1\\2\\-1\end{pmatrix}$

$$P=\begin{pmatrix}1&1&1\\1&2&0\\0&-1&0\end{pmatrix}\Rightarrow P^{-1}AP=J$$

注： A 中元素很小的变化可能引起 Jordan 形很大变化（Butterfly Effect?）

（这就是为什么不能用计算机求 J）

例：$A(\varepsilon)=\begin{pmatrix}\varepsilon & 1\\ 0 & 0\end{pmatrix}$，$(\varepsilon\neq 0)$，可知 $A(\varepsilon)\backsim J(\varepsilon)=\begin{pmatrix}\varepsilon & 0\\ 0 & 0\end{pmatrix}$

令 $\varepsilon\to 0$，求极限 $A(\varepsilon)\to\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}=A_0$，$J(\varepsilon)\to\begin{pmatrix}0 & 0\\ 0 & 0\end{pmatrix}=0$

例 1：求矩阵 $A=\begin{pmatrix}-2 & -1 & -1 & -1\\ 2 & 1 & 3 & 2\\ 1 & 1 & 0 & 1\\ -1 & -1 & -2 & -2\end{pmatrix}$ 的 Jordan 标准形 J

解：求出 A 的特征多项式 $|\lambda I-A|=\lambda(\lambda+1)^3$，全体特征值为 0, -1, -1, -1

若 A 与相似于 Jordan 标准形 J：$A\backsim J$，则它们有相同的特征值，从而有

$$J=\begin{pmatrix}0 & & & \\ & -1 & * & \\ & & -1 & * \\ & & & -1\end{pmatrix}$$，其中的*等于 1 或 0

注： 若 A 的特征值 λ 是单根，则必有 1 阶 Jordan 块(λ)。

由相似关系 $A+I\backsim J+I=\begin{pmatrix}1 & & & \\ & 0 & * & \\ & & 0 & * \\ & & & 0\end{pmatrix}$

可得秩数：$r(J+I)=r(A+I)=rank\begin{pmatrix}-1 & -1 & -1 & -1\\ 2 & 2 & 3 & 2\\ 1 & 1 & 1 & 1\\ -1 & -1 & -2 & -1\end{pmatrix}=2$

可知 $J+I$ 中的 2 个*只有一个等于1，另一个为0，因此

$$J=\begin{pmatrix}0 & & & \\ & -1 & 1 & \\ & & -1 & 0 \\ & & & -1\end{pmatrix} \text{或} J=\begin{pmatrix}0 & & & \\ & -1 & 0 & \\ & & -1 & 1 \\ & & & -1\end{pmatrix}$$

这两个 J 本质上是相同的（都含有 3 个 Jordan 块），只是 Jordan 块的排列次序不同。

注： 如果两个 Jordan 矩阵只是 Jordan 块的次序不同，则认为它们本质上相同。在这个意

义上

本题中的 J 由 A 唯一决定.可写 $A \backsim J = \begin{pmatrix} 0 & & & \\ & -1 & 1 & \\ & & -1 & \\ & & & -1 \end{pmatrix}$

另外，可找到一个可逆阵 $P = \begin{pmatrix} -1 & 0 & 1 & 1 \\ 3 & -1 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ -2 & 1 & 0 & -1 \end{pmatrix}$ 使得

$AP = P\begin{pmatrix} 0 & & & \\ & -1 & 1 & \\ & & -1 & \\ & & & -1 \end{pmatrix} = PJ$ ，即 $P^{-1}AP = J$

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

（1）求 Jordan 标准形 J，并判断 A 可否对角化；

（2）求相似变换阵 P，使 $P^{-1}AP = J$

解：A 的特征多项式为：$|\lambda I - A| = (\lambda-2)(\lambda-1)^2$，特征值为 2, 1, 1。所以

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

注： 若 A 的特征值 λ 是单根，则必有 1 阶 Jordan 块(λ)。

由于 J 含有 2 阶 Jordan 块，可知 A 不能对角化.

令 $P = (X_1, X_2, X_3)$，$X_i(i = 1, 2, 3)$为列向量，则 $AP = PJ$，即

$$A(X_1, X_2, X_3) = (X_1, X_2, X_3)\begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix},$$

即 $AX_1 = 2X_1, AX_2 = X_2, AX_3 = X_2 + X_3$。

所以 X_1 为 A 的关于 $\lambda = 2$ 的特征向量；X_2 为 A 的关于 $\lambda = 1$ 的特征向量；

X_3 是非齐次方程$(A - I)\,X_3 = X_2$ 的解（广义特征向量）。

由$(2I - A)\,X_1 = 0$ 解出 $X_1 = (0, 0, 1)^T$,

由$(I-A)X_2=0$解出$X_2=(1,2,-1)^T$，

由$(A-I)X_3=X_2$解出$X_3=(-1,-1,0)^T$，或$X_3=(0,1,-1)^T$

令$P=(X_1,X_2,X_3)=\begin{pmatrix}0&1&-1\\0&2&-1\\1&-1&0\end{pmatrix}$，或$P=\begin{pmatrix}0&1&0\\0&2&1\\1&-1&-1\end{pmatrix}$可知

$$AP=P\begin{pmatrix}2&0&0\\0&1&1\\0&0&1\end{pmatrix}=PJ\text{，即 }P^{-1}AP=J.$$

例 3　试证：每个 Jordan 块J_k都相似于它的转置$J_k^{\ T}$.

Pf：计算可知

$$\begin{bmatrix}0&&&1\\&&\iddots&\\&1&&\\1&&&0\end{bmatrix}\begin{bmatrix}\lambda&1&&0\\&\ddots&\ddots&\\&&\ddots&1\\0&&&\lambda\end{bmatrix}\begin{bmatrix}0&&&1\\&&\iddots&\\&1&&\\1&&&0\end{bmatrix}=\begin{bmatrix}\lambda&&&0\\1&\ddots&&\\&\ddots&\ddots&\\0&&1&\lambda\end{bmatrix}.$$

注： 由此例可知，每个 Jordan 矩阵J都相似于它的转置：$J\backsim J^T$（下三角矩阵）

利用此例 3 与 Jordan 标准形定理可得：

推论 3： 每个方阵A都相似于它的转置A^T：　$A\backsim A^T$

例 4　设k为自然数，$A^k=0$，试证：$|A+I|=1$

证：由$A^k=0$知A的特征值全为零，从而 Jordan 标准形J的主对角线元素全为零。利用$A=PJP^{-1}$，可知$|A+I|=|PJP^{-1}+I|=|P||J+I||P^{-1}|=1$

补充结论：

每个 Jordan 块$J_k(b)=\begin{pmatrix}b&1&&\\&b&\ddots&\\&&\ddots&1\\O&&&b\end{pmatrix}_{k\times k}$的极小式为$m(x)=(x-b)^k$

每个块$J_k(b)$相似于转置$J_k(b)^T=\begin{pmatrix}b&&&O\\1&b&&\\&\ddots&\ddots&\\&&1&b\end{pmatrix}$

Pf：取 $P=\begin{pmatrix} O & & & 1 \\ & & 1 & \\ & \iddots & & \\ 1 & & & O \end{pmatrix}_{k\times k}$ 可知 $P^{-1}=P$（正交阵）

计算知：$\begin{pmatrix} b & 1 & & \\ & b & \ddots & \\ & & \ddots & 1 \\ & & & b \end{pmatrix}\begin{pmatrix} O & & & 1 \\ & & 1 & \\ & \iddots & & \\ 1 & & & O \end{pmatrix}=\begin{pmatrix} O & & & 1 \\ & & 1 & \\ & \iddots & & \\ 1 & & & O \end{pmatrix}\begin{pmatrix} b & & & \\ 1 & b & & \\ & \ddots & \ddots & \\ & & 1 & b \end{pmatrix}$

$J_kP=PJ_k^T \Rightarrow P^{-1}J_kP=J_k^T$，$J_k \backsim J_k^T$

练习：$J=\begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_s \end{pmatrix} \backsim \begin{pmatrix} J_1^T & & & \\ & J_2^T & & \\ & & \ddots & \\ & & & J_s^T \end{pmatrix}=J^T$

每个 A 相似于 A^T

$\because A\backsim J \Leftrightarrow A^T \backsim J^T \backsim J \Rightarrow A \backsim A^T$

Ex.1.已知 5 阶阵 A 有条件 $r(A)=3, r(A^2)=2, r(A+I)=4, r(A+I)^2=3$，求 Jordan 形。

2.求下列 Jordan 形

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

Jordan 形公式与结论

参考书：（1）Horn and Johnson: “Matrix Analysis”（矩阵分析）

§3 Jordan 形的一个证明（用分块矩阵方法）

（2）李尚志《线性代数》P370 定理 1（差分格式求 Jordan 形）

利用$(xI-A)$的初等因子求 Jordan 形

定义：若 $g(x)=(x-b_1)^{k_1}(x-b_2)^{k_2}\cdots(x-b_s)^{k_s}$，$b_1, b_2, \dots, b_s$ 互不相同

称 $(x-b_1)^{k_1},(x-b_2)^{k_2},\cdots,(x-b_s)^{k_s}$ 为 $g(x)$的初等因子。

定义：（1）若 Jordan 块 $J_k=\begin{pmatrix} b & 1 & & \\ & b & \ddots & \\ & & \ddots & 1 \\ & & & b \end{pmatrix}_{n_k\times n_k}$，称 $(x-b)^{n_k}$ 为 J_k 的初等因子

（2）若 $A\backsim J=\begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_s \end{pmatrix}$（Jordan 形），称 $J_1, J_2, \dots ,J_s$ 的初等因子

$(x-b_1)^{n_1},(x-b_2)^{n_2},\cdots,(x-b_s)^{n_s}$ 为 A 的全体初等因子。

注：A 的初等因子$(x-b)^k$ 与 Jordan 块一一对应

例如：因子 $(x-b)^k\leftrightarrow\begin{pmatrix} b & 1 & & \\ & b & \ddots & \\ & & \ddots & 1 \\ & & & b \end{pmatrix}_{k\times k}$

特别单因子 $(x-b)\leftrightarrow(b)$，（1 阶块）

初等因子定理：若$(xI-A)$可用初等变换化为对角形

$$(xI-A)\to\begin{pmatrix} g_1(x) & & & \\ & g_2(x) & & \\ & & \ddots & \\ & & & g_n(x) \end{pmatrix}_{n\times n}$$

则（1）$g_1(x), g_2(x), \dots ,g_n(x)$的全体初等因子（含重复）恰为 A 的初等因子。

（2）行列式$|xI-A|=g_1(x)g_2(x)\dots g_n(x)=$ 全体初等因子的积。

$(xI-A)$有 3 类初等变换

（1）互换行（或列）（2）用常数 $k\neq 0$ 乘某一行（或列）

（3）倍加法：用多项式 $k(x)$乘第 j 行后加到第 i 行。（记 $r_i+k(x)r_j$）（**注**：第 j 行不变）

Eg. $A=\begin{pmatrix} 2 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & -1 & 3 \end{pmatrix}$，求 Jordan 形 J

$$\left(xI-A\right)=\begin{pmatrix} x-2 & 0 & 0 \\ -1 & x-1 & -1 \\ -1 & 1 & x-3 \end{pmatrix}\xrightarrow{\text{互换}r_1,r_2}\begin{pmatrix} 1 & -(x-1) & 1 \\ x-2 & 0 & 0 \\ -1 & 1 & x-3 \end{pmatrix}$$

解：令 $\xrightarrow[r_3+r_1]{r_2-(x-2)r_1}\begin{pmatrix} 1 & -(x-1) & 1 \\ 0 & (x-1)(x-2) & -(x-2) \\ 0 & -(x-2) & (\mathrm{x}-2) \end{pmatrix}\xrightarrow{\text{列变换}}\begin{pmatrix} 1 & 0 & 0 \\ 0 & (x-1)(x-2) & -(x-2) \\ 0 & -(x-2) & (x-2) \end{pmatrix}$

$$\xrightarrow{c_2+c_3}\begin{pmatrix} 1 & 0 & 0 \\ 0 & (x-1)(\mathrm{x}-2) & -(x-2) \\ 0 & 0 & (x-2) \end{pmatrix}\xrightarrow{r_2+r_3}\begin{pmatrix} 1 & 0 & 0 \\ 0 & (x-2)^2 & 0 \\ 0 & 0 & (x-2) \end{pmatrix}$$

全体初等因子为$(x-2)^2, (x-2)$

$\Rightarrow A \backsim J=\begin{pmatrix} 2 & 1 & \\ 0 & 2 & \\ & & 2 \end{pmatrix}$，（Jordan 形）

Eg.把 $A=\begin{pmatrix} b & 1 & & \\ & b & \ddots & \\ & & \ddots & 1 \\ & & & b \end{pmatrix}_{n\times n}$ 的$(xI - A)$化成对角形。

解：

$$\left(xI - b\right)=\begin{pmatrix} -b & -1 & \cdots & 0 \\ 0 & x-b & \ddots & \vdots \\ \vdots & \vdots & \ddots & -1 \\ 0 & 0 & \cdots & x-b \end{pmatrix}\xrightarrow{\text{行变换}}\begin{pmatrix} x-b & -1 & \cdots & 0 \\ (x-b)^2 & 0 & \ddots & \vdots \\ \vdots & \vdots & \ddots & -1 \\ (x-b)^n & 0 & \cdots & 0 \end{pmatrix}$$

$$\xrightarrow{\text{列变换}}\begin{pmatrix} 0 & -1 & \cdots & 0 \\ 0 & 0 & \ddots & \vdots \\ \vdots & \vdots & \ddots & -1 \\ (x-b)^n & 0 & \cdots & 0 \end{pmatrix}\xrightarrow{\text{互换行}}\begin{pmatrix} -1 & & & \\ & -1 & & \\ & & \ddots & \\ & & & (x-b)^{\mathrm{n}} \end{pmatrix}$$

Jordan 形与极小式

引理：（1）Jordan 块 $J_k\left(b\right)=\begin{pmatrix} b & 1 & & \\ & b & \ddots & \\ & & \ddots & 1 \\ & & & b \end{pmatrix}_{k\times k}$ 的极小式为 $m(x)=(x-b)^k$

（2）设 $A \backsim J=\begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_s \end{pmatrix}$（Jordan 形），则 A 的极小式 $m(x)=$ 全体初等因子的最小公倍

推论：A 的极小式 $m(x)$分解后的初等因子是 A 的部分初等因子，可用极小式求出 3 阶阵 $A = A_{3\times 3}$ 的 Jordan 形

Eg. $A=\begin{pmatrix}1&2&-1\\0&1&1\\0&0&-2\end{pmatrix}$，$|xI-A|=(x-1)^2(x+2)$

计算：$(A-I)(A+2I)=\begin{pmatrix}0&2&-1\\0&0&1\\0&0&-3\end{pmatrix}\begin{pmatrix}3&2&-1\\0&3&1\\0&0&0\end{pmatrix}\neq 0$

只有$(A-I)^2(A+2I)=0$，（Cayley 公式）

$\Rightarrow A$的极小式 $m(x)=(x-1)^2(x+2)$

$\Rightarrow A$的初等因子$(x-1)^2, (x+2)$

$\Rightarrow A\backsim J=\begin{pmatrix}1&1&\\0&1&\\&&2\end{pmatrix}$，（Jordan 形）

Eg. $A=\begin{pmatrix}2&0&0\\1&1&1\\1&-1&3\end{pmatrix}$，$|xI-A|=(x-2)^3$

计算：$(A-2I)(A-2I)=0\Rightarrow m(x)=(x-2)^2$ 有一个 Jordan 块$\begin{pmatrix}2&1\\0&2\end{pmatrix}$

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

引理：设 $A=\begin{pmatrix}A_1&&&\\&A_2&&\\&&\ddots&\\&&&A_s\end{pmatrix}$，则 A 的极小式=各块极小式的最小公倍

且各块 $A_1, A_2, \ldots, A_n$ 的 Jordan 块也是 A 的 Jordan 块

Ex.求 $A=\begin{pmatrix}A_1&\\&A_2\end{pmatrix}_{6\times 6}$ 的 Jordan 形 $A_1=\begin{pmatrix}1&2&-1\\0&1&1\\0&0&2\end{pmatrix}$，$A_2=\begin{pmatrix}2&0&0\\1&1&1\\1&-1&3\end{pmatrix}$

对角化的条件（判定）

定义： 若有 P 使得 $P^{-1}AP=\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{pmatrix}$ 称 A 为可对角化的（也称 A 是单纯的）

引理： 阶数大于 1 的 Jordan 块 J_k 不可对角化（Jordan 块可对角化 $\Leftrightarrow$ 阶数为 1）

Pf：设 $J_k=\begin{pmatrix}b & 1 & & \\ & b & \ddots & \\ & & \ddots & 1 \\ & & & b\end{pmatrix}_{k\times k}$，$(k>1)$，特征根为 $b, b, \dots, b$

若 J_k 可对角化：$P^{-1}J_kP=\begin{pmatrix}b & & & \\ & b & & \\ & & \ddots & \\ & & & b\end{pmatrix}=bI\Rightarrow J_k=P(bI)P^{-1}=bI$，矛盾。

定理：（1）若方阵 A 的 Jordan 形中有阶数大于 1 的块，则 A 不能对角化。

（2）A 可对角化 $\Leftrightarrow$ Jordan 块都是 1 阶的，此时 $A\backsim J=\begin{pmatrix}(\lambda_1) & & & \\ & (\lambda_2) & & \\ & & \ddots & \\ & & & (\lambda_n)\end{pmatrix}$

（3）A 可对角化 $\Leftrightarrow$ A 的极小式无重根。

（因为：极小式中的初等因子是全体初等因子的公倍）

（4）若 $f(x)$ 是 A 的一个零化式且 $f(x)$ 无重根，则 A 可对角化

（因为零化式为极小式的倍式）

Eg. $A=\begin{pmatrix}4 & 6 & 0\\ -3 & -5 & 0\\ -3 & -6 & 1\end{pmatrix}$ 判定 A 可否对角化

解：$\because |xI-A|=(x-1)^2(x+2)$

计算：$(A-I)(A+2I)=0\Rightarrow$ 极小式 $m(x)=(x-1)(x+2)\Rightarrow A$ 可对角化：

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

Ex.1.若 $A^2-3A+2I_n=0$，则 A 可对角化

2.若 $A^2=2A$，则 A 可对角化

3.《矩阵分析（史荣昌等）》P111 7（1）（3） 8（1）（3） P110 3 4

§2 线性变换与矩阵

线性空间（向量空间）定义：

集合 V 中有加法“+”与数乘“$k(\bullet)$” $k\in R(C)$，具有 8 条规则（公理）：其中 V 中元素叫“向量”（广元）。

子空间条件：

设 $W\subset V$（空间），若 W 对加法与倍数（数乘）封闭，则 W 是 V 的子空间，生成（张成）自空间，任取 $\alpha_1, \alpha_2, \dots, \alpha_s\in V$，称 $W=\text{span}(\alpha_1, \alpha_2, \dots, \alpha_s)=\{$全体组合 $\alpha=k_1\alpha_1+k_2\alpha_2+\dots+k_s\alpha_s\}$，$(k_1, k_2, \dots, k_s\in R)$，为 $\alpha_1, \alpha_2, \dots, \alpha_s$ 的生成空间。

可验证： W 对加法与倍数都封闭。

Eg.（1）$m\times n$ 矩阵空间：$R^{m\times n}$，$C^{m\times n}$

方阵：$R^{n\times n}$，$C^{n\times n}$

（2）数组空间：$R^n=\left\{x=\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}, x_i\in R\right\}$，$C^n=\left\{z=\begin{pmatrix}z_1\\z_2\\\vdots\\z_n\end{pmatrix}, z_i\in C\right\}$

子空间例子：

（1）核空间（零空间，解空间）：

设 $A=A_{m\times n}\in R^{m\times n}$，规定：$N(A)=A^{-1}(0)\underline{\underline{\Delta}}\{x\in R^n | Ax=0\}$（对加法、倍数封闭）

（2）值空间（列空间）：$R(A)=\{$ 全体 $y=Ax | x\in R^n\}$

注： 把 $A=A_{m\times n}$ 按列 $\alpha_1, \alpha_2, \dots, \alpha_n$ 改写 $A=(\alpha_1, \alpha_2, \dots, \alpha_n)$，令 $x=\begin{pmatrix}k_1\\k_2\\\vdots\\k_n\end{pmatrix}\in R^n$，

写 $Ax=(\alpha_1,\alpha_2,\cdots,\alpha_n)\begin{pmatrix}k_1\\k_2\\\vdots\\k_n\end{pmatrix}=\sum_{i=1}^n k_i\alpha_i \Rightarrow$ 值空间 $R(A)=\{Ax | x\in R^n\}=\left\{y=\sum_{i=1}^n k_i\alpha_i \middle| k_i\in R\right\}$

（全体线性组合），即 $R(A)=\text{span}(\alpha_1, \alpha_2, \dots, \alpha_n)$，（由 A 的列生成）也叫 A 的列向量。

注： $R^n = span(e_1, e_2, \cdots, e_n) = \left\{ x = \sum_{i=1}^{n} x_i e_i \middle| x_i \in R \right\}$

"相关组"与"无关组"定义。

"表示"与"组合"： 若 $\alpha = \sum_{i=1}^{s} k_i \alpha_i$，称 α 可由 $\alpha_1, \alpha_2, \ldots, \alpha_s$ "表示"

也说 α 是 $\alpha_1, \alpha_2, \ldots, \alpha_s$ 的"组合"

极大无关组： 若大组 S 中有 r 个无关向量 $\alpha_1, \alpha_2, \ldots, \alpha_r$，且任何 $r+1$ 个向量都相关

则称 $\alpha_1, \alpha_2, \ldots, \alpha_r$ 是一个极大无关组，r 叫 S 的秩数 $\mathrm{rank}(S) = r$

注： 大组中任 2 个极大无关组互相表示（等价）

唯一表示定理： 若 $\alpha_1, \alpha_2, \ldots, \alpha_s$ 无关，且 $\alpha_1, \alpha_2, \ldots, \alpha_s, \beta$ 相关，则有唯一表示：$\beta = \sum_{i=1}^{s} k_i \alpha_i$

（系数唯一），此时，规定 $k_1, k_2, \ldots, k_s$ 为 β 的坐标。

注： 坐标常写成列 $\begin{pmatrix} k_1 \\ k_2 \\ \vdots \\ k_s \end{pmatrix} = (k_1, k_2, \cdots, k_s)^T$

基、维数、坐标定义： 设空间 V 中有 n 个无关向量 $\alpha_1, \alpha_2, \ldots, \alpha_n$，且任何 $n+1$ 个元都相关，则称$(\alpha_1, \alpha_2, \ldots, \alpha_n)$（有次序）为 V 中一个基，且 n 叫维数，记 $\dim V = n$

注： 空间的基 $\alpha_1, \alpha_2, \ldots, \alpha_n$ 就是 V 中的一个极大无关组（有次序），且维数就是秩数：

$\dim V = \mathrm{rank}(V) = n$

坐标定义： 设空间 V 中有 n 个无关向量 $\alpha_1, \alpha_2, \ldots, \alpha_n$，且任何 $n+1$ 个向量必相关，则

任一 $\alpha \in V$ 必有唯一表示 $\alpha = \sum_{i=1}^{n} x_i \alpha_i$，称列向量 $(x_1, x_2, \cdots, x_n)^T = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ 为 α

的坐标（此时 V 的基为 $\alpha_1, \alpha_2, \ldots, \alpha_n$）

注： 向量 α 与坐标是一一对应（唯一表示定理）

基元 $\alpha_1, \alpha_2, \ldots, \alpha_n$ 与单位向量 $e_1, e_2, \ldots, e_n$ 对应，设 $\alpha_1, \alpha_2, \ldots, \alpha_n$ 为 V 中基，

$$\text{则}\begin{cases}\alpha_1 = 1\bullet\alpha_1 + 0\bullet\alpha_2 + \cdots + 0\bullet\alpha_n \xrightarrow{\text{对应}} e_1 = \begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix} \in R^n \\ \alpha_2 = 0\bullet\alpha_1 + 1\bullet\alpha_2 + \cdots + 0\bullet\alpha_n \xrightarrow{\text{对应}} e_2 = \begin{pmatrix}0\\1\\\vdots\\0\end{pmatrix} \in R^n \\ \qquad\qquad\vdots \\ \alpha_n = 0\bullet\alpha_1 + 0\bullet\alpha_2 + \cdots + 1\bullet\alpha_n \xrightarrow{\text{对应}} e_n = \begin{pmatrix}0\\0\\\vdots\\1\end{pmatrix} \in R^n \end{cases}$$

空间同构：若 V 与 W 是空间，$\varphi: V \to W$ 是映射

（1）φ 是一一对应，

（2）$\varphi(\alpha+\beta) = \varphi(\alpha)+\varphi(\beta)$，$\varphi(k\alpha) = k\varphi(\alpha)$，（保加法、保倍数）

称 φ 是 V 到 W 的同构，记 $V\underset{=}{\varphi}W$

性质：（1）同构 φ 把无关组变成无关组（把基变成基）$\Rightarrow$（保坐标）

（2）φ 把相关组变成相关组

Pf:（1）设 $\alpha_1, \alpha_2, \ldots, \alpha_n$ 为无关组，若 $\sum_{i=1}^{n} k_i\alpha_i = 0$，则必有 $k_1 = 0, k_2 = 0, \ldots, k_n = 0$

设 $\sum_{i=1}^{n} k_i\varphi(\alpha_i) = 0$，（$\varphi$ 是同构）$\Leftrightarrow \sum_{i=1}^{n}\varphi(k_i\alpha_i) = 0 \Leftrightarrow \varphi\left(\sum_{i=1}^{n}k_i\alpha_i\right) = 0 = \varphi(0)$

$\Leftrightarrow \sum_{i=1}^{n} k_i\alpha_i = 0$（一一对应）$\Rightarrow k_1 = k_2 = \cdots = k_n = 0 \Rightarrow \varphi(\alpha_1), \varphi(\alpha_2), \cdots, \varphi(\alpha_n)$无关

同构定理：（1）任何 n 维（实）空间 V 都与 R^n 同构

（2）任 2 个 n 维空间 V 与 W 同构（利用（1）与传递性）

Pf:（1）任取 $\alpha \in V$ $(\alpha_1, \alpha_2, \ldots, \alpha_n)$是个固定的基，有 $\alpha = \sum_{i=1}^{n} x_i\alpha_i$

规定坐标映射 $\varphi\colon V\to R^n$ 使得 $\varphi(\alpha)=\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}=x\in R^n$

可知：φ 是同构①φ 是一一的（唯一定理）

②设 $\beta=\sum_{i=1}^n y_i\alpha_i$ ， $\alpha=\sum_{i=1}^n x_i\alpha_i$

$$\alpha+\beta=\sum_{i=1}^n(x_i+y_i)\alpha_i ，\ \varphi(\alpha+\beta)=\begin{pmatrix}x_1+y_1\\x_2+y_2\\\vdots\\x_n+y_n\end{pmatrix}=x+y=\varphi(\alpha)+\varphi(\beta)$$

且 $\varphi(k\alpha)=\begin{pmatrix}kx_1\\kx_2\\\vdots\\kx_n\end{pmatrix}=kx=k\varphi(\alpha)$

注：利用同构可用 $R^n(C^n)$代表空间 V

线性映射：若 V 与 W 是空间，$\varphi\colon V\to W$ 是映射，且 $\varphi(\alpha+\beta)=\varphi(\alpha)+\varphi(\beta)$，$\varphi(k\alpha)=k\varphi(\alpha)$，

（保加法、保倍数），称 φ 是 V 到 W 的线性映射

特别：若 $V=W$（同一空间）称线性映射 $\varphi\colon V\to W$ 为线性变换

记号：$L(V, W)$（V 到 W 的全体线性映射），$L(V, V)$（全体线性变换），

可写 $\varphi\in L(V, W)$或 $\varphi\colon V\to W$

例子：

恒同映射：$I_V\colon V\to W$ 使得 $I_V(\alpha)=\alpha$，$\alpha\in V$（是线性的）

零射：$0\colon V\to W$ 使得 $0(\alpha)=\vec{0}\in W$ （$\forall\alpha\in V$）（是线性的）

矩阵映射：令 $A=A_{m\times n}\in R^{m\times n}$，任取 $x=\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}\in R^n$

规定 $\mathscr{A}: R^n \to R^m$ 如下 $\mathscr{A}(x) \triangleq Ax = A\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in R^n$

$\mathscr{A}(\alpha+\beta) = A(\alpha+\beta) = A\alpha + A\beta = \mathscr{A}(\alpha) + \mathscr{A}(\beta)$

且 $\mathscr{A}(k\alpha) = k\mathscr{A}(\alpha) \Rightarrow \mathscr{A}$ 为线性的

以后常把 $\mathscr{A}$ 写成映射：$\mathscr{A}: R^n \to R^m \qquad x \to Ax$

值空间 $R(A) = \{Ax \mid x \in R^n\} \subset R^m$

核：$N(A) = A^{-1}(0) = \{x \mid Ax = 0\} \subset R^n$

特别：n 阶方阵 $A = A_{n\times n} \in R^{n\times n}$，有线性变换 $A: R^n \to R^m \quad A(x) = Ax \in R^n$

Ex.预习《矩阵分析（史荣昌等）》P24-44

P68 1 3 4 6 8 9

线性映射（变换）性质： 设 $\varphi: V \to W$ 为线性

保组合

把相关组变成相关组

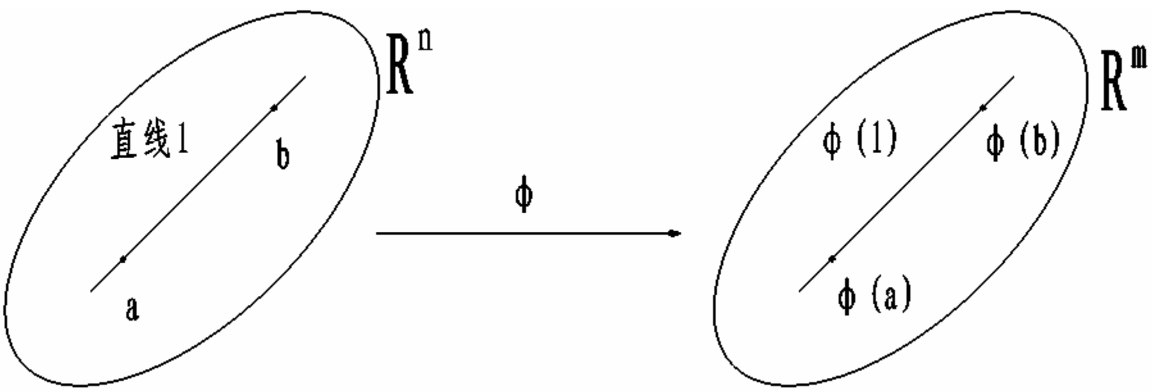
例如：$\varphi\left(\sum_{i=1}^{n} k_i \alpha_i\right) = \sum_{i=1}^{n} \varphi(k_i \alpha_i)$（保组合系数）

几何定义： 线性映射（变换）$\varphi: R^n \to R^m$

①φ 把直线变成直线（或退化直线成一点）

②φ 把平行线变成平行（重合）线

Pf:



设 a、b 决定直线 $l = \{a + t(b-a) \mid t \in R\}$

$\Rightarrow$像$\varphi(l)=\{\varphi(a)+t[\varphi(b)-\varphi(a)] | t\in R\}$也是直线或退为一点

再设α、β是2条直线l_1、l_2的方向向量，若$l_1 /\!/ l_2 \Rightarrow \alpha /\!/ \beta \Rightarrow \alpha = k\beta$

$\varphi(\alpha)=k\varphi(\beta) \Rightarrow \varphi(\alpha) /\!/ \varphi(\beta) \Rightarrow \varphi(l_1) /\!/ \varphi(l_2)$（或重合）

命题： 若φ为线性的，且$\varphi(\alpha_1), \varphi(\alpha_2), \ldots, \varphi(\alpha_n)$线性无关，则$\alpha_1, \alpha_2, \ldots, \alpha_n$也无关

Pf：若$\alpha_1,\alpha_2,\cdots,\alpha_n \Rightarrow \varphi(\alpha_1),\varphi(\alpha_2),\cdots,\varphi(\alpha_n)$相关

注： 若$\varphi: V\to W$为线性，且φ为一一的，则φ“把无关组变成无关组”

规定： ①任取广元$\otimes_1,\otimes_2,\cdots,\otimes_n$称记号$(\otimes_1,\otimes_2,\cdots,\otimes_n)$为一个广行

②若有“组合”$\alpha=\sum_{i=1}^{n}k_i\otimes_i$，称公式$\alpha=(\otimes_1,\otimes_2,\cdots,\otimes_n)\begin{pmatrix}k_1\\k_2\\\vdots\\k_n\end{pmatrix}$为广阵格式

要点： 要把组合系数写成列

若$\alpha_1, \alpha_2, \ldots, \alpha_p$可由$\otimes_1,\otimes_2,\cdots,\otimes_n$组合表示：$\begin{cases}\alpha_1=\sum_{i=1}^{n}k_{i1}\otimes_i\\\alpha_2=\sum_{i=1}^{n}k_{i2}\otimes_i\\\vdots\\\alpha_p=\sum_{i=1}^{n}k_{ip}\otimes_i\end{cases}$

规定广阵格式如下：$(\alpha_1,\alpha_2,\cdots,\alpha_p)=(\otimes_1,\otimes_2,\cdots,\otimes_n)B_{n\times p}$

$$B_{n\times p}=\left(\begin{pmatrix}k_{11}\\k_{12}\\\vdots\\k_{1p}\end{pmatrix}\begin{pmatrix}k_{21}\\k_{22}\\\vdots\\k_{2p}\end{pmatrix}\cdots\begin{pmatrix}k_{n1}\\k_{n2}\\\vdots\\k_{np}\end{pmatrix}\right)_{n\times p}\underline{\underline{\text{记为}}}(\beta_1,\beta_2,\cdots,\beta_p)$$

即$(\alpha_1,\alpha_2,\cdots,\alpha_p)=(\otimes_1,\otimes_2,\cdots,\otimes_n)\begin{pmatrix}k_{11}&k_{21}&\cdots&k_{n1}\\k_{12}&k_{22}&\cdots&k_{n2}\\\vdots&\vdots&\ddots&\vdots\\k_{1p}&k_{2p}&\cdots&k_{np}\end{pmatrix}$

广阵原理： ①一切线性组合都有广阵格式。

②若广元 $\alpha_1, \alpha_2, \ldots, \alpha_p$ 可由 $e_1, e_2, \ldots, e_n$ 表示，

则有广阵格式$(\alpha_1, \alpha_2, \ldots, \alpha_p) = (e_1, e_2, \ldots, e_n)\ B_{n\times p}$

要求： 系数阵 $B_{n\times p}$ 中的列就是组合系数

性质：（引理 1） 设 $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ 是广元，I_n 为单位阵，则

（1）$(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n) = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)\ I_n$

（2）结合公式：$[(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)B]C = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)(BC)$

（3）消去律：（唯一性公式）：若 $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ 无关（基元），且

$$(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)B = (\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)C \Leftrightarrow B = C$$

要证：$(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)B \underline{\underline{\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n\text{无关}}} (\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)C \Rightarrow B = C$

先证 B、C 只有一列 $B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$、$C = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$

令 $\alpha = (\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)B = (\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)C \Rightarrow \alpha = \sum_{i=1}^{n} b_i\varepsilon_i = \sum_{i=1}^{n} c_i\varepsilon_i \Rightarrow b_i = c_i \Rightarrow B = C$

设 B、C 恰有 2 列 $B = (\beta_1, \beta_2)_{n\times 2}$、$C = (\gamma_1, \gamma_2)_{n\times 2}$

由$(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)(\beta_1, \beta_2) = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)(\gamma_1, \gamma_2)$

$\Leftrightarrow (\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)\beta_1 = (\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)\gamma_1, (\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)\beta_2 = (\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)\gamma_2$

$\Rightarrow \beta_1 = \gamma_1, \beta_2 = \gamma_2 \Rightarrow B = (\beta_1,\beta_2) = (\gamma_1,\gamma_2) = C$

记号规定： $\varphi(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n) \underline{\underline{\Delta}} (\varphi(\varepsilon_1),\varphi(\varepsilon_2),\cdots,\varphi(\varepsilon_n))$（$\varphi$: $V \to W$ 是线性映射）

性质（4）： 若 φ 是线性映射，则$(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)B = [\varphi(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)]B$（右提取公式）

Pf：先设 B 只有 1 列 $B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$

$(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)B = \sum_{i=1}^{n} b_i\varepsilon_i$

$$\Rightarrow \varphi[(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)B]=\varphi\left(\sum_{i=1}^{n}b_i\varepsilon_i\right)=\sum_{i=1}^{n}b_i\varphi(\varepsilon_i)=(\varphi(\varepsilon_1),\varphi(\varepsilon_2),\cdots,\varphi(\varepsilon_n))\begin{pmatrix}b_1\\b_2\\\vdots\\b_n\end{pmatrix}$$

$$\Rightarrow \varphi[(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)B]=(\varphi(\varepsilon_1),\varphi(\varepsilon_2),\cdots,\varphi(\varepsilon_n))B=\varphi(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)B$$

若 $B=(\beta_1,\beta_2)_{n\times 2}$ 恰有 2 列，可同样证明

Eg.设 $\varepsilon_1=\begin{pmatrix}1&0\\0&1\end{pmatrix}$，$\varepsilon_2=\begin{pmatrix}0&1\\1&0\end{pmatrix}$，$\alpha=\begin{pmatrix}1&2\\2&1\end{pmatrix}\in R^{2\times 2}$

$\alpha=\varepsilon_1+2\varepsilon_2\Rightarrow\alpha=(\varepsilon_1,\varepsilon_2)\begin{pmatrix}1\\2\end{pmatrix}$，或 $\alpha=\left(\begin{pmatrix}1&0\\0&1\end{pmatrix},\begin{pmatrix}0&1\\1&0\end{pmatrix}\right)\begin{pmatrix}1\\2\end{pmatrix}$

Eg. R^3（行向量）中取 $\varepsilon_1=\overrightarrow{(1,0,0)},\varepsilon_2=\overrightarrow{(0,1,0)},\varepsilon_3=\overrightarrow{(0,0,1)}$

$\alpha_1=\overrightarrow{(1,1,2)},\alpha_2=\overrightarrow{(0,1,1)}\Rightarrow\alpha_1=1\bullet\varepsilon_1+1\bullet\varepsilon_2+2\bullet\varepsilon_3,\alpha_2=0\bullet\varepsilon_1+1\bullet\varepsilon_2+1\bullet\varepsilon_3$

$(\alpha_1,\alpha_2)=(\varepsilon_1,\varepsilon_2,\varepsilon_3)\begin{pmatrix}1&0\\1&1\\2&1\end{pmatrix}$ 或 $\left(\overrightarrow{(1,1,2)},\overrightarrow{(0,1,1)}\right)=\left(\overrightarrow{(1,0,0)},\overrightarrow{(0,1,0)},\overrightarrow{(0,0,1)}\right)\begin{pmatrix}1&0\\1&1\\2&1\end{pmatrix}$

改为“列向量” $\left(\begin{pmatrix}1\\1\\2\end{pmatrix},\begin{pmatrix}0\\1\\1\end{pmatrix}\right)=\left(\begin{pmatrix}1\\0\\0\end{pmatrix},\begin{pmatrix}0\\1\\0\end{pmatrix},\begin{pmatrix}0\\0\\1\end{pmatrix}\right)\begin{pmatrix}1&0\\1&1\\2&1\end{pmatrix}$

例如：线性组合必有广阵

(甲) = 3(红) + 7(白), (乙) = 4(红) + 6(白) $\Leftrightarrow$ (甲,乙) = (红,白)$\begin{pmatrix}3&4\\7&6\end{pmatrix}$

应用：线性变换矩阵公式：设 $\varphi: V\to W$ 为线性的$(\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n)$为基

则有公式 $\varphi(\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n)=(\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n)A$

其中 $A_{n\times n}$ 叫 φ 在基$(\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n)$下的矩阵（表示阵）

Pf：设：$\begin{cases}\varphi(\varepsilon_1)=\sum_{i=1}^{n}a_{i1}\varepsilon_i\\\varphi(\varepsilon_2)=\sum_{i=1}^{n}a_{i2}\varepsilon_i\\\vdots\\\varphi(\varepsilon_n)=\sum_{i=1}^{n}a_{in}\varepsilon_i\end{cases}$，即 $\varphi(\varepsilon_1),\varphi(\varepsilon_2),\ldots,\varphi(\varepsilon_n)$可由 $\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n$ 表示

由广阵原理$\Rightarrow(\varphi(\varepsilon_1),\varphi(\varepsilon_2),\cdots,\varphi(\varepsilon_n))=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)A_{n\times n}$（$A$中列就是组合系数）

注： 在给定基下，φ有唯一矩阵A（用消去律）$\Rightarrow\varphi\Leftrightarrow A$是一一对应的

V到V全体线性变换集合$L(V, V)$与方阵集合$R^{n\times n}$可等同有$L(V,V)\underline{\underline{(同构)}}R^{n\times n}$

同理：若$\varphi: V\to W$ ($\dim V=n, \dim W=m$)为线性的

且$(\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n)$为$V$中基，$(g_1,g_2,\ldots,g_n)$为$W$中基

$\varphi(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)=(g_1,g_2,\cdots,g_n)A_{m\times n}\Rightarrow\varphi\Leftrightarrow A_{m\times n}$为一一对应

广阵格式及应用

引理：（广阵原理）： 一切线性组合都有广阵格式。

若广元$\otimes_1,\otimes_2,\cdots,\otimes_p$可由$\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n$ “表示”

则有$(\otimes_1,\otimes_2,\cdots,\otimes_p)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)B_{n\times p}$

其中系数阵B中列就是原组合系数

线性映射的矩阵（表示阵）

设$\varphi: V\to W$线性变换，固定基$(\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n)$，$\dim V=n$

则$\varphi(\varepsilon_1),\varphi(\varepsilon_2),\ldots,\varphi(\varepsilon_n)$（在$V$中）可由$\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n$表示

有广阵格式$(\varphi(\varepsilon_1),\varphi(\varepsilon_2),\ldots,\varphi(\varepsilon_n))=(\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n)A_{n\times n}$

称$A=A_{n\times n}$为φ在固定基下的矩阵（表示阵）

注： 固定基：每个线性变换$\varphi: V\to W$对应一个唯一矩阵A

即$\varphi\leftrightarrow A$ 是一一对应（双射）

（利用消去法：$(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)A=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)B\Rightarrow A=B$）

消去前提：线性无关（基就是线性无关的）

推论： 全体线性变换空间$L(V,V)\leftrightarrow R^{n\times n}$（方阵空间）是同构

可写$L(V,V)\underline{\underline{同构}}R^{n\times n}$（实域上），$L(V, V)=C^{n\times n}$（复域上）

注： 固定基$\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n\in V$，$\forall\alpha\in V$，$\alpha=\sum_{i=1}^{n}a_i\varepsilon_i$

$\alpha=\left(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n\right)\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix}$，引入坐标同构 $\sigma: V\to R^n(C^n)$

使得 $\sigma(\alpha)=\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix}\in R^n$，$\alpha=(\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n)\,\sigma(\alpha)$

$$
\text{令}\begin{cases}
\varphi(\varepsilon_1)=\sum\limits_{i=1}^{n}a_{i1}\varepsilon_i=\left(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n\right)\begin{pmatrix}a_{11}\\a_{21}\\\vdots\\a_{n1}\end{pmatrix}=\left(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n\right)\sigma(\varepsilon_1)\\
\varphi(\varepsilon_2)=\sum\limits_{i=1}^{n}a_{i2}\varepsilon_i=\left(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n\right)\begin{pmatrix}a_{12}\\a_{22}\\\vdots\\a_{n2}\end{pmatrix}=\left(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n\right)\sigma(\varepsilon_2)\\
\qquad\qquad\qquad\qquad\vdots\\
\varphi(\varepsilon_n)=\sum\limits_{i=1}^{n}a_{in}\varepsilon_i=\left(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n\right)\begin{pmatrix}a_{1n}\\a_{2n}\\\vdots\\a_{nn}\end{pmatrix}=\left(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n\right)\sigma(\varepsilon_n)
\end{cases}
$$

$\Rightarrow\left(\varphi(\varepsilon_1),\varphi(\varepsilon_2),\cdots,\varphi(\varepsilon_n)\right)=\left(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n\right)\left(\sigma(\varepsilon_1),\sigma(\varepsilon_2),\cdots,\sigma(\varepsilon_n)\right)_{n\times n}$

令 $\sigma: V\to R^n$ 为坐标同构映射，则 $V\underset{=}{\sigma}R^n$， $L(V,V)\underset{=}{\sigma}R^{n\times n}$

推广： 设 $\varphi: V\to W$ 线性映射，记为 $\varphi\in L(V, W)$（全体线性映射）

固定基 $\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n\in V$，再固定基 $g_1, g_2, \ldots, g_m\in W$，（$\dim V=n$, $\dim W=m$）

$\varphi(\varepsilon_1), \varphi(\varepsilon_2), \ldots, \varphi(\varepsilon_n)$（在 W 中）可由 $g_1, g_2, \ldots, g_m$ 表示

用广阵格式$(\varphi(\varepsilon_1), \varphi(\varepsilon_2), \ldots, \varphi(\varepsilon_n))=(g_1, g_2, \ldots, g_m)A_{m\times n}$

称 $A=A_{m\times n}$ 为 φ 在固定基下的矩阵（表示阵）

可知： 每个 $\varphi\in L(V, W)$对应唯一的矩阵 $A=A_{m\times n}$

推论：（全体线性映射）$L(V, W)$在固定基下与 $R^{m\times n}$ 或 $C^{m\times n}$ 同构

可写 $L(V,W)\underset{=}{\sigma}R^{m\times n}$ 或 $L(V,W)\underset{=}{\sigma}C^{m\times n}$

规定： V^n 表示 n 维空间，W^m 表示 m 维空间

（全体线性映射）$L\left(V^n, W^m\right)\underline{\underline{\text{同构}}}R^{m\times n}$或$C^{m\times n}$

坐标公式： 固定基$(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n) \in V$，$(g_1, g_2, \ldots, g_m) \in W$

$\forall\alpha = \sum_{i=1}^{n} a_i\varepsilon_i$，$\varphi(\alpha) = \sum_{i=1}^{m} b_i g_i$

取坐标：$\alpha \leftrightarrow x = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \in R^n$，$\varphi(\alpha) \leftrightarrow y = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} \in R^m$，则 $y = A_{m\times n}$，即$\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} = A_{m\times n}$

Pf：$\alpha = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)x \Rightarrow \varphi(\alpha) = (\varphi(\varepsilon_1), \varphi(\varepsilon_2), \cdots, \varphi(\varepsilon_n))x = (g_1, g_2, \cdots, g_m)A_{m\times n}x$

又写$\varphi(\alpha) = (g_1, g_2, \cdots, g_m)y \Rightarrow (g_1, g_2, \cdots, g_m)Ax = (g_1, g_2, \cdots, g_m)y \Rightarrow Ax = y$

结论： 设 $\varphi\in L(V^n, W^m)$（固定 2 个基），则$(\varphi(\varepsilon_1), \varphi(\varepsilon_2), \ldots, \varphi(\varepsilon_n)) = (g_1, g_2, \ldots, g_m)A_{m\times n}$

$\forall\alpha \in V$，$\alpha = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)x$，$\varphi(\alpha) = (g_1, g_2, \ldots, g_m)y$

则$\varphi \leftrightarrow A$互相对应，$\alpha \to \varphi(\alpha)$可用$x \to Ax$代替

即：若$\beta = \varphi(\alpha)$则可写$y = Ax$

注： $A: R^n \to R^m$是线性映射，可代替 $\varphi: V \to W$

Eg.零映射 $\theta: V^n \to W^m$，固定基$(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)$和$(g_1, g_2, \ldots, g_m)$

$\forall\alpha \in V$，$\theta(\alpha) = O\in W$

$$\begin{cases} \theta(\varepsilon_1) = 0 = 0\bullet g_1 + 0\bullet g_2 + \cdots + 0\bullet g_m \\ \theta(\varepsilon_2) = 0 = 0\bullet g_1 + 0\bullet g_2 + \cdots + 0\bullet g_m \\ \qquad\vdots \\ \theta(\varepsilon_n) = 0 = 0\bullet g_1 + 0\bullet g_2 + \cdots + 0\bullet g_m \end{cases} \Rightarrow (\theta(\varepsilon_1), \theta(\varepsilon_2), \cdots, \theta(\varepsilon_n)) = (g_1, g_2, \cdots, g_m)O_{m\times n}$$

$\theta \leftrightarrow O_{m\times n} \in R^{m\times n}$

Eg.恒同映射：$I_V: V \to V$，$\forall\alpha \in V$，$I_V(\alpha) = \alpha$

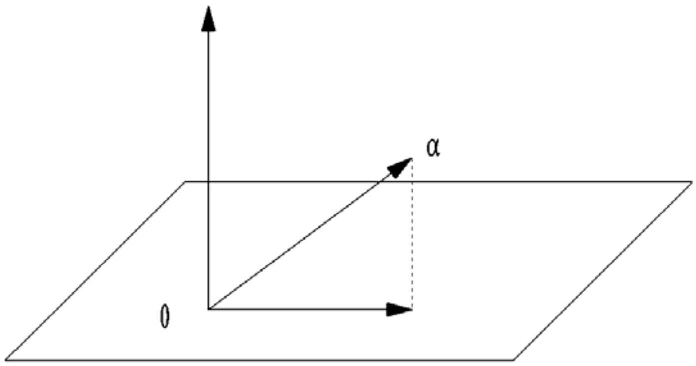
$$\begin{cases} I_V(\varepsilon_1) = \varepsilon_1 = 1\bullet\varepsilon_1 + 0\bullet\varepsilon_2 + \cdots + 0\bullet\varepsilon_n \\ I_V(\varepsilon_2) = \varepsilon_2 = 0\bullet\varepsilon_1 + 1\bullet\varepsilon_2 + \cdots + 0\bullet\varepsilon_n \\ \qquad\vdots \\ I_V(\varepsilon_n) = \varepsilon_n = 0\bullet\varepsilon_1 + 0\bullet\varepsilon_2 + \cdots + 1\bullet\varepsilon_n \end{cases} \Rightarrow I_V(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n) = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)I_n$$

$I_V \leftrightarrow I_n$（单位阵）

设 $\alpha = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)x$，则 $\varphi(\alpha)$坐标 $y = I_n x = x$

Eg.设 $V=\mathrm{span}(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)$，$\varphi: V\to V$，$\alpha=\sum_{i=1}^{n} a_i\varepsilon_i \in V$

使得 $\varphi\left(\sum_{i=1}^{n} x_i\varepsilon_i\right)=x_1\varepsilon_1+x_2\varepsilon_2$（投影）



$$\because \begin{cases}\varphi(\varepsilon_1)=\varepsilon_1=1\bullet\varepsilon_1+0\bullet\varepsilon_2+\cdots+0\bullet\varepsilon_n\\ \varphi(\varepsilon_2)=\varepsilon_2=0\bullet\varepsilon_1+1\bullet\varepsilon_2+\cdots+0\bullet\varepsilon_n\\ \varphi(\varepsilon_3)=0=0\bullet\varepsilon_1+0\bullet\varepsilon_2+\cdots+0\bullet\varepsilon_n\\ \qquad\vdots\\ \varphi(\varepsilon_n)=0=0\bullet\varepsilon_1+0\bullet\varepsilon_2+\cdots+0\bullet\varepsilon_n\end{cases} \Rightarrow \varphi(\varepsilon_1,\varepsilon_2)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)A,$$

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

坐标公式：$y=Ax$，$x\in R^n$

Eg. V 基$(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)$，取 $W=\mathrm{span}(\varepsilon_1, \varepsilon_2)$，$\varphi: V\to W$，$\varphi\in L(V, W)$

使得 $\varphi(\alpha)=\varphi\left(\sum_{i=1}^{n} x_i\varepsilon_i\right)=x_1\varepsilon_1+x_2\varepsilon_2\in W$ （投影）

$$\begin{cases}\varphi(\varepsilon_1)=\varepsilon_1=1\bullet\varepsilon_1+0\bullet\varepsilon_2\\ \varphi(\varepsilon_2)=\varepsilon_2=0\bullet\varepsilon_1+1\bullet\varepsilon_2\\ \varphi(\varepsilon_3)=0=0\bullet\varepsilon_1+0\bullet\varepsilon_2\\ \qquad\vdots\\ \varphi(\varepsilon_n)=0=0\bullet\varepsilon_1+0\bullet\varepsilon_2\end{cases} \Rightarrow \varphi(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)=(\varepsilon_1,\varepsilon_2,)\begin{pmatrix}1&0&0&\cdots&0\\0&1&0&\cdots&0\end{pmatrix}_{2\times n}$$

Ex.令 $V_n(x)=span\left(1,x,\cdots,x^{n-1}\right)=\left\{f=a_0+a_1x+\cdots+a_{n-1}x^{n-1}\middle| a_i\in R\right\}$

（全体次数小于 n 的多项式空间）

（1）令 $\varphi=\frac{d}{dx}$：$V\to V$（求导），求 φ 在基$(1, x, \ldots, x^{n-1})$的矩阵 A

（2）令 $\varphi = \frac{d}{dx}: V_n(x) \to V_{n-1}(x)$ （求导），求 φ 在基$(1, x, \dots, x^{n-1})$与基$(1, x, \dots, x^{n-2})$下的矩阵。

换基公式：设 V 中 2 个基$(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)$与$(g_1, g_2, \dots, g_n)$

则它们互换表示（由广阵格式）可写：$(g_1, g_2, \dots, g_n) = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)P$，$(P = P_{n\times n})$

$(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n) = (g_1, g_2, \dots, g_n)Q$，$(Q = Q_{n\times n})$

则 P 可逆，且 $Q = P^{-1}$，$(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n) = (g_1, g_2, \dots, g_n)\, P^{-1}$

Pf：$\because (g_1, g_2, \cdots, g_n) = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)P \underline{\underline{\text{代入}}} (g_1, g_2, \cdots, g_n)QP$，（消去）

$\therefore I_n = QP \Rightarrow Q = P^{-1}$

称 P 是(ε)到(g)的过度阵

规定记号：$(\varepsilon) = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)$，$(g) = (g_1, g_2, \dots, g_n)$。（2 个坐标系）

换基公式：$(g) = (\varepsilon)P$，$(\varepsilon) = (g)\, P^{-1}$

换坐标公式：若 $\alpha = \sum_{i=1}^{n} x_i \varepsilon_i = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)x = (\varepsilon)x$

且 $\alpha = \sum_{i=1}^{n} y_i g_i = (g_1, g_2, \cdots, g_n)y = (g)y$

则有坐标公式 $x = Py$ 或 $y = P^{-1}x$，$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in R^n$

Pf：$\alpha = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \underline{\underline{\text{写}}} (\varepsilon)x$，且 $\alpha = (g_1, g_2, \cdots, g_n)\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = (g)y$

$\Rightarrow \alpha = (g)y \underline{\underline{(g) = (\varepsilon)P}} (\varepsilon)Py = (\varepsilon)x = \varepsilon$ （消去）$\Rightarrow Py = x$

记号：$(\varepsilon) = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)$，$(g) = (g_1, g_2, \dots, g_n)$，$\varphi(\varepsilon) = (\varphi(\varepsilon_1), \varphi(\varepsilon_2), \dots, \varphi(\varepsilon_n))$

换基相似定理：设 $\varphi: V \to V$ 为线性变换，固定 2 个基

$(\varepsilon) = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)$，$(g) = (g_1, g_2, \dots, g_n)$

令 $\varphi(\varepsilon) = (\varepsilon)A$（表示公式），$\varphi(g) = (g)B$（第 2 表示）

$(g) = (\varepsilon)P$ 或$(\varepsilon) = (g)\ P^{-1}$（**换基公式**）

则：$B = P^{-1}AP$ 相似

Pf：$\varphi(g) = (g)B$，且 $\varphi(g) = \varphi((\varepsilon)P) = \varphi(\varepsilon)P = (\varepsilon)AP = (g)P^{-1}AP$

$(g)B = (g)P^{-1}AP \Rightarrow B = P^{-1}AP$

Ex. 《矩阵分析（史荣昌等）》P68 9 10 12 19

坐标与广阵格式应用

广阵原理：若广元$\otimes_1,\otimes_2,\cdots,\otimes_p$可由 $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ “表示”，

则由公式$(\otimes_1,\otimes_2,\cdots,\otimes_p) = (\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)B_{n\times p}$

Eg. $V = \mathrm{span}(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)$，$W = \mathrm{span}(g_1, g_2)$

固定基$\{\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n\}$，$\{g_1, g_2\}$

令线性映射 $\varphi: V \to W$ $\quad \forall \alpha = \sum_{i=1}^{n} x_i\varepsilon_i \in V$

使得$\varphi(\alpha) = \varphi\left(\sum_{i=1}^{n} x_i\varepsilon_i\right) = x_1g_1 + x_2g_2$

$$\begin{cases} \varphi(\varepsilon_1) = g_1 = 1\bullet g_1 + 0\bullet g_2 \\ \varphi(\varepsilon_2) = g_2 = 0\bullet g_1 + 1\bullet g_2 \\ \varphi(\varepsilon_3) = 0 = 0\bullet g_1 + 0\bullet g_2 \\ \qquad\vdots \\ \varphi(\varepsilon_n) = 0 = 0\bullet g_1 + 0\bullet g_2 \end{cases}$$

$$(\varphi(\varepsilon_1),\varphi(\varepsilon_2),\cdots,\varphi(\varepsilon_n)) = (g_1, g_2,)\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \end{pmatrix}_{2\times n}$$

简写$\varphi(\varepsilon) = (g)\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \end{pmatrix}$

$(\varepsilon) = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)$，$(g) = (g_1, g_2)$，$\varphi(\varepsilon) = (\varphi(\varepsilon)) = (\varphi(\varepsilon_1), \varphi(\varepsilon_2), \ldots, \varphi(\varepsilon_n))$

令 $\varphi: V^n \to W^m$ 为线性的，固定基$(\varepsilon) = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)$，$(g) = (g_1, g_2, \ldots, g_m)$

则 $\varphi(\varepsilon) = (g)A_{m\times n}$，$(\varphi(\varepsilon_1), \varphi(\varepsilon_2), \ldots, \varphi(\varepsilon_n)) = (g_1, g_2, \ldots, g_m)\ A_{m\times n}$

$A_{m\times n}$ 叫 φ 的表示阵（在固定基下）

换基相似公式：设 $\varphi: V \to V$ 为线性的或 $\varphi \in L(V, V)$

固定基$(\varepsilon) = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)$，$(g) = (g_1, g_2, \ldots, g_n)$

记：$\varphi(\varepsilon)=(\varepsilon)A_{n\times n}$ 或$(\varphi(\varepsilon_1),\varphi(\varepsilon_2),\dots,\varphi(\varepsilon_n))=(\varepsilon_1,\varepsilon_2,\dots,\varepsilon_n)\,A_{n\times n}$

$\varphi(g)=(g)B_{n\times n}$ 或$(\varphi(g_1),\varphi(g_2),\dots,\varphi(g_n))=(g_1,g_2,\dots,g_n)\,B$

$(g)=(\varepsilon)P$（换基公式）或$(\varepsilon)=(g)P^{-1}$

则：$B=P^{-1}AP$（相似）

推论：（1）线性变换 φ: $V\to V$ 在不同基下的矩阵是相似关系

（2）在复数域上可取一个基$(g)=(g_1,g_2,\dots,g_n)$，使 φ 在该基下的矩阵 B 是 Jordan 形，即 $B=P^{-1}AP=\begin{pmatrix}J_1&&&\\&J_2&&\\&&\ddots&\\&&&J_s\end{pmatrix}$（Jordan 形）

小结：固定基下常用下列“替换”（替身）

（1）V^n 用 R^n 或 C^n 代替

（2）广元 $\alpha\in V^n$ 可用坐标 $x=\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}$ 代替（$\alpha=\sum_{i=1}^{n}x_i\varepsilon_i$）

（3）线性变换 φ: $V\to V$ 用矩阵代替

（4）相元 $\varphi(\alpha)$用 Ax 代替

固定基(ε)下$\begin{cases}(\alpha)\leftrightarrow(x)\in R^n\\ \varphi(\alpha)\leftrightarrow Ax\\ \varphi\leftrightarrow A\\ V^n\underline{\underline{\text{同构}}}R^n\text{或}C^n\\ L(V,V)=R^{n\times n}\end{cases}$

Eg.设 $V=R(x)_n=\{f=a_0+a_1x_1+\cdots+a_{n-1}x_{n-1}\,|\,a_i\in R\}$（全体次数小于 n 的多项式）

$\dim V=n$，$\{1,x,\dots,x^{n-1}\}$是一个基

另 $b_1,b_2,\dots,b_n$ 为互不相同的数

$$\begin{aligned}f_1(x)&=(x-\hat{b}_1)(x-b_2)\cdots(x-b_n)\\f_2(x)&=(x-b_1)(x-\hat{b}_2)\cdots(x-b_n)\\&\vdots\\f_n(x)&=(x-b_1)(x-b_2)\cdots(x-\hat{b}_n)\end{aligned}$$

“∧”表示删掉一项

$$f_j(x)=(x-b_1)\cdots(x-\hat{b}_j)\cdots(x-b_n)$$

$$g_1(x)=\frac{f_1(x)}{f_1(b_1)}, g_2(x)=\frac{f_2(x)}{f_2(b_2)}, \cdots, g_n(x)=\frac{f_n(x)}{f_n(b_n)} \in V$$

$$\text{取值}\begin{cases} g_1(b_1)=\dfrac{f_1(b_1)}{f_1(b_1)}=1, g_1(b_2)=0, \cdots, g_1(b_n)=0 \\ g_2(b_1)=0, g_2(b_2)=1, \cdots, g_2(b_n)=0 \\ \vdots \\ g_n(b_1)=0, g_n(b_2)=0, \cdots, g_n(b_n)=1 \end{cases}$$

证明：（1）$g_1, g_2, \ldots, g_n$ 是 V 的基

（2）求$(g_1, g_2, \ldots, g_n)$到$(1, x, \ldots, x^{n-1})$的过度阵 P

Pf：引入映射 $\varphi: V \to R^n$ $\forall f \in V$

$$\varphi(f) \triangleq \begin{pmatrix} f(b_1) \\ f(b_2) \\ \vdots \\ f(b_n) \end{pmatrix}，\varphi \text{ 是线性的：} \begin{matrix} \varphi(f+g)=\varphi(f)+\varphi(g) \\ \varphi(kf)=k\varphi(f) \end{matrix}$$

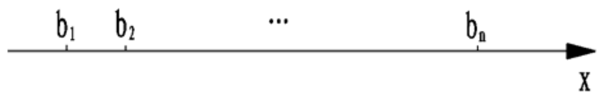
$$\Rightarrow \varphi(g_1)=\begin{pmatrix} g_1(b_1) \\ g_1(b_2) \\ \vdots \\ g_1(b_n) \end{pmatrix}=\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}=e_1, \varphi(g_2)=\begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}=e_2, \cdots, \varphi(g_n)=\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}=e_n \in R^n$$

$\Rightarrow \{\varphi(g_1)=e_1, \varphi(g_2)=e_2, \cdots, \varphi(g_n)=e_n\}$为无关组$\{ g_1, g_2, \ldots, g_n \}$也无关（是基）

设换基公式$(1, x, \ldots, x^{n-1}) = (g_1, g_2, \ldots, g_n)P$

$$\Rightarrow (\varphi(1), \varphi(x), \cdots, \varphi(x^{n-1})) = (\varphi(g_1), \varphi(g_2), \cdots, \varphi(g_n))P$$

$$\begin{pmatrix} 1 & b_1 & b_1^2 & \cdots & b_1^{n-1} \\ 1 & b_2 & b_2^2 & \cdots & b_2^{n-1} \\ 1 & b_3 & b_3^2 & \cdots & b_3^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & b_n & b_n^2 & \cdots & b_n^{n-1} \end{pmatrix} = I_n P \Rightarrow P = \begin{pmatrix} 1 & b_1 & b_1^2 & \cdots & b_1^{n-1} \\ 1 & b_2 & b_2^2 & \cdots & b_2^{n-1} \\ 1 & b_3 & b_3^2 & \cdots & b_3^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & b_n & b_n^2 & \cdots & b_n^{n-1} \end{pmatrix}$$



注： 固定 n 个不同点 $b_1, b_2, \ldots, b_n$；

$$\text{规定“取值映射”}\varphi(f)=\begin{pmatrix} f(b_1) \\ f(b_2) \\ \vdots \\ f(b_n) \end{pmatrix} \in R^n，\varphi: V \to R^n \text{ 为线性}$$

令 $\varphi(h)=\begin{pmatrix} h(b_1) \\ h(b_2) \\ \vdots \\ h(b_n) \end{pmatrix} \Rightarrow \varphi(f+h)=\begin{pmatrix} f(b_1)+h(b_1) \\ f(b_2)+h(b_2) \\ \vdots \\ f(b_n)+h(b_n) \end{pmatrix}=\begin{pmatrix} f(b_1) \\ f(b_2) \\ \vdots \\ f(b_n) \end{pmatrix}+\begin{pmatrix} h(b_1) \\ h(b_2) \\ \vdots \\ h(b_n) \end{pmatrix}$

f ≡ 1 时，$\varphi(f)=\varphi(1)\begin{pmatrix} f(b_1) \\ f(b_2) \\ \vdots \\ f(b_n) \end{pmatrix}=\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$

实际上：$\varphi\colon V=R(x)\to R^n$ 是同构，$R(x)_n \underline{\underline{\text{同构}\varphi}} R^n$

0 **点引理**：固定 $b_1, b_2, \dots, b_n$（互异）；$g_1(x), g_2(x), \dots, g_n(x)$同上

则：（1）$1=\sum_{i=1}^{n} g_i(x)$；（2）$x=\sum_{i=1}^{n} b_i g_i(x)$；

（3）$g_i(x)\, g_j(x)$含有因子$(x-b_1)(x-b_2)\dots(x-b_n)$ $(i\neq j)$

Pf：∵$(1, x, \dots, x^{n-1})=(g_1, g_2, \dots, g_n)P$；$P=\begin{pmatrix} 1 & b_1 & b_1^2 & \cdots & b_1^{n-1} \\ 1 & b_2 & b_2^2 & \cdots & b_2^{n-1} \\ 1 & b_3 & b_3^2 & \cdots & b_3^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & b_n & b_n^2 & \cdots & b_n^{n-1} \end{pmatrix}$

$$\Rightarrow \begin{cases} 1=\sum_{i=1}^{n} g_i(x) \\ x=\sum_{i=1}^{n} b_i g_i(x) \\ \vdots \\ x^{n-1}=\sum_{i=1}^{n} b_i^{n-1} g_i(x) \end{cases}\text{，有 } x^k=\sum_{i=1}^{n} b_i^k g_i(x)$$

（3）例如

$$\left.\begin{aligned} g_1(x)&=\frac{f_1(x)}{f_1(b_1)}=\frac{(x\hat{-}b_1)(x-b_2)\cdots(x-b_n)}{f_1(b_1)} \\ g_2(x)&=\frac{f_2(x)}{f_2(b_2)}=\frac{(x-b_1)(x\hat{-}b_2)\cdots(x-b_n)}{f_2(b_2)} \end{aligned}\right. \Rightarrow g_1(x)g_2(x)=(x-b_1)(x-b_2)\cdots(x-b_n)(\cdots)$$

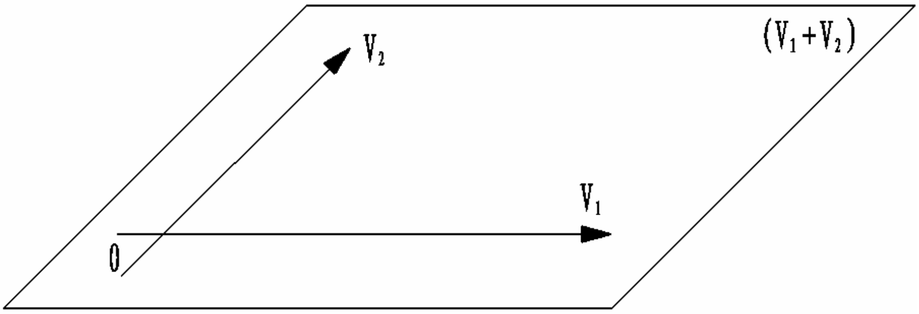
推论：固定 $b_1, b_2, \dots, b_n$ 与 $g_1(x), g_2(x), \dots, g_n(x)$任取方阵 $A=A_{p\times p}$

有（1）$I=\sum_{i=1}^{n} g_i(A)$；（2）$A=\sum_{i=1}^{n} b_i g_i(A)$

和空间定义：设 V_1，V_2 是子空间

称 $V_1+V_2 \underline{\underline{\Delta}} \{全体\alpha_1+\alpha_2 | \alpha_1 \in V_1, \alpha_2 \in V_2\}$ 为 V_1，V_2 的和（可知 V_1+V_2 是子空间）

注：并集 $V_1 \cup V_2$ 一般不是子空间，$V_1 \cup V_2 \subset V_1+V_2$



同理 V_1，V_2，V_3 为子空间，可定义 $V_1+V_2+V_3=\{全体\alpha_1+\alpha_2+\alpha_3 | \alpha_i \in V_i\}$

维数公式：$\dim(V_1+V_2)=\dim V_1+\dim V_2-\dim(V_1 \cap V_2)$

或 $\operatorname{rank}(V_1+V_2)=\operatorname{rank} V_1+\operatorname{rank} V_2-\operatorname{rank}(V_1 \cap V_2)$

直和定义：设 V_1，V_2 为子空间，且 0 元具有唯一分解性

即：$0=\alpha_1+\alpha_2$ $(\alpha_1 \in V_1$，$\alpha_2 \in V_2)$ 必有 $\alpha_1=0$，$\alpha_2=0$

称 V_1+V_2 为直和，记为 $V_1 \oplus V_2$

定理：V_1+V_2 为直和 $V_1 \oplus V_2 \Leftrightarrow V_1 \cap V_2=\{0\}$

同理：3 个子空间 V_1，V_2，V_3

若 0 元具有唯一分解：$0=\alpha_1+\alpha_2+\alpha_3 \Rightarrow \alpha_1=0, \alpha_2=0, \alpha_3=0$ $(\alpha_i \in V_i)$

则称 $V_1+V_2+V_3$ 为直和，记为 $V_1 \oplus V_2 \oplus V_3$

直和维数公式：$\dim(V_1 \oplus V)_2 \Leftrightarrow \dim V_1+\dim V_2$

$\dim(V_1 \oplus V_2 \oplus V_3)=\dim V_1+\dim V_2+\dim V_3$

补空间：若 $V_1+V_2=V$（全空间）且 $V_1 \cap V_2=\{0\}$，即 $V_1 \oplus V_2=V$

称 V_2 为 V_1 的补空间

注：V_1 的补空间可能很多

生成元公式：设 $V_1=\operatorname{span}(\alpha_1, \alpha_2, \ldots, \alpha_s)$，$V_2=\operatorname{span}(\beta_1, \beta_2, \ldots, \beta_t)$

则 $V_1+V_2=\operatorname{span}(\alpha_1, \alpha_2, \ldots, \alpha_s, \beta_1, \beta_2, \ldots, \beta_t)$

注：$\{\alpha_1, \alpha_2, \dots, \alpha_s, \beta_1, \beta_2, \dots, \beta_t\}$未必无关

同构方法（替身法）

先固定基$(\varepsilon) = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)$，空间为 $V = \mathrm{span}(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)$

固定基$(g) = (g_1, g_2, \dots, g_m)$，空间为 $W = \mathrm{span}(g_1, g_2, \dots, g_m)$

可用下列代替法：

（1）$\alpha \in V$ 写成坐标 $x = (x_1, x_2, \dots, x_n)^T \in R^n$ 或 C^n（$\because V \underline{\underline{同构}} R^n$ 或 C^n）

$\beta \in W$ 写成坐标 $y = (y_1, y_2, \dots, y_m)^T \in R^m$ 或 C^m（$\because W \underline{\underline{同构}} R^m$ 或 C^m）

（2）线性变换：$\varphi \in L(V, V)$写成方阵 $A_{n\times n}$ 且有表示公式：$\varphi(\varepsilon) = (\varepsilon) A_{n\times n}$

公式：$\varphi(\varepsilon) = \lambda\alpha$ 写成 $Ax = \lambda x$

（3）$\varphi(\alpha)$写成 Ax

（4）映射 $\varphi \in L(V, W)$写成矩阵 $A_{m\times n} = A$，有表示公式：$\varphi(\varepsilon) = (g) A_{m\times n}$

（5）$\varphi(\alpha)$写成 $A_{m\times n}x$，公式 $\varphi(\alpha) = \beta$ 写成 $A_{m\times n}x = y$（坐标公式）

注： 若 $\alpha_1, \alpha_2, \dots, \alpha_s$ 无关，则坐标 $X_1, X_2, \dots, X_s$ 也无关

一般秩 $\mathrm{rank}(\alpha_1, \alpha_2, \dots, \alpha_s) = \mathrm{rank}(X_1, X_2, \dots, X_s)$

Ex.取 $\alpha_1 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \alpha_2 = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}, \alpha_3 = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}, \alpha_4 = \begin{pmatrix} 1 & 3 \\ 1 & 0 \end{pmatrix}$ 为 $R^{2\times 2}$ 中基，且 φ 是线性的

$$\varphi(\alpha_1) = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \varphi(\alpha_2) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \varphi(\alpha_3) = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}, \varphi(\alpha_4) = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}$$

求 φ 的表示矩阵（公式）

解：利用“拉直同构”可写

$\alpha_1 = (1, 0, 1, 1)^T, \alpha_2 = (0, 1, 1, 1)^T, \alpha_3 = (1, 1, 0, 2)^T, \alpha_4 = (1, 3, 1, 0)^T \in R^4$

$\varphi(\alpha_1) = (1, 1, 0, 0)^T, \varphi(\alpha_2) = (0, 0, 0, 0)^T, \varphi(\alpha_3) = (0, 0, 1, 1)^T, \varphi(\alpha_4) = (0, 1, 0, 1)^T \in R^4$

设表示公式：$\varphi(\alpha_1, \alpha_2, \alpha_3, \alpha_4) = (\alpha_1, \alpha_2, \alpha_3, \alpha_4)A$

$(\varphi(\alpha_1), \varphi(\alpha_2), \varphi(\alpha_3), \varphi(\alpha_4)) = (\alpha_1, \alpha_2, \alpha_3, \alpha_4)A$

$$A = (\alpha_1, \alpha_2, \alpha_3, \alpha_4)^{-1}(\varphi(\alpha_1), \varphi(\alpha_2), \varphi(\alpha_3), \varphi(\alpha_4)) = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 3 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 2 & 0 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

映射：$\varphi \in L(V, W)$的相空间（值域）与核

规定：相空间为 $\mathscr{R}(\varphi)=\varphi(V)=\{全体\varphi(\alpha)|\alpha\in V\}\subset W$

核空间为 $\mathscr{N}(\varphi)=\varphi^{-1}(0)=\{全体\alpha|\varphi(\alpha)=0\}\subset V$

（1）相空间的秩为：$\mathrm{rank}(\varphi)=\dim\mathscr{R}(\varphi)=\mathrm{rank}\,\mathscr{R}(\varphi)$，也叫映射 φ 的秩数

（2）核空间的维数（秩数）：$\mathrm{rank}\,\mathscr{N}(\varphi)=\dim\mathscr{N}(\varphi)$也叫 φ 的 0 度

0 度公式：$\dim\mathscr{N}(\varphi)+\dim\mathscr{R}(\varphi)=n \qquad \varphi\in L(V^n, W^m)$

或 $\mathrm{rank}(\varphi^{-1}(0))+\mathrm{rank}(\varphi)=n$

注： φ 写成 $A=A_{m\times n}$

$\varphi^{-1}(0)=\{\alpha|\varphi(\alpha)=0\}$写成 $A^{-1}(0)=\{x|Ax=0\}$（解空间）

相空间 $\mathscr{R}(\varphi)=\{\varphi(\alpha)|\alpha\in V\}$写成 $\mathscr{R}(A)=\{Ax|x\in R^n\}$

规定： $A=A_{m\times n}$的列空间为 $\mathscr{R}(A)=\{Ax|x\in R^n\}\subset R^m$

$A=A_{m\times n}$的核空间 $\mathscr{N}(A)=A^{-1}(0)=\{x|Ax=0\}\subset R^n$（解空间）

改写 $A_{m\times n}=(\alpha_1, \alpha_2, \ldots, \alpha_n)$，$A_{m\times n}x=\sum_{i=1}^{n}x_i\alpha_i$，$x=\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}$

$\Rightarrow\mathscr{R}(A)=\{Ax\}=\left\{全体\sum_{i=1}^{n}x_i\alpha_i\right\}=span(\alpha_1,\alpha_2,\cdots,\alpha_n)$（由 $\alpha_1, \alpha_2, \ldots, \alpha_n$ 生成）

$\Rightarrow\dim\mathscr{R}(A)=\mathrm{rank}\,\mathscr{R}(A)=\mathrm{rank}\,(\alpha_1, \alpha_2, \ldots, \alpha_n)=\mathrm{rank}(A)$

由公式 $rankA^{-1}(0)+rank(A)=n\Rightarrow rank\varphi^{-1}(0)+rank(\varphi)=n$

$A_{m\times n}x=0$ 的基础解有$(n-r)$个 $\xi_1, \xi_2, \ldots, \xi_{n-1}$，$r=\mathrm{rank}(A)$，通解：$x=\sum_{i=1}^{n-r}c_i\xi_i$

$\Rightarrow$核 $\mathscr{N}(A)=A^{-1}(0)=\{x|Ax=0\}=span(\xi_1,\xi_2,\cdots,\xi_{n-1})$（解空间）

$\Rightarrow\dim\mathscr{N}(A)=\mathrm{rank}\,\mathscr{N}(A)=\mathrm{rank}(\xi_1, \xi_2, \ldots, \xi_{n-1})=n-r$

$\Rightarrow rank\,\mathscr{N}(A)=n-rank(A)\Leftrightarrow rank\ \mathscr{N}(A)+\mathrm{rank}(A)=0$

引理： $\varphi\in L(V^n, W^m)$写成 $A=A_{m\times n}$，$R^n\to R^m$

则：（1）$\mathrm{rank}\,(\varphi)=\mathrm{rank}(A)=\mathrm{rank}\,\mathscr{R}(A)$（列空间维数）

（2）$\mathrm{rank}(\varphi^{-1}(0))=\mathrm{rank}(A^{-1}(0))$或 $\mathrm{rank}\,\mathscr{N}(\varphi)=\mathrm{rank}\,\mathscr{N}(A)$

φ 与 A 的不变子空间：

设 $\varphi \in L(V, V)$固定基下，可写 $A \in L(R^n \to R^m)$

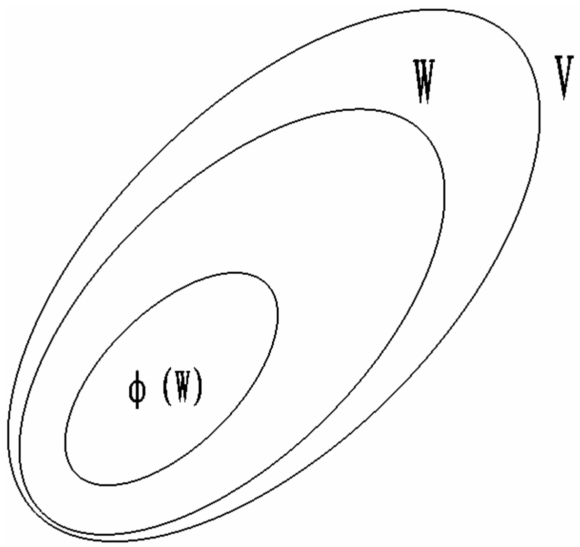
若子空间 $W \subset V$ 使得 $\forall \alpha \in W$，$\varphi(\alpha) \in W$

即 $\varphi(W) \subset W$，称 W 是 φ 的不变子空间

平凡不变子空间$\{0\}$与 V 都是 φ 的不变子空间

取特征子空间 $V(\lambda) = \{\alpha | \varphi(\alpha) = \lambda\alpha\} = \{\alpha | (\varphi - \lambda I)\alpha = 0\}$（$\lambda$ 的特征向量含 $\vec{0}$）

$V(\lambda)$是 φ 的不变子空间，若 $\alpha \in V(\alpha)$，验证：$\varphi(\alpha) \in V(\lambda)$



$A = A_{n\times n}$ 的不变子空间 $W \subset R^n$（或 C^n）

使得 $A(W) \subset W$，即任何 $x \in W$，$Ax \in W$

特征子空间 $V(\lambda) = \{x | Ax = \lambda x\} = \{x | (A - \lambda I)x = 0\}$是 A 的不变子空间

引理： 若 $W = \mathrm{span}(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_r)$是 $V = \mathrm{span}(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)$中子空间

$(\varepsilon) = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_r)$为基，$\varphi \in L(V, V)$

设 W 是 φ 的不变子空间，则有表示阵 $A = \begin{pmatrix} A_{r\times r} & (*) \\ 0 & (*) \end{pmatrix}_{n\times n}$

Pf：由定义：$\begin{cases} \varphi(\varepsilon_1)=(*)\varepsilon_1+(*)\varepsilon_2+\cdots+(*)\varepsilon_r+0\bullet\varepsilon_{r+1}+\cdots+0\bullet\varepsilon_n \\ \varphi(\varepsilon_2)=(*)\varepsilon_1+(*)\varepsilon_2+\cdots+(*)\varepsilon_r+0\bullet\varepsilon_{r+1}+\cdots+0\bullet\varepsilon_n \\ \vdots \\ \varphi(\varepsilon_r)=(*)\varepsilon_1+(*)\varepsilon_2+\cdots+(*)\varepsilon_r+0\bullet\varepsilon_{r+1}+\cdots+0\bullet\varepsilon_n \\ \varphi(\varepsilon_{r+1})=(*)\varepsilon_1+(*)\varepsilon_2+\cdots+(*)\varepsilon_n \\ \vdots \\ \varphi(\varepsilon_n)=(*)\varepsilon_1+(*)\varepsilon_2+\cdots+(*)\varepsilon_n \end{cases}$

$$\Rightarrow(\varphi(\varepsilon_1),\varphi(\varepsilon_2),\cdots,\varphi(\varepsilon_n))=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_r)\begin{pmatrix}(A_{r\times r}) & (*)\\ 0 & (*)\end{pmatrix}$$

定理：若 $\varphi\in L(V,V)$有 2 个不变子空间$W_1,W_2\subset V$，且 $W_1\oplus W_2=V$（直和）

可设 $W_1=\mathrm{span}(\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_r)$，$W_2=\mathrm{span}(\varepsilon_{r+1},\varepsilon_{r+2},\ldots,\varepsilon_n)$

则 φ 的矩阵为 $A=\begin{pmatrix}A_1 & 0\\ 0 & A_2\end{pmatrix}\underline{\text{记为}}A=A_1\oplus A_2$

注：若 W 是 φ 的不变子空间，限制映射$\varphi|_W$：$W\to W$ 是 W 到 W 的线性变换

引理：（1）复数域上方阵 $A=A_{n\times n}$ 必有特征值与特征向量，使得 $Ax=\lambda x$（$x\neq\vec{0}$）

（2）复数域上，线性变换 $\varphi\in L(V,V)$，$\dim V=n$，必有特征向量$\exists\alpha$：$\varphi(\alpha)=\lambda\alpha$ $(\alpha\neq0)$

（3）设 W 是 $\varphi\in L(V,V)$的不变子空间，则 φ 在 W 上必有特征向量$\exists\alpha\in W$：$\varphi(\alpha)=\lambda\alpha$ $(\alpha\neq0)$（$\because\varphi|_W$：$W\to W$ 也是线性变换）

Ex.若 $AB=BA$（A，B 是方阵），令$V(\lambda)=\{x|Ax=\lambda x\}$（特征子空间）

证明：（1）$V(\lambda)$是 A 与 B 的不变子空间

（2）$V(\lambda)$中有一个 $x\neq0$ 是 B 的特征向量（用引理（3））

（3）A、B 有公共特征向量

线性变换的规范表示

R^n 中规范基 $e_1=(1,0,\ldots,0)^T, e_2=(0,1,\ldots,0)^T,\ldots,e_n=(0,0,\ldots,1)^T\in R^n$

R^m 中规范基$\tilde{e}_1=(1,0,\cdots,0)^T,\tilde{e}_2=(0,1,\cdots,0)^T,\cdots,\tilde{e}_m=(0,0,\cdots,1)^T\in R^m$

$x=(x_1,x_2,\ldots,x_n)^T\in R^n$，$x=\sum_{i=1}^{n}x_ie_i$

$y=(y_1, y_2, \dots, y_m)^T \in R^m$， $y=\sum_{i=1}^{m} y_i \widetilde{e}_i$

规范公式： 每个线性的 $\varphi: R^n \to R^m$ 或 $\varphi \in L(R^n, R^m)$都有一个（唯一的）矩阵

$A=A_{m\times n} \in R^{m\times n}$，使得 $\varphi(x)=Ax$，$x\in R^n$，其中 $A=(\varphi(e_1), \varphi(e_2), \dots, \varphi(e_n))_{m\times n}$

Pf：$\because x=(x_1, x_2, \cdots, x_n)^T=\sum_{i=1}^{n} x_i e_i$ ， $\varphi(x)=\sum_{i=1}^{n} x_i \varphi(e_i)=(\varphi(e_1), \varphi(e_2), \cdots, \varphi(e_n))\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$

令 $A=(\varphi(e_1), \varphi(e_2), \cdots, \varphi(e_n))=A_{m\times n} \in R^{m\times n} \Rightarrow \varphi(x)=A\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}=Ax$

实际上： φ 在规范基$(e_1, e_2, \dots, e_n)$与$(\widetilde{e}_1, \widetilde{e}_2, \cdots, \widetilde{e}_m)$下的表示公式

$\varphi(e_1, e_2, \cdots, e_n)=(\widetilde{e}_1, \widetilde{e}_2, \cdots, \widetilde{e}_m)A_{m\times n}$，其中$(\widetilde{e}_1, \widetilde{e}_2, \cdots, \widetilde{e}_m)=I_m$ （单位阵）

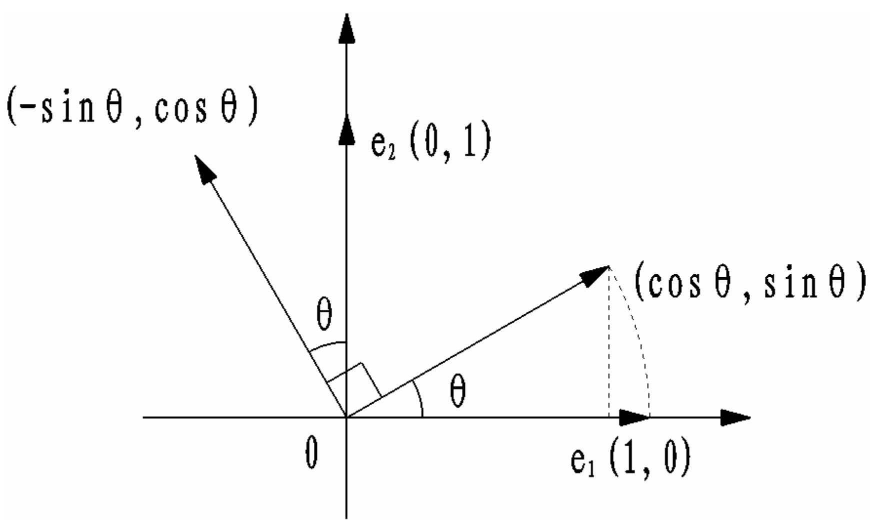
$\Rightarrow \varphi(e_1, e_2, \cdots, e_n)=A_{m\times n}$

在实用中，可把 $\varphi: R^n \to R^m$ 写成 $A: R^n \to R^m$（可写 $\varphi=A$）

矩阵 $A=A_{m\times n}$ 有双重身份：（1）A 是矩阵；（2）$A: R^n \to R^m$（$A\in L(R^n, R^m)$）为线性映射

注： 若 R^n 中为行向量，在公式中应该为列向量

Eg.令 θ 旋转 $\varphi: R^2 \to R^2$，求 $A=A_{2\times 2}$ 使得 $\varphi(x)=Ax$



$\because\ \varphi(e_1)=(\cos\theta,\sin\theta)^T=\begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix}$，$\varphi(e_2)=\begin{pmatrix}-\sin\theta\\ \cos\theta\end{pmatrix}$

令 $A=(\varphi(e_1),\varphi(e_2))=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}$

$\Rightarrow$ 公式 $\varphi(x)=Ax=A\begin{pmatrix}x_1\\ x_2\end{pmatrix}=\begin{pmatrix}x_1\cos\theta-x_2\sin\theta\\ x_1\sin\theta+x_2\cos\theta\end{pmatrix}$

令 $y=Ax$，$y=\begin{pmatrix}y_1\\ y_2\end{pmatrix}\Rightarrow\begin{cases}y_1=x_1\cos\theta-x_2\sin\theta\\ y_2=x_1\sin\theta+x_2\cos\theta\end{cases}$

Eg.令 $\varphi\in L(R^3,R^2)$，即 $\varphi: R^3\to R^2$ 为线性

使得：$\varphi(x)=(x_1+x_2,x_2+x_3)^T$，$\forall x=(x_1,x_2,x_3)^T\in R^3$

求 $A=A_{2\times3}$ 使得 $\varphi(x)=Ax$

解：$\varphi(e_1)=\varphi(1,0,0)^T=(1,0)^T$，$\varphi(e_2)=\varphi(0,1,0)^T=(1,1)^T$，$\varphi(e_3)=\varphi(0,0,1)^T=(0,1)^T$

$A=(\varphi(e_1),\varphi(e_2),\varphi(e_3))=\begin{pmatrix}1&1&0\\0&1&1\end{pmatrix}_{2\times3}$

计算 $\varphi(x)=A\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}=\begin{pmatrix}x_1+x_2\\ x_2+x_3\end{pmatrix}$

一般表示公式：设$(\varepsilon)=(\varepsilon_1,\varepsilon_2,\dots,\varepsilon_n)$，$(g)=(g_1,g_2,\dots,g_m)$分别为

$\varphi: R^n\to R^m$ 为线性的

设表示式：$\varphi(\varepsilon_1,\varepsilon_2,\dots,\varepsilon_n)=(g_1,g_2,\dots,g_m)A$

则有：$A=(g_1,g_2,\dots,g_m)^{-1}(\varphi(\varepsilon_1),\varphi(\varepsilon_2),\dots,\varphi(\varepsilon_n))$

其中$(g_1,g_2,\dots,g_m)$为可逆方阵

方法：可用行变换 $(g_1,g_2,\cdots,g_n|\varphi(\varepsilon_1),\varphi(\varepsilon_2),\cdots,\varphi(\varepsilon_n))\xrightarrow{\text{行变}}(I_m|A)$ 求出 A

注： $R^{m\times n}$ 中的矩阵 $A=(a_{ij})_{m\times n}$，$B=(b_{ij})_{m\times n}$

可用"拉直法"：$A\to\vec{A}=(a_{11},a_{12},\cdots,a_{mn})^T\in R^{mn}$，$B\to\vec{B}=(b_{11},b_{12},\cdots,b_{mn})^T\in R^{mn}$

"$\to$"：$R^{m\times n}\to R^{mn}$ 为线性（同构）

$\overrightarrow{(A+B)}=\overrightarrow{(a_{ij}+b_{ij})}=(a_{11}+b_{11},\cdots,a_{mn}+b_{mn})^T=(a_{11},\cdots,a_{mn})^T+(b_{11},\cdots,b_{mn})^T=\vec{A}+\vec{B}$

同理 $\overrightarrow{(kA)}=k\vec{A}$

应用： 若 $\varphi: R^{2\times 2}\to R^{2\times 2}$ 为线性，取基 $(\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4)$，$(g_1,g_2,g_3,g_4)$

由公式 $\varphi(\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4)=(g_1,g_2,g_3,g_4)A_{4\times 4}$

拉直 $\left(\overrightarrow{\varphi(\varepsilon_1)},\overrightarrow{\varphi(\varepsilon_2)},\overrightarrow{\varphi(\varepsilon_3)},\overrightarrow{\varphi(\varepsilon_4)}\right)=(\vec{g}_1,\vec{g}_2,\vec{g}_3,\vec{g}_4)A$

$\Rightarrow A=(\vec{g}_1,\vec{g}_2,\vec{g}_3,\vec{g}_4)^{-1}\left(\overrightarrow{\varphi(\varepsilon_1)},\overrightarrow{\varphi(\varepsilon_2)},\overrightarrow{\varphi(\varepsilon_3)},\overrightarrow{\varphi(\varepsilon_4)}\right)$

实用中可写： $\alpha=\begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix}=(a_1,a_2,a_3,a_4)^T$

Ex.1.令 $\varphi: R^2\to R^3$ 为线性的

且 $\forall x\in R^2$，$\varphi(x)=(x_2,x_1+x_2,x_1-x_2)^T$

（1） 求规范公式 $\varphi(x)=Ax$ 中的 A

（2） 若取基$(\varepsilon_1,\varepsilon_2)$与$(g_1,g_2,g_3)$，其中 $\varepsilon_1=(1,2)^T$，$\varepsilon_2=(3,1)^T$，$g_1=(1,0,0)^T$，$g_2=(1,1,0)^T$，$g_3=(1,1,1)^T$，求公式 $\varphi(\varepsilon_1,\varepsilon_2)=(g_1,g_2,g_3)B$ 中的表示阵 B（可用初等行变换求 B）

线性变换应用参考书： Steven Leon《线性代数与应用》

§4 应用 1：计算机图形与动画设计；应用 2：飞机运动矩阵表示

§3 欧式空间与 QR 分解

标准欧式空间：R^n 中引入标准内积（点积）

标准内积（点积） $x\bullet y=(x,y)=(x_1,x_2,\cdots,x_n)\bullet(y_1,y_2,\cdots,y_n)=\sum_{i=1}^{n}x_iy_i$

有公式：$x\bullet y=(x,y)=x^Ty=(x_1,x_2,\cdots,x_n)\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}=\sum_{i=1}^{n}x_iy_i$，$x\bullet x=x^Tx=\sum_{i=1}^{n}x_i^2$

长度公式： $|x|=\sqrt{x\bullet x}=\sqrt{\sum_{i=1}^{n}x_i^2}$，$|x|^2=x\bullet x=\sum_{i=1}^{n}x_i^2$

正交（垂直）： $x\perp y\Leftrightarrow x\bullet y=(x,y)=0$

勾股定理：（1） $x\perp y\Rightarrow(x\pm y)^2=|x|^2+|y|^2$

（2） $x \perp y \Rightarrow (kx \pm ly)^2 = k^2|x|^2 + l^2|y|^2$ （$\because kx \perp ly$）

正交组与正交基：若 $\alpha_1, \alpha_2, \ldots,\alpha_s \in R^n$ 互相正交（且非 0），($\alpha_1 \perp \alpha_2 \perp \ldots \perp \alpha_s$)

称 $\alpha_1, \alpha_2, \ldots,\alpha_s$ 为一个正交组

称生成空间 $W = \mathrm{span}(\alpha_1, \alpha_2, \ldots,\alpha_s)$中有正交基 $\alpha_1, \alpha_2, \ldots,\alpha_s$

若单位化：$\varepsilon_1 = \dfrac{\alpha_1}{|\alpha_1|}, \varepsilon_2 = \dfrac{\alpha_2}{|\alpha_2|}, \cdots, \varepsilon_s = \dfrac{\alpha_s}{|\alpha_s|}$，可得单位（规范）正交组（基）$\varepsilon_1, \varepsilon_2, \ldots,\varepsilon_s$

定义：若 $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_s \in R^n$ 为单位正交组（基）

称矩阵 $A = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_s)_{n\times s}$ 为正交高阵（次正交阵）$(s \le n)$

特别：$s = n$ 时 $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ 为 R^n 中正交基

称 $A = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)_{n\times n}$ 为正交阵

例：$A = (\varepsilon_1, \varepsilon_2) = \dfrac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & -1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \\ 1/2 & -1/2 \\ 1/2 & -1/2 \end{pmatrix}$为正交高阵

$\varepsilon_1 \perp \varepsilon_2 \Leftrightarrow \varepsilon_1 \bullet \varepsilon_2 = 0$

计算 $A^T A = \begin{pmatrix} \varepsilon_1^T \\ \varepsilon_2^T \end{pmatrix} (\varepsilon_1, \varepsilon_2) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2$

正交高阵性质：$A = A_{n\times s}$ 为次正交 $A^TA = I_s$

Pf：$A = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_s)$，$A^T = \begin{pmatrix} \varepsilon_1^T \\ \varepsilon_2^T \\ \vdots \\ \varepsilon_s^T \end{pmatrix} \Rightarrow A^T A = \begin{pmatrix} \varepsilon_1^T \\ \varepsilon_2^T \\ \vdots \\ \varepsilon_s^T \end{pmatrix} (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_s) = \begin{pmatrix} 1 & & 0 \\ & 1 & \\ & & \ddots & \\ 0 & & & 1 \end{pmatrix}$

特别：$A = A_{n\times n}$ 为正交阵 $\Leftrightarrow A^T A = I_n$ （此时 $A^T = A^{-1}$）

***QR* 公式**：若 $A = A_{n\times s} = (\alpha_1, \alpha_2, \ldots,\alpha_s)$，秩为 $\mathrm{rank}(A) = s$，（$\alpha_1, \alpha_2, \ldots,\alpha_s$ 无关）

则有正交高阵 $Q = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_s)_{n\times s}$ 与上三角阵 $R = R_{n\times s}$

使得 $A=QR=\left(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_s\right)\begin{pmatrix} t_1 & & (*) \\ & t_2 & \ddots \\ 0 & & \\ & & t_s \end{pmatrix}$

Ex.《矩阵分析》P70 12（1）（2） 13 P68 3 6

QR 分解：若 $A=A_{n\times s}=(\alpha_1,\alpha_2,\ldots,\alpha_s)$为高阵，（rank($A$) = 列数）

则分解 $A=QR$ 基中 $Q=Q_{n\times s}$ 为正交高阵（次正交阵），R 为上三角

Pf：由许米特（Schmidt）正交公式

令
$$\begin{aligned}
\beta_1 &= \alpha_1 \\
\beta_2 &= \alpha_2-\frac{\left(\alpha_2\bullet\beta_1\right)}{\left|\beta_1\right|^2}\beta_1 \\
&\vdots \\
\beta_s &= \alpha_s-\frac{\left(\alpha_s\bullet\beta_1\right)}{\left|\beta_1\right|^2}\beta_1-\frac{\left(\alpha_s\bullet\beta_2\right)}{\left|\beta_2\right|^2}\beta_2-\cdots-\frac{\left(\alpha_s\bullet\beta_{s-1}\right)}{\left|\beta_{s-1}\right|^2}\beta_{s-1}
\end{aligned}$$

注：此时 $\beta_1\perp\beta_2\perp\ldots\perp\beta_s$（互正交）

且 $\alpha_1,\alpha_2,\ldots,\alpha_s$ 与 $\beta_1,\beta_2,\ldots,\beta_s$ 互相表示

$$\begin{cases}\alpha_1=\beta_1 \\ \alpha_2=(*)\beta_1+\beta_2 \\ \vdots \\ \alpha_s=(*)\beta_1+(*)\beta_2+\cdots+\beta_s\end{cases} \Rightarrow \left(\alpha_1,\alpha_2,\cdots,\alpha_s\right)=\left(\beta_1,\beta_2,\cdots,\beta_s\right)\begin{pmatrix}1 & * & \cdots & * \\ 0 & 1 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1\end{pmatrix}$$

单位化 $\varepsilon_1=\frac{\beta_1}{\left|\beta_1\right|},\varepsilon_2=\frac{\beta_2}{\left|\beta_2\right|},\cdots,\varepsilon_s=\frac{\beta_s}{\left|\beta_s\right|}$ 或 $\beta_1=|\beta_1|\varepsilon_1,\beta_2=|\beta_2|\varepsilon_2,\ldots,\beta_s=|\beta_s|\varepsilon_s$

$$\Rightarrow\left(\beta_1,\beta_2,\cdots,\beta_s\right)=\left(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_s\right)\begin{pmatrix}\left|\beta_1\right| & & O \\ & \left|\beta_2\right| & \ddots \\ O & & \\ & & \left|\beta_s\right|\end{pmatrix}_{s\times s}$$ 代入上式

$$\Rightarrow A=\left(\alpha_1,\alpha_2,\cdots,\alpha_s\right)\left(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_s\right)\begin{pmatrix}\left|\beta_1\right| & & * \\ & \left|\beta_2\right| & \ddots \\ O & & \\ & & \left|\beta_s\right|\end{pmatrix}\begin{pmatrix}1 & & (*) \\ & 1 & \ddots \\ O & & \\ & & 1\end{pmatrix}_{s\times s}$$

$$令\ R=\begin{pmatrix}|\beta_1| & & & * \\ & |\beta_2| & & \\ & & \ddots & \\ O & & & |\beta_s|\end{pmatrix}\begin{pmatrix}1 & & & (*) \\ & 1 & & \\ & & \ddots & \\ O & & & 1\end{pmatrix}=\begin{pmatrix}|\beta_1| & & & (*)' \\ & |\beta_2| & & \\ & & \ddots & \\ O & & & |\beta_s|\end{pmatrix}上三角$$

$Q=(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_s)_{n\times s}$ 为正交高阵

$\Rightarrow A=QR=Q_{n\times s}R_{s\times s}$

特别：$A=A_{n\times n}$ 为可逆方阵，也有 $A=Q_{n\times n}R_{n\times n}$

注：若 $Q=Q_{n\times s}$ 为正交高阵，则 $Q^TQ=I_s$

由 $A=QR\Rightarrow R=Q^TA$（$\because Q^TQ=I$）

Eg. $A=(\alpha_1,\alpha_2,\alpha_3)=\begin{pmatrix}1 & -1 & 4\\ 1 & 4 & -2\\ 1 & 4 & 2\\ 1 & -1 & 0\end{pmatrix}_{4\times 3}$（高阵）

解：$\alpha_1=(1, 1, 1, 1)^T$，$\alpha_2=(-1, 4, 4, -1)^T$，$\alpha_3=(4, -2, 2, 0)^T$

令 $\beta_1=\alpha_1=(1, 1, 1, 1)^T$，$|\beta_1|^2=4$，$|\beta_1|=2$

$$\beta_2=\alpha_2-\frac{(\alpha_2\bullet\beta_1)}{|\beta_1|^2}\beta_1=\left(-\frac{5}{2},\frac{5}{2},\frac{5}{2},-\frac{5}{2}\right)^T=\frac{5}{2}(-1,1,1,-1)^T，|\beta_2|^2=25，|\beta_2|=5$$

$$\beta_3=\alpha_3-\frac{(\alpha_3\bullet\beta_1)}{|\beta_1|^2}\beta_1-\frac{(\alpha_3\bullet\beta_2)}{|\beta_2|^2}\beta_2=(2,-2,2,-2)^T，|\beta_3|^2=16，|\beta_3|=4$$

单位化：$\varepsilon_1=\dfrac{\beta_1}{|\beta_1|}=\dfrac{1}{2}(1,1,1,1)^T,\varepsilon_2=\dfrac{\beta_2}{|\beta_2|}=\dfrac{1}{2}(-1,1,1,-1)^T,\varepsilon_3=\dfrac{\beta_3}{|\beta_3|}=\dfrac{1}{2}(1,-1,1,-1)^T$

$$令\ Q=(\varepsilon_1,\varepsilon_2,\varepsilon_3)=\begin{pmatrix}1/2 & -1/2 & 1/2\\ 1/2 & 1/2 & -1/2\\ 1/2 & 1/2 & 1/2\\ 1/2 & -1/2 & -1/2\end{pmatrix}_{4\times 3}（正交高阵）$$

$$令\ R=Q^TA=\frac{1}{2}\begin{pmatrix}1 & 1 & 1 & 1\\ -1 & 1 & 1 & -1\\ 1 & -1 & 1 & -1\end{pmatrix}A=\begin{pmatrix}2 & 3 & 2\\ 0 & 5 & -2\\ 0 & 0 & 4\end{pmatrix}（上三角）$$

$\Rightarrow A=QR$

Ex.求 QR 分解

（1）$A=\begin{pmatrix}0 & 3 & 1\\ 0 & 4 & -2\\ 2 & 1 & 2\end{pmatrix}$ （2）$A=\begin{pmatrix}1 & 1\\ 1 & -1\\ 1 & 1\\ 1 & 0\end{pmatrix}_{4\times 2}$

正交阵定义：若方阵 $A=A_{n\times n}$ 的 n 个列 $\alpha_1,\alpha_2,\ldots,\alpha_n$ 为单位正交组（基）

性质：设 $A=(\alpha_1,\alpha_2,\ldots,\alpha_n)_{n\times n}$ 为正交阵$(\alpha_1\perp\alpha_2\perp\ldots\perp\alpha_n)$

（1）$A^TA=I_n$ 且 $A^{-1}=A^T$ 或 $AA^T=I_n$

（2）长度公式：$|Ax|^2=|x|^2$（$x\in R^n$）（$\because |Ax|^2=(Ax)^T(Ax)=x^Tx$）

复欧空间（酉空间）C^n

设复 n 元数组空间 $C^n=\left\{x=(x_1,x_2,\cdots,x_n)^T \middle| x_1,x_2,\cdots,x_n\in C\right\}$

任取 $x=(x_1,x_2,\ldots,x_n)^T,\ y=(y_1,y_2,\ldots,y_n)^T\in C^n$

规定：标准内积（点积）如下：$(x,y)=x\bullet y=y^Hx=(\bar{y}_1,\bar{y}_2,\cdots,\bar{y}_n)\begin{pmatrix}x_1\\ x_2\\ \vdots\\ x_n\end{pmatrix}=\sum_{i=1}^n x_i\bar{y}_i$

注：y^H 表示复共轭转置也叫 Hermite 转置

复内积性质：

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

（2）$(kx,y)=k(x,y)$，$(x,ky)=\bar{k}(x,y)$或 $x\bullet(ky)=\bar{k}(x\bullet y)$

（3）$(x,y+z)=(x,y)+(x,z)$或 $x\bullet(y+z)=x\bullet y+x\bullet z$

（4）正定性：$(x,x)=x^Hx\geq 0$，长度公式：$|x|=\sqrt{(x,x)}=\sqrt{x^Hx}=\sqrt{\sum_{i=1}^n|x_i|^2}$，$x\in C^n$

注：$x^Hx=\sum_{i=1}^n\bar{x}_ix_i=\sum_{i=1}^n|x_i|^2$

许互次（Schwarz）不等式：$|(x,y)|\leq|x|\bullet|y|$

正交定义：$x\perp y\Leftrightarrow(x,y)=x\bullet y=0$（$\sum_{i=1}^n x_i\bar{y}_i=0$）

注：$(x,y)=x\bullet y=0$ 必有$(y,x)=y\bullet x=0$，$\because (y,x)=\overline{(x,y)}=\bar{0}=0$

引理：$x \perp y \Leftrightarrow (x,y)=0 \Leftrightarrow (y,x)=0$

勾股定理：$x \perp y \Rightarrow |(x \pm y)|^2 = |x|^2 + |y|^2$

Pf: $|(x+y)|^2 = (x+y) \bullet (x+y) = x \bullet x + y \bullet y + x \bullet y + y \bullet x = |x|^2 + |y|^2$

次酉阵定义：若 $A = A_{n\times s}$ 中列 $\alpha_1, \alpha_2, \ldots, \alpha_s$ 是单位正交组，则称 A 为次酉阵

称 $A = (\alpha_1, \alpha_2, \ldots, \alpha_s)_{n\times s}$ 为次酉阵，$\alpha_1 \perp \alpha_2 \perp \ldots \perp \alpha_s$，$|\alpha_1|^2 = |\alpha_2|^2 = \ldots = \alpha_s|^2 = 1$

性质：$A = A_{n\times s}$ 为次酉阵 $\Leftrightarrow \overline{A}^T A = I_s$ 记为 $A^H A = I_s$

注：$A^H = \overline{A}^T = \overline{A^T}$ 表示 Hermite 转置

Pf: $\because \alpha_1 \perp \alpha_2 \perp \cdots \perp \alpha_s \Rightarrow \alpha_1^H \alpha_2 = 0, \cdots, \alpha_s^H \alpha_{s-1} = 0$

$$\Rightarrow A^H A = \begin{pmatrix} \alpha_1^H \\ \alpha_2^H \\ \vdots \\ \alpha_s^H \end{pmatrix} (\alpha_1, \alpha_2, \cdots, \alpha_s) = \begin{pmatrix} 1 & & O \\ & 1 & \\ & & \ddots \\ O & & & 1 \end{pmatrix} = I_s$$

特别对方阵 $A = A_{n\times n} = (\alpha_1, \alpha_2, \ldots, \alpha_n)$

若各列互正交且长度为 1，则称 A 为酉阵

酉阵性质：$A = A_{n\times n}$ 为酉阵 $\Rightarrow A^H A = I_n$ 或 $A^{-1} = A^H$

引理：$A = A_{n\times n}$ 为酉阵 $\Leftrightarrow A^H A = AA^H = I_n \Leftrightarrow A^{-1} = A^H$

注：用“$\mathscr{U}$”表示“酉”

$\mathscr{U}R$ 分解公式：每个高阵 $A = A_{n\times s} = (\alpha_1, \alpha_2, \ldots, \alpha_s)$（rank($A$) = 列数）

都有分解 $A = QR$，$Q = Q_{n\times s}$ 为次酉，R 为上三角（正交线性）

注：许 Schmidt 正交公式在 $\mathscr{U}$空间 C^n 中也成立

若 $\alpha_1, \alpha_2, \ldots, \alpha_s$ 为无关组

令$\begin{cases} \beta_1 = \alpha_1 \\ \beta_2 = \alpha_2 - \dfrac{(\alpha_2 \bullet \beta_1)}{|\beta_1|^2}\beta_1 \\ \vdots \\ \beta_s = \alpha_s - \dfrac{(\alpha_s \bullet \beta_1)}{|\beta_1|^2}\beta_1 - \dfrac{(\alpha_s \bullet \beta_2)}{|\beta_2|^2}\beta_2 - \cdots - \dfrac{(\alpha_s \bullet \beta_{s-1})}{|\beta_{s-1}|^2}\beta_{s-1} \end{cases}$，则 $\beta_1 \perp \beta_2 \perp \ldots \perp \beta_s$

Eg. $A=\begin{pmatrix}1 & i\\ i & 1\end{pmatrix}=(\alpha_1,\alpha_2)$，求 QR 分解

$\alpha_1=(1,i)^T$，$\alpha_2=(i,1)^T$

$\beta_1=\alpha_1=(1,i)^T$，$|\beta_1|^2=2$，$|\beta_1|=\sqrt{2}$

$$\beta_2=\alpha_2-\frac{(\alpha_2\bullet\beta_1)}{|\beta_1|^2}\beta_1=\alpha_2-0\bullet\alpha_1=\alpha_2=(i,1)^T$$

$\beta_1\perp\beta_2$ $(\alpha_1\perp\alpha_2)$

$$\varepsilon_1=\frac{\beta_1}{|\beta_1|}=\frac{1}{\sqrt{2}}\beta_1,\quad \varepsilon_2=\frac{\beta_2}{|\beta_2|}=\frac{1}{\sqrt{2}}\beta_2$$

$$Q=(\varepsilon_1,\varepsilon_2)=\begin{pmatrix}\frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}}\\ \frac{i}{\sqrt{2}} & \frac{1}{\sqrt{2}}\end{pmatrix}$$（为 $\mathscr{U}$阵）

令 $A=QR\Rightarrow R=Q^HA$

C^n 中标准内积（点积）（称 C^n 为复欧空间或 $\mathscr{U}$空间）

$x=(x_1,x_2,\ldots,x_n)^T, y=(y_1,y_2,\ldots,y_n)^T\in C^n$

内积为： $$x\bullet y=(x,y)=\sum_{i=1}^{n}x_i\bar{y}_i=(\bar{y}_1,\bar{y}_2,\cdots,\bar{y}_n)\begin{pmatrix}x_1\\ x_2\\ \vdots\\ x_n\end{pmatrix}=y^Hx$$

特别： $x,y\in R^n\subset C^n$，有 $x\bullet y=(x,y)=\sum_{i=1}^{n}x_iy_i$

性质：（1）$y\bullet x=\overline{x\bullet y}$；（2）$x\bullet(y+z)=x\bullet y+y\bullet z$;

（3）$(x,ky)=\bar{k}(x,y)$；（4）$x\bullet x=x^Hx=\sum_{i=1}^{n}|x_i|^2$（长度平方）

C^n 中的正交条件 “$x\perp y$”

定义： $x\perp y\Leftrightarrow x\bullet y=0$ 或 $y\bullet x=0$

勾股定理： $x\perp y\Rightarrow |(x\pm y)|^2=|x|^2+|y|^2$

Eg. $\alpha=(1,i,i)^T$，$\beta=(2,-i,-i)^T$，则 $\alpha\perp\beta$

$$\because \alpha \bullet \beta = \beta^H \alpha = (\overline{2}, \overline{-i}, \overline{-i})\begin{pmatrix}1\\ i\\ i\end{pmatrix} = 2 + i^2 + i^2 = 0$$

验证：$|(\alpha+\beta)|^2 = |\alpha|^2 + |\beta|^2$

次酉阵定义：若 $A = A_{n\times s}$ 中列 $\alpha_1, \alpha_2, \ldots, \alpha_s$ 是单位正交组($\alpha_1 \perp \alpha_2 \perp \ldots \perp \alpha_s$)

则称 $A = (\alpha_1, \alpha_2, \ldots, \alpha_s)_{n\times s}$ 为次酉阵

引理：$A = A_{n\times s}$ 为次酉阵 $\Leftrightarrow \overline{A}^T A = I_s$

酉阵定义：若方阵 $A = A_{n\times n} = (\alpha_1, \alpha_2, \ldots, \alpha_n)$的列构成单位正交基，称 A 为 $\mathscr{U}$阵

引理：$A = A_{n\times n}$ 为 $\mathscr{U}$阵 $\Rightarrow A^H A = I_n$ 或 $A^{-1} = A^H$

特别：实正交阵（$A \in R^{n\times n}$，$A^T A = I_n$）都是 $\mathscr{U}$阵

例：$A = \dfrac{1}{\sqrt{2}}\begin{pmatrix}1 & i\\ i & 1\end{pmatrix} = \begin{pmatrix}1/\sqrt{2} & i/\sqrt{2}\\ i/\sqrt{2} & 1/\sqrt{2}\end{pmatrix}$，$B = \begin{pmatrix}0 & -2/\sqrt{6} & 1/\sqrt{3}\\ 1/\sqrt{2} & i/\sqrt{6} & i/\sqrt{3}\\ i/\sqrt{2} & 1/\sqrt{6} & 1/\sqrt{3}\end{pmatrix}$为 $\mathscr{U}$阵

注：Schmidt 正交化公式仍成立

设 $\alpha_1, \alpha_2, \ldots, \alpha_s$ 为无关组，则 $\beta_1, \beta_2, \ldots, \beta_s$ 互相正交

其中：
$$\begin{cases}\beta_1 = \alpha_1\\ \beta_2 = \alpha_2 - \dfrac{(\alpha_2 \bullet \beta_1)}{|\beta_1|^2}\beta_1\\ \vdots\\ \beta_s = \alpha_s - \dfrac{(\alpha_s \bullet \beta_1)}{|\beta_1|^2}\beta_1 - \dfrac{(\alpha_s \bullet \beta_2)}{|\beta_2|^2}\beta_2 - \cdots - \dfrac{(\alpha_s \bullet \beta_{s-1})}{|\beta_{s-1}|^2}\beta_{s-1}\end{cases}$$

QR（或 $\mathscr{U}R$）分解

（1）若 $A = (\alpha_1, \alpha_2, \ldots, \alpha_s)_{n\times s}$ 为高阵（$\mathrm{rank}(A) = s$）

则 $A = QR$，$Q = Q_{n\times s}$ 为次酉阵（$Q^H Q = I_s$），R 为上三角阵

（2）若方阵 $A = (\alpha_1, \alpha_2, \ldots, \alpha_n)_{n\times n}$ 为可逆方阵

则 $A = QR$，$Q = Q_{n\times n}$ 为酉阵（$Q^H Q = I_n$），R 为上三角阵

方法：先把 A 中列正交单位化可得 Q，设 $A = QR$ 解出 $R = Q^H A$

Ex.求 QR 分解

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

许尔公式：每个方阵 $A=A_{n\times n}$ 相似于上三角阵

即：$P^{-1}AP=\begin{pmatrix}\lambda_1 & & (*)\\ & \lambda_2 & \ddots \\ O & & & \lambda_n\end{pmatrix}$

许尔公式 2：每个方阵 $A=A_{n\times n}$ 都酉相似于上三角阵

即存在 $\mathscr{U}$ 阵 Q 使得 $Q^{-1}AQ=Q^{H}AQ=\begin{pmatrix}\lambda_1 & & (*)\\ & \lambda_2 & \ddots \\ O & & & \lambda_n\end{pmatrix}$

Pf：$\because P^{-1}AP=\begin{pmatrix}\lambda_1 & & (*)\\ & \lambda_2 & \ddots \\ O & & & \lambda_n\end{pmatrix}$

用 QR 分解 $P=QR$ 写 $R=\begin{pmatrix}t_1 & & (*)\\ & t_2 & \ddots \\ O & & & t_n\end{pmatrix}$，$Q^{H}Q=I_n$，$Q^{-1}=Q^{H}$

$$P^{-1}AP=R^{-1}\left(Q^{-1}AQ\right)R=\begin{pmatrix}\lambda_1 & & (*)\\ & \lambda_2 & \ddots \\ O & & & \lambda_n\end{pmatrix}$$

$$\Rightarrow Q^{-1}AQ=R\begin{pmatrix}\lambda_1 & & (*)\\ & \lambda_2 & \ddots \\ O & & & \lambda_n\end{pmatrix}R^{-1}$$

$$=\begin{pmatrix}t_1 & & (*)\\ & t_2 & \ddots \\ O & & & t_n\end{pmatrix}\begin{pmatrix}\lambda_1 & & (*)\\ & \lambda_2 & \ddots \\ O & & & \lambda_n\end{pmatrix}\begin{pmatrix}t_1^{-1} & & (*)\\ & t_2^{-1} & \ddots \\ O & & & t_n^{-1}\end{pmatrix}=\begin{pmatrix}\lambda_1 & & (*)'\\ & \lambda_2 & \ddots \\ O & & & \lambda_n\end{pmatrix}$$

Hermite 阵定义：若 $A^H = A$ 称 A 为 Hermite 阵

注：若 $A^H = A$ 则 A 为方阵

例：$A=\begin{pmatrix}1 & 1-i\\ 1+i & 2\end{pmatrix}$，$B=\begin{pmatrix}3 & i\\ -i & 5\end{pmatrix}$为 Hermite 阵

$$A^H=\begin{pmatrix}\overline{1} & \overline{1+i}\\ \overline{1-i} & \overline{2}\end{pmatrix}=\begin{pmatrix}1 & 1-i\\ 1+i & 2\end{pmatrix}=A，B^H=\begin{pmatrix}\overline{3} & \overline{-i}\\ \overline{i} & \overline{5}\end{pmatrix}=\begin{pmatrix}3 & i\\ -i & 5\end{pmatrix}=B$$

特别：是对称阵 $A=A^T\in R^{n\times n}$ 也是 Hermite 阵（$\because A^H=\overline{A}^T=A^T$）

反 Hermite 阵定义：若 $A^H=-A$

实反对称阵 $A^H=-A\in R^{n\times n}$ 也是反 Hermite 阵

引理：（1）A 为 Hermite 阵$\Leftrightarrow iA$ 为反 Hermite 阵或$\dfrac{A}{i}$为反 Hermite 阵

（2）A 为反 Hermite 阵$\Leftrightarrow iA$ 为 Hermite 阵

Pf:（1）$\because (iA)^H=(\bar{i})A^H=(-i)A=-iA$

例：$A=i\begin{pmatrix}1 & 1-i\\ 1+i & 2\end{pmatrix}=\begin{pmatrix}i & 1+i\\ -1+i & 2i\end{pmatrix}$为反 Hermite 阵

注：反 Hermite 阵对角线为纯虚的（或 0）

注：$(AB)^H=B^HA^H$，$(A+B)^H=A^H+B^H$

Hermite 阵对角线为实数

Eg.设 $\varepsilon=(\alpha_1, \alpha_2, \ldots, \alpha_n)^T\in C^n$，$|\varepsilon|^2=\varepsilon^H\varepsilon=\sum_{j=1}^{n}|a_j|^2=1$（单位长）

令 $Q=I_n-2\varepsilon\varepsilon^H$

则（1）$Q^H=Q$；（2）$Q^HQ=I_n$，即 Q 为 $\mathscr{U}$阵；（3）$Q^{-1}=Q$

解：（1）$Q^H=(I_n-2\varepsilon\varepsilon^H)^H=(I_n)^H-2(\varepsilon\varepsilon^H)^H=I_n-2\varepsilon\varepsilon^H=Q$

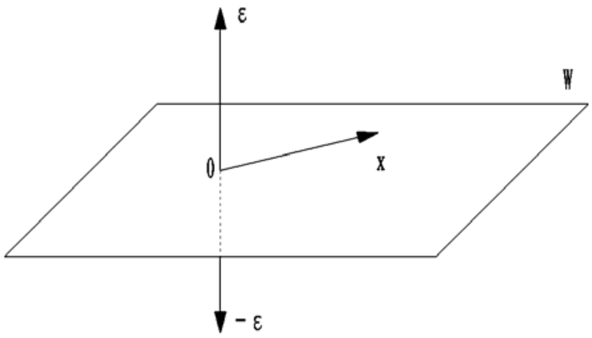
（2）$Q^HQ=Q\bullet Q=(I_n-2\varepsilon\varepsilon^H)^2=I_n+4(\varepsilon\varepsilon^H)(\varepsilon\varepsilon^H)-4\varepsilon\varepsilon^H=I_n$，$Q$ 为 $\mathscr{U}$阵

称这种 $\mathscr{U}$阵 Q 为镜面阵（或 Householder 阵）

镜面阵性质：设 $Q=I_n-2\varepsilon\varepsilon^H$，$|\varepsilon|^2=\varepsilon^H\varepsilon=1$

△（在空间 R^n 中 $\varepsilon^H=\varepsilon^T$；$Q=I_n-2\varepsilon\varepsilon^H$）

如图：以 ε 为法向做一个“平面”（正交补空间）W



（1）$Q\varepsilon = -\varepsilon$，$\varepsilon$ 是属于-1 的特征向量

（2）若 $x\perp\varepsilon$，则 $Qx = x$，属于 1 的特征向量

Pf:（1）$Q\varepsilon = (I_n - 2\varepsilon\varepsilon^T)\varepsilon = \varepsilon - 2\varepsilon\varepsilon^T\varepsilon = \varepsilon - 2\varepsilon = -\varepsilon$

（2）若 $x \perp \varepsilon \Rightarrow \varepsilon^T x = 0$，$Qx = (I_n - 2\varepsilon\varepsilon^T)x = \mathrm{x} - 2\varepsilon(\varepsilon^T x) = x$

（3）Q 恰有 n 个特征向量：$\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{n-1}$（无关）

其中 $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{n-1}$ 为 $W = \varepsilon^{\perp}$中的基，属于 1 的特征向量

令 $P = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{n-1})$为可逆

$$\Rightarrow P^{-1}QP = \begin{pmatrix} -1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix} \Rightarrow Q\text{有 } n \text{ 个特征值为}\{-1, 1, \ldots, 1\}（n-1\text{ 重}）$$

$$\Rightarrow |Q| = \left|I_n - 2\varepsilon\varepsilon^H\right| = (-1)\bullet 1\bullet\cdots\bullet 1 = -1$$

注： 若 $A = A_{n\times n}$ 为 $\mathscr{U}$阵（或正交阵），则有：

（1）保长度：$|Ax|^2 = |x|^2$

（2）保内积：$(Ax, Ay) = (x, y)$

Pf:（1）$|Ax|^2 = (Ax)^H(Ax) = y^H(A^HA)x = x^Hx = |x|^2$

（2）$(Ax, Ay) = (Ay)^H(Ax) = y^H(A^HA)x = y^Hx = (x, y)$

推论： 若 $x\perp y$，则 $Ax\perp Ay$，A 为 $\mathscr{U}$阵

正规阵条件： $A^HA = AA^H$

注： 正规阵必为方阵

可知：Hermite 阵；反 Hermite 阵；$\mathscr{U}$阵（正交阵）都是正规的

例：$A = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$，$A^HA = AA^H$，$A$ 为正规的

引理：上三角正规阵一定是对角阵

Pf：设 $A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ & a_{22} & \cdots & a_{2n} \\ & & \ddots & \vdots \\ O & & & a_{nn} \end{pmatrix}$，$A^H=\begin{pmatrix} \overline{a_{11}} & & & O \\ \overline{a_{12}} & \overline{a_{22}} & & \\ \vdots & \vdots & \ddots & \\ \overline{a_{1n}} & \overline{a_{2n}} & \cdots & \overline{a_{nn}} \end{pmatrix}$

由条件：$AA^H=A^HA\Rightarrow\sum_{i=1}^{n}\left|a_{1i}\right|^2=\left|a_{11}\right|^2\Rightarrow\sum_{i=2}^{n}\left|a_{1i}\right|^2=0$

同理：$a_{23}=a_{24}=\ldots=a_{2n}=0$

$$\Rightarrow A=\begin{pmatrix} a_{11} & & & O \\ & a_{22} & & \\ & & \ddots & \\ O & & & a_{nn} \end{pmatrix}$$

推论：若 A 为上三角正交阵，则 A 为对角阵。

正规阵理论（$A^HA=AA^H$）

引理：（1）每个上三角正规阵一定是对角阵

（2）正规阵经过 $\mathscr{U}$变换仍是正规阵：A 为正规阵，且 Q 为 $\mathscr{U}$阵$\Rightarrow Q^HAQ$为正规阵

Pf：$\because A^HA=AA^H\Rightarrow Q^HA^HAQ=Q^HAA^HQ\Rightarrow\left(Q^HA^HQ\right)\left(Q^HAQ\right)=\left(Q^HAQ\right)\left(Q^HA^HQ\right)$

正规分解：$A=A_{n\times n}$为正规阵，则有阵 Q（$Q^HQ=I_n$，$Q^{-1}=Q^H$）

使得 $Q^HAQ=Q^{-1}AQ=\begin{pmatrix} \lambda_1 & & & O \\ & \lambda_2 & & \\ & & \ddots & \\ O & & & \lambda_n \end{pmatrix}$

Pf：用许尔（第 2 公式）$\Rightarrow$存在 $\mathscr{U}$阵 Q 使得 $Q^HAQ=\begin{pmatrix} \lambda_1 & & & (*) \\ & \lambda_2 & & \\ & & \ddots & \\ O & & & \lambda_n \end{pmatrix}$（上三角）

且 Q^HAQ 也正规，由引理 Q^HAQ 为对角阵 $Q^HAQ=\begin{pmatrix} \lambda_1 & & & O \\ & \lambda_2 & & \\ & & \ddots & \\ O & & & \lambda_n \end{pmatrix}$

正规阵结论

写 $Q=(q_1, q_2, \ldots, q_n)$ （$q_1, q_2, \ldots, q_n$ 互正交）

$$Q^{-1}AQ=\begin{pmatrix}\lambda_1 & & & O\\ & \lambda_2 & & \\ & & \ddots & \\ O & & & \lambda_n\end{pmatrix}\Leftrightarrow AQ=Q\begin{pmatrix}\lambda_1 & & & O\\ & \lambda_2 & & \\ & & \ddots & \\ O & & & \lambda_n\end{pmatrix}$$

$$\Leftrightarrow A(q_1,q_2,\cdots,q_n)=(q_1,q_2,\cdots,q_n)\begin{pmatrix}\lambda_1 & & & O\\ & \lambda_2 & & \\ & & \ddots & \\ O & & & \lambda_n\end{pmatrix}$$

$(Aq_1, Aq_2, \ldots, Aq_n)=(\lambda_1q_1, \lambda_2q_2, \ldots,\lambda_nq_n)$

$Aq_1=\lambda_1q_1, Aq_2=\lambda_2q_2, \ldots, Aq_\mathrm{n}=\lambda_\mathrm{n}q_\mathrm{n}$

（1）正规阵 $A=A_{n\times n}$ 有 n 个互相正交的特征向量 $q_1, q_2, \ldots, q_n$

注： $x\perp y$（正交）$\Leftrightarrow y^Hx=0$ 或 $x^Hy=0$

$q_1\perp q_2\perp\cdots\perp q_n\Leftrightarrow q_k^Hq_l=0$（$k\neq l$）

$Q=(q_1, q_2, \ldots, q_n)$， $Q^H=\begin{pmatrix}q_1^H\\ q_2^H\\ \vdots\\ q_n^H\end{pmatrix}$，（$QQ^H=Q^HQ=I_n$）

$$\Rightarrow QQ^H=(q_1,q_2,\cdots,q_n)\begin{pmatrix}q_1^H\\ q_2^H\\ \vdots\\ q_n^H\end{pmatrix}=\sum_{i=1}^n q_iq_i^H=I_n$$

令 $Q_1=q_1q_1^H, Q_2=q_2q_2^H,\cdots,Q_n=q_nq_n^H$ 都是 Hermite 阵

$Q_1^H=Q_1, Q_2^H=Q_2,\cdots,Q_n^H=Q_n$； 且 $Q_1^2=Q_1, Q_2^2=Q_2,\cdots,Q_n^2=Q_n$

$$Q_1^2=Q_1Q_1=(q_1q_1^H)(q_1q_1^H)=q_1(q_1^Hq_1)q_1^H=q_1q_1^H=Q_1$$

（2）$A=A_{n\times n}$ 为正规阵，则有分解公式： $A=\sum_{i=1}^n\lambda_iQ_i$ （谱分解）

且 $Q_1^2=Q_1=Q_1^H, Q_2^2=Q_2=Q_2^H,\cdots,Q_n^2=Q_n=Q_n^H$， $Q_1+Q_2+\ldots+Q_n=I_n$

Pf：$\because A = Q\begin{pmatrix} \lambda_1 & & & O \\ & \lambda_2 & & \\ & & \ddots & \\ O & & & \lambda_n \end{pmatrix} Q^H = (q_1, q_2, \cdots, q_n)\begin{pmatrix} \lambda_1 & & & O \\ & \lambda_2 & & \\ & & \ddots & \\ O & & & \lambda_n \end{pmatrix}\begin{pmatrix} q_1^H \\ q_2^H \\ \vdots \\ q_n^H \end{pmatrix}$

$$\Rightarrow A = \sum_{i=1}^{n} \lambda_i \left(q_i q_i^H\right) = \sum_{i=1}^{n} \lambda_i Q_i$$

注：分解中的 $Q_1, Q_2, \ldots, Q_n$ 叫投影阵

性质：（1）$Q_1 + Q_2 + \ldots + Q_n = I_n$

（2）$Q_1^2 = Q_1 = Q_1^H, Q_2^2 = Q_2 = Q_2^H, \cdots, Q_n^2 = Q_n = Q_n^H$

（3）$Q_1Q_2 = 0, \ldots, Q_kQ_l = 0$（$k \neq l$）

$\because Q_1Q_2 = \left(q_1q_1^H\right)\left(q_2q_2^H\right) = q_1\left(q_1^Hq_2\right)q_2^H = 0$，（$q_1 \perp q_2$）

（4）$AQ_1 = \lambda_1 Q_1, AQ_2 = \lambda_2 Q_2, \ldots, AQ_n = \lambda_n Q_n$

（5）$A^k = \sum_{i=1}^{n} \lambda_i^k Q_i$

（6）$f(A) = \sum_{i=1}^{n} f(\lambda_i) Q_i$，（$f(x)$为多项式）

Pf：$\because A = \sum_{i=1}^{n} \lambda_i Q_i$

$$\Rightarrow AQ_1 = \left(\sum_{i=1}^{n} \lambda_i Q_i\right)Q_1 = \lambda_1 Q_1^2 + \lambda_2 Q_2 Q_1 + \cdots + \lambda_n Q_n Q_1 = \lambda_1 Q_1^2 + 0 + \cdots + 0$$

$AQ_1 = \lambda_1 Q_1^2 = \lambda_1 Q_1$，同理 $AQ_2 = \lambda_2 Q_2$

Pf：（5）若 $A^k = \sum_{i=1}^{n} \lambda_i^k Q_i$ （归纳法）

$$\Rightarrow A^{k+1} = A \bullet A^k = A\left(\sum_{i=1}^{n} \lambda_i^k Q_i\right) = \sum_{i=1}^{n} \lambda_i^k \left(AQ_i\right) = \sum_{i=1}^{n} \lambda_i^{k+1} Q_i$$

Pf：（6）写 $f(x) = a_0 + a_1x + \ldots + a_mx^m$

$$f(A) = a_0 I_n + a_1 A + \cdots + a_m A^m = a_0\left(\sum_{i=1}^{n} Q_i\right) + a_1\left(\sum_{i=1}^{n} \lambda_i Q_i\right) + \cdots + a_m\left(\sum_{i=1}^{n} \lambda_i^m Q_i\right)$$

$$=\left(a_0+a_1\lambda_1+\cdots+a_m\lambda_1^m\right)Q_1+\left(a_0+a_1\lambda_2+\cdots+a_m\lambda_2^m\right)Q_2+\cdots+\left(a_0+a_1\lambda_n+\cdots+a_m\lambda_n^m\right)Q_2$$

$$=\sum_{i=1}^{n}f\left(\lambda_i\right)Q_i$$

注： $\lambda_1,\lambda_2,\ldots,\lambda_n$ 有重根时，可合并部分 $Q_1, Q_2, \ldots, Q_n$

例如：$\lambda_1=\lambda_2$ 时：$\lambda_1Q_1+\lambda_2Q_2=\lambda_1(Q_1+Q_2)$

写 $G_1=Q_1+Q_2$，且 $G_1^H=G_1=G_1^2=\left(Q_1+Q_2\right)^2=Q_1+Q_2$

正规谱分解公式： 设 A 为正规阵，$\lambda_1,\lambda_2,\ldots,\lambda_s$ 为互异特征值，则存在 $G_1,G_2,\ldots,G_s$ 使得

（1）$A=\sum_{i=1}^{s}\lambda_iG_i$（注：$G_1,G_2,\ldots,G_s$ 由 $Q_1,Q_2,\ldots,Q_n$ 合并）

（2）$G_1+G_2+\ldots+G_s=I_n$

（3）$G_1^2=G_1=G_1^H,G_2^2=G_2=G_2^H,\cdots,G_s^2=G_s=G_s^H$

（4）$G_1G_2=0,\ldots,G_kG_l=0$，（$k\neq l$）

（5）$AG_1=\lambda_1G_1, AG_2=\lambda_2G_2,\ldots,AG_s=\lambda_sG_s$

（6）$A^k=\sum_{i=1}^{s}\lambda_i^kG_i$

（7）$f\left(A\right)=\sum_{i=1}^{s}f\left(\lambda_i\right)G_i$

$k=0$ 时，$A^0=I_n=G_1+G_2+\ldots+G_s$

$k=1$ 时，$A^1=\sum_{i=1}^{s}\lambda_iG_i$，（$G_1G_2=0,\ldots,G_sG_{s-1}=0$）

注： 其中 $G_1,G_2,\ldots,G_s$ 叫 A 的投影阵

引入 $g\left(x\right)=\prod_{i=1}^{s}\left(x-\lambda_i\right)$，（$\lambda_1,\lambda_2,\ldots,\lambda_s$ 互异）

$$g_1(x)=(x-\hat{\lambda}_1)(x-\lambda_2)\cdots(x-\lambda_s)\quad (\text{去掉}(x-\lambda_1))$$
$$g_2(x)=(x-\lambda_1)(x-\hat{\lambda}_2)\cdots(x-\lambda_s)\quad (\text{去掉}(x-\lambda_2))$$
$$\vdots$$
$$g_s(x)=(x-\lambda_1)(x-\lambda_2)\cdots(x-\hat{\lambda}_s)\quad (\text{去掉}(x-\lambda_s))$$

则 $g_1(\lambda_1)\neq 0, g_2(\lambda_2)\neq 0, \ldots, g_s(\lambda_s)\neq 0$

令 $G_1=\varphi_1(A)=\dfrac{g_1(A)}{g_1(\lambda_1)}, G_2=\varphi_2(A)=\dfrac{g_2(A)}{g_2(\lambda_2)},\cdots,G_s=\varphi_s(A)=\dfrac{g_s(A)}{g_s(\lambda_s)}$

则 $A=\sum\limits_{i=1}^{s}\lambda_i G_i$； $A^k=\sum\limits_{i=1}^{s}\lambda_i^k G_i$

Pf：由公式 $f(A)=\sum\limits_{i=1}^{s}f(\lambda_i)G_i$，（$f(x)$为任取）

取 $f(x)=g_1(x)\Rightarrow g_1(A)=\sum\limits_{i=1}^{s}g_1(\lambda_i)G_i\Rightarrow g_1(A)=g_1(\lambda_1)G_1\Rightarrow G_1=\dfrac{g_1(A)}{g_1(\lambda_1)}=\varphi_1(A)$

同理：取 $f(x)=g_2(x)\Rightarrow G_2=\dfrac{g_2(A)}{g_2(\lambda_2)}=\varphi_2(A)$

取 $f(x)=g_s(x)\Rightarrow G_s=\dfrac{g_s(A)}{g_s(\lambda_s)}=\varphi_s(A)$

Eg. $A=\begin{pmatrix}2 & 1-i\\ 1+i & 1\end{pmatrix}$，（$i=\sqrt{-1}$，$i^2=-1$），$A^H=A$（正规）

解：$|xI-A|=(x-3)x$，$\sigma(A)=\{3,0\}$，$\lambda_1=3$，$\lambda_2=0$

$\lambda_1=3$：特征向量 $q_1=\dfrac{1}{\sqrt{3}}\begin{pmatrix}1-i\\ 1\end{pmatrix}$；$\lambda_2=0$：特征向量 $q_2=\dfrac{1}{\sqrt{3}}\begin{pmatrix}-1\\ 1+i\end{pmatrix}$

令 $Q=(q_1,q_2)=\dfrac{1}{\sqrt{3}}\begin{pmatrix}1-i & -1\\ 1 & 1+i\end{pmatrix}$，$Q$ 为 $\mathscr{U}$阵（$Q^HQ=I$）

$$\Rightarrow Q^HAQ=\begin{pmatrix}\frac{1}{\sqrt{3}}(1+i) & \frac{1}{\sqrt{3}}\\ -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}}(1-i)\end{pmatrix}\begin{pmatrix}2 & 1-i\\ 1+i & 1\end{pmatrix}\begin{pmatrix}\frac{1}{\sqrt{3}}(1-i) & -\frac{1}{\sqrt{3}}\\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}}(1+i)\end{pmatrix}=\begin{pmatrix}3 & 0\\ 0 & 0\end{pmatrix}$$

$$\Rightarrow A=\lambda_1q_1q_1^H+\lambda_2q_2q_2^H=3\cdot G_1+0\cdot G_2$$

方法 2：用投影阵公式：$g(x)=(\mathrm{x}-\lambda_1)(\mathrm{x}-\lambda_2)=(\mathrm{x}-3)(\mathrm{x}-0)$

$g_1(x)=(x-\hat{\lambda_1})(x-\lambda_2)=x$，$g_2(x)=(x-\lambda_1)(x-\hat{\lambda_2})=(x-3)$

$G_1=\dfrac{g_1(A)}{g_1(\lambda_1)}=\dfrac{A}{3}$，$G_2=\dfrac{g_2(A)}{g_2(\lambda_2)}=\dfrac{A-3I}{0-3}$

$$A=\lambda_1G_1+\lambda_2G_2=\lambda_1\left(\frac{A}{3}\right)+\lambda_2\left(\frac{A-3I}{-3}\right)\Rightarrow A^{100}=3^{100}\left(\frac{A}{3}\right)+0^{100}\left(\frac{A-3I}{-3}\right)$$

Ex.判定下列矩阵为正规，并写出谱分解公式（$f(A)=?$）

（1）$A=\begin{pmatrix}1&-1\\1&1\end{pmatrix}$　（2）$A=\begin{pmatrix}1&1-i\\1+i&0\end{pmatrix}$　（3）$A=\begin{pmatrix}1&-i\\i&1\end{pmatrix}$

（4）$A=\begin{pmatrix}1&-i&0\\i&1&0\\0&0&1\end{pmatrix}$　（5）$A=\begin{pmatrix}0&0&1\\0&1&0\\1&0&0\end{pmatrix}$　（6）$A=\begin{pmatrix}1&1&1\\1&1&1\\1&1&1\end{pmatrix}$

《矩阵分析》P179-180：2 5（2）（4）　3（1）　8（仿例 3.6.5）

Eg.证明：若 $A^H=A$（Hermite），则 A 的特征值全为实数

Pf：∵$A^H=A$ 为正规阵

$$\Rightarrow Q^HAQ=\begin{pmatrix}\lambda_1&&&O\\&\lambda_2&&\\&&\ddots&\\O&&&\lambda_n\end{pmatrix}，（Q 为 \mathscr{U} 阵）$$

$$\left(Q^HAQ\right)^H=\begin{pmatrix}\lambda_1&&&O\\&\lambda_2&&\\&&\ddots&\\O&&&\lambda_n\end{pmatrix}^H=\begin{pmatrix}\overline{\lambda_1}&&&O\\&\overline{\lambda_2}&&\\&&\ddots&\\O&&&\overline{\lambda_n}\end{pmatrix}$$

左边：$Q^HA^HQ=Q^HAQ=\begin{pmatrix}\lambda_1&&&O\\&\lambda_2&&\\&&\ddots&\\O&&&\lambda_n\end{pmatrix}\Rightarrow\overline{\lambda_1}=\lambda_1,\overline{\lambda_2}=\lambda_2,\cdots,\overline{\lambda_n}=\lambda_n$

∴$\lambda_1,\lambda_2,\ldots,\lambda_n$ 为实数

Ex.斜 Hermite 阵 $A=-A^H$ 的特征值全为纯虚或 0（可用 iA 为 Hermite 阵）

正规阵应用（正规条件 $A^HA=AA^H$）

注： 实对称阵：Hermite 阵、正交阵、$\mathscr{U}$阵都为正规阵

正规分解：A 正规$\Leftrightarrow Q^H AQ=\begin{pmatrix}\lambda_1 & & & O\\ & \lambda_2 & & \\ & & \ddots & \\ O & & & \lambda_n\end{pmatrix}$，（$Q$ 为某 $\mathscr{U}$阵）

正规谱分解：A 正规$\Leftrightarrow A=\lambda_1 G_1+\lambda_2 G_2+\cdots+\lambda_s G_s$；（$\lambda_1,\lambda_2,\ldots,\lambda_s$ 互异）

$$A^k=\lambda_1^k G_1+\lambda_2^k G_2+\cdots+\lambda_s^k G_s$$

$$f(A)=f(\lambda_1)G_1+f(\lambda_2)G_2+\ldots+f(\lambda_s)G_s\text{；（}f(x)\text{为任意多项式）}$$

其中：$G_1, G_2, \ldots, G_s$ 叫 A 的投影阵

注：（1）$G_1+G_2+\ldots+G_s=I_n$

（2）$G_1^2=G_1=G_1^H, G_2^2=G_2=G_2^H,\cdots,G_s^2=G_s=G_s^H$

（3）$G_iG_j=0$；（$i\neq j$）

特别有投影公式如下：$G_1=\dfrac{g_1(A)}{g_1(\lambda_1)}, G_2=\dfrac{g_2(A)}{g_2(\lambda_2)},\cdots,G_s=\dfrac{g_s(A)}{g_s(\lambda_s)}$

其中：
$$\begin{aligned}g_1(x)&=(x-\hat{\lambda_1})(x-\lambda_2)\cdots(x-\lambda_s)\\ g_2(x)&=(x-\lambda_1)(x-\hat{\lambda_2})\cdots(x-\lambda_s)\\ &\vdots\\ g_s(x)&=(x-\lambda_1)(x-\lambda_2)\cdots(x-\hat{\lambda_s})\end{aligned}$$

eg.任取 $A=A_{m\times n}\in C^{m\times n}$ 则 A^HA、AA^H 都为 Hermite（正规）

pf：$(A^H A)^H=A^H(A^H)^H=A^H A\Rightarrow A^H A$ 为 Hermite

引理：（1）$y^Hy=|y_1|^2+|y_2|^2+\ldots+|y_n|^2\geq 0$；$y=(y_1, y_2, \ldots, y_n)^T\in C^n$

（2）$y^H y=0\Leftrightarrow y=0$

由于 $x^H(A^HA)x=(Ax)^H(Ax)\geq 0$（引理），令 $y=Ax$

$\Rightarrow$2 次型 $x^H(A^HA)x$ 为半正定$\Rightarrow A^H A$ 的特征值非负。

可设$Q^H(A^H A)Q=\begin{pmatrix}\lambda_1 & & & O\\ & \lambda_2 & & \\ & & \ddots & \\ O & & & \lambda_n\end{pmatrix}$，（$Q$ 为 $\mathscr{U}$阵）

$$x^H Q^H \left(A^H A\right) x Q = (Qx)^H \left(A^H A\right)(Qx) = x^H \begin{pmatrix} \lambda_1 & & & O \\ & \lambda_2 & & \\ & & \ddots & \\ O & & & \lambda_n \end{pmatrix} x$$

$$= \lambda_1 |x_1|^2 + \lambda_2 |x_2|^2 + \cdots + \lambda_n |x_n|^2 \geq 0 \Rightarrow \lambda_1 \geq 0, \lambda_2 \geq 0, \cdots, \lambda_n \geq 0$$

定义： 设 $A^H A$ 的特征值为 $\lambda_1 \geq 0, \lambda_2 \geq 0, \ldots, \lambda_n \geq 0$，称 $\sqrt{\lambda_1}, \sqrt{\lambda_2}, \cdots, \sqrt{\lambda_n}$ 为 A 的奇异值

eg.任一半正定（或正定）Hermite 阵 A 存在平方根 B

使得 $B^2 = A$（可写 $B = A^{1/2}$）且 B 为半正定（正定）

Pf：证法 1：$\because Q^H A Q = \begin{pmatrix} \lambda_1 & & & O \\ & \lambda_2 & & \\ & & \ddots & \\ O & & & \lambda_n \end{pmatrix}$，$\lambda_1 \geq 0, \lambda_2 \geq 0, \ldots, \lambda_n \geq 0$

$$\Rightarrow AQ = Q \begin{pmatrix} \sqrt{\lambda_1} & & & O \\ & \sqrt{\lambda_2} & & \\ & & \ddots & \\ O & & & \sqrt{\lambda_n} \end{pmatrix} \begin{pmatrix} \sqrt{\lambda_1} & & & O \\ & \sqrt{\lambda_2} & & \\ & & \ddots & \\ O & & & \sqrt{\lambda_n} \end{pmatrix} Q^H$$

$$= Q \begin{pmatrix} \sqrt{\lambda_1} & & & O \\ & \sqrt{\lambda_2} & & \\ & & \ddots & \\ O & & & \sqrt{\lambda_n} \end{pmatrix} Q^H Q \begin{pmatrix} \sqrt{\lambda_1} & & & O \\ & \sqrt{\lambda_2} & & \\ & & \ddots & \\ O & & & \sqrt{\lambda_n} \end{pmatrix} Q^H$$

令 $B = Q \begin{pmatrix} \sqrt{\lambda_1} & & & O \\ & \sqrt{\lambda_2} & & \\ & & \ddots & \\ O & & & \sqrt{\lambda_n} \end{pmatrix} Q^H$ 且 B 为 Hermite 半正定

$\Rightarrow A = B^2$ 或 $B = A^{1/2}$

证法 2：$\because A = \lambda_1 G_1 + \lambda_2 G_2 + \ldots + \lambda_s G_s$；（$\lambda_1 \geq 0, \lambda_2 \geq 0, \ldots, \lambda_s \geq 0$ 互异）

令 $B = \sqrt{\lambda_1} G_1 + \sqrt{\lambda_2} G_2 + \cdots + \sqrt{\lambda_s} G_s$

$$\Rightarrow B = \left(\sqrt{\lambda_1}\right)^2 G_1^2 + \left(\sqrt{\lambda_2}\right)^2 G_2^2 + \cdots + \left(\sqrt{\lambda_s}\right)^2 G_s^2 = \lambda_1 G_1 + \lambda_2 G_2 + \cdots + \lambda_s G_s = A$$

$\because G_1, G_2, \ldots, G_s$ 为 A 的多项式 $\Rightarrow B = \sqrt{\lambda_1} G_1 + \sqrt{\lambda_2} G_2 + \cdots + \sqrt{\lambda_s} G_s$ 为 A 的多项式

可对角化矩阵的谱分解

注： 方阵 $A = A_{n\times n}$ 可对角化条件：A 相似于对角阵

引理：A 相似于对角阵 $\Leftrightarrow$ A 有 n 个无关的特征向量 $\Leftrightarrow$ A 的极小式无重根

谱分解公式：设 A 可对角化（极小式无重根）

令 A 的全体互异特征值为 $\lambda_1,\lambda_2,\ldots,\lambda_s$

令 $G_1=\dfrac{g_1(A)}{g_1(\lambda_1)}, G_2=\dfrac{g_2(A)}{g_2(\lambda_2)},\cdots,G_s=\dfrac{g_s(A)}{g_s(\lambda_s)}$

其中 $\begin{cases} g_1(x)=(x\hat{-}\lambda_1)(x-\lambda_2)\cdots(x-\lambda_s) \\ g_2(x)=(x-\lambda_1)(x\hat{-}\lambda_2)\cdots(x-\lambda_s) \\ \vdots \\ g_s(x)=(x-\lambda_1)(x-\lambda_2)\cdots(x\hat{-}\lambda_s) \end{cases}$；$g_1(\lambda_1)\neq 0, g_2(\lambda_2)\neq 0, \ldots, g_s(\lambda_s)\neq 0$

则 $A=\lambda_1G_1+\lambda_2G_2+\ldots+\lambda_sG_s$

$A^k=\lambda_1^kG_1+\lambda_2^kG_2+\cdots+\lambda_s^kG_s$

$f(A)=f(\lambda_1)G_1+f(\lambda_2)G_2+\ldots+f(\lambda_s)G_s$；（$f(x)$为任意多项式）

其中 $G_1, G_2, \ldots, G_s$ 叫投影阵

性质：（1）$G_1+G_2+\ldots+G_s=I_n$

（2）$G_1^2=G_1, G_2^2=G_2,\cdots,G_s^2=G_s$

（3）$G_iG_j=0$；（$i\neq j$）

（4）$AG_1=\lambda_1G_1, AG_2=\lambda_2G_2, \ldots, AG_s=\lambda_sG_s$

eg. $A=\begin{pmatrix}1&1&0\\0&2&0\\0&0&3\end{pmatrix}$，求谱分解

解：A 有 3 个互异特征值 1、2、3（可对角化）

$$\text{令}\ \begin{aligned} g_1(x)&=(x\hat{-}1)(x-2)(x-3)=(x-2)(x-3) \\ g_2(x)&=(x-1)(x\hat{-}2)(x-3)=(x-1)(x-3) \\ g_3(x)&=(x-1)(x-2)(x\hat{-}3)=(x-1)(x-2) \end{aligned}$$

$$令\begin{cases} G_1=\dfrac{g_1(A)}{g_1(\lambda_1)}=\dfrac{(A-2I)(A-3I)}{(1-2)(1-3)} \\ G_2=\dfrac{g_2(A)}{g_2(\lambda_2)}=\dfrac{(A-I)(A-3I)}{(2-1)(2-3)} \\ G_3=\dfrac{g_3(A)}{g_3(\lambda_3)}=\dfrac{(A-I)(A-2I)}{(3-1)(3-2)} \end{cases}\Rightarrow A=1\cdot G_1+2\cdot G_2+3\cdot G_3$$

谱分解习题：

1. $A=\begin{pmatrix} 0 & -i & 1 \\ i & 0 & i \\ 1 & -i & 0 \end{pmatrix}$，求谱分解（$i^2=-1$）

2.若 A 正规，且 $A = \lambda_1 G_1 + \lambda_2 G_2 + \ldots + \lambda_s G_s$ 为其谱分解，则 A^H 的谱分解为 $A^H=\overline{\lambda_1}G_1+\overline{\lambda_2}G_2+\cdots+\overline{\lambda_s}G_s$

3. 若 A 有谱分解：$A = \lambda_1 G_1 + \lambda_2 G_2 + \ldots + \lambda_s G_s$ 则 A^T 有谱分解 $A^T=\lambda_1 G_1^T+\lambda_2 G_2^T+\cdots+\lambda_s G_s^T$

另有，《矩阵分析》P212，3、4

条件：A 可对角化（极小式 $g_1(x) = (x-\lambda_1)(x-\lambda_2)\ldots(x-\lambda_s)$无重根）

则 $A = \lambda_1 G_1 + \lambda_2 G_2 + \ldots + \lambda_s G_s$，且 $f(A) = f(\lambda_1)G_1 + f(\lambda_2)G_2 + \ldots + f(\lambda_s)G_s$，（$f(x)$为任意多项式）

其中 $G_1=\dfrac{g_1(A)}{g_1(\lambda_1)},G_2=\dfrac{g_2(A)}{g_2(\lambda_2)},\cdots,G_s=\dfrac{g_s(A)}{g_s(\lambda_s)}$

（1）$G_1 + G_2 + \ldots + G_s = I_n$

（2）$G_1^2=G_1,G_2^2=G_2,\cdots,G_s^2=G_s$

（3）$G_iG_j = 0$；（$i\neq j$）

（4）$AG_1 =\lambda_1 G_1, AG_2 =\lambda_2 G_2, \ldots, AG_s =\lambda_s G_s$

Pf：1：$A=P\begin{pmatrix} \lambda_1 & & O \\ & \lambda_2 & \\ & & \ddots & \\ O & & & \lambda_n \end{pmatrix}P^{-1}$，令 $P = (p_1, p_2, \ldots, p_n)$，$P^{-1}=\begin{pmatrix} q_1 \\ q_2 \\ \vdots \\ q_n \end{pmatrix}$（按行）

$$\Rightarrow A=(p_1,p_2,\cdots,p_n)\begin{pmatrix}\lambda_1 & & & O\\ & \lambda_2 & & \\ & & \ddots & \\ O & & & \lambda_n\end{pmatrix}\begin{pmatrix}q_1\\ q_2\\ \vdots\\ q_n\end{pmatrix}=\lambda_1(p_1q_1)+\lambda_2(p_2q_2)+\cdots+\lambda_s(p_sq_s)$$

注： $I=P^{-1}P=\begin{pmatrix}q_1\\ q_2\\ \vdots\\ q_n\end{pmatrix}(p_1,p_2,\cdots,p_n)=\begin{pmatrix}q_1p_1 & & & *\\ & q_2p_2 & & \\ & & \ddots & \\ * & & & q_np_n\end{pmatrix}\Rightarrow\begin{matrix}q_1p_1=\cdots=q_np_n=1\\ \text{其它}q_ip_j=0(i\neq j)\end{matrix}$

$\Rightarrow$令 $P_1=p_1q_1, P_2=p_2q_2, \ldots,P_n=p_nq_n$,

则 $A=\lambda_1P_1+\lambda_2P_2+\ldots+\lambda_nP_n$

且 $P_1^2=(p_1q_1)(p_1q_1)=p_1(q_1p_1)q_1=p_1q_1=P_1$，同理 $P_2^2=P_2,\cdots,P_n^2=P_n$

$P_1P_2=(p_1q_1)(p_2q_2)=p_1(q_1p_2)q_2=0$，同理 $P_iP_j=0$（$i\neq j$）

$$P_1+P_2+\cdots+P_n=p_1q_1+p_2q_2+\cdots+p_nq_n=(p_1,p_2,\cdots,p_n)\begin{pmatrix}q_1\\ q_2\\ \vdots\\ q_n\end{pmatrix}=PP^{-1}=I_n$$

注： 若 $\lambda_1=\lambda_2$（重根）$\lambda_1P_1+\lambda_2P_2=\lambda_1(P_1+P_2)\underline{\underline{\text{记为}}}\lambda_1G_1$

且 $G_1=(P_1+P_2)$，$G_1^2=(P_1+P_2)^2=P_1^2+P_2^2=P_1+P_2=G_1$

合并重根$\Rightarrow A=\lambda_1G_1+\lambda_2G_2+\cdots+\lambda_sG_s$（$\lambda_1,\lambda_2,\ldots,\lambda_s$互异）

Pf：2：若 A 的极小式 $g_1(x)=(x-\lambda_1)(x-\lambda_2)\ldots(x-\lambda_s)$无重根

$$\text{令}\begin{cases}g_1(x)=(x\hat{-}\lambda_1)(x-\lambda_2)\cdots(x-\lambda_s)\\ g_2(x)=(x-\lambda_1)(x\hat{-}\lambda_2)\cdots(x-\lambda_s)\\ \quad\vdots\\ g_s(x)=(x-\lambda_1)(x-\lambda_2)\cdots(x\hat{-}\lambda_s)\end{cases}$$

$$\text{令}\ \tilde{g}_1=\frac{g_1(x)}{g_1(\lambda_1)},\tilde{g}_2=\frac{g_2(x)}{g_2(\lambda_2)},\cdots,\tilde{g}_s=\frac{g_s(x)}{g_s(\lambda_s)}$$

则有 0 点公式：$\tilde{g}_1+\tilde{g}_2+\cdots+\tilde{g}_s=1$；$\lambda_1\tilde{g}_1+\lambda_2\tilde{g}_2+\cdots+\lambda_s\tilde{g}_s=x$

$$\text{取}\ x=A\Rightarrow\begin{cases}\tilde{g}_1(A)+\tilde{g}_2(A)+\cdots+\tilde{g}_s(A)=I\\ \lambda_1\tilde{g}_1(A)+\lambda_2\tilde{g}_2(A)+\cdots+\lambda_s\tilde{g}_s(A)=A\end{cases}$$

令$G_1=\widetilde{g}_1(A)=\dfrac{g_1(A)}{g_1(\lambda_1)}, G_2=\widetilde{g}_2(A)=\dfrac{g_2(A)}{g_2(\lambda_2)},\cdots,G_s=\widetilde{g}_s(A)=\dfrac{g_s(A)}{g_s(\lambda_s)}$

则有 $G_1+G_2+\ldots+G_s=I_n$；$A=\lambda_1G_1+\lambda_2G_2+\ldots+\lambda_sG_s$

$\because g_1(x)\,g_2(x)=(x-\lambda_1)\,(x-\lambda_2)\ldots(x-\lambda_s)(\ldots)=g(x)(\ldots)$

$\Rightarrow g_1(A)g_2(A)=g(A)(\cdots)=0(\cdots)=0$

同理 $g_i(A)g_j(A)=0$（$i\neq j$）$\Rightarrow G_1G_2=\dfrac{g_1(A)g_2(A)}{g_1(\lambda_1)g_2(\lambda_2)}=0$

又 $I=G_1+G_2+\ldots+G_s$

$\Rightarrow G_1=G_1I=G_1(G_1+G_2+\cdots+G_s)=G_1^2+G_1G_2+\cdots+G_1G_s=G_1^2$

$\Rightarrow G_1^2=G_1,G_2^2=G_2,\cdots,G_s^2=G_s$

注：（1）极小式无重根$\Leftrightarrow A$可对角化

（2）某 0 化式无重根$\Rightarrow A$可对角化

eg. $A=\begin{pmatrix}2&2&1\\1&3&1\\1&2&2\end{pmatrix}$求极小式与谱分解

解：$|\lambda I-A|\underline{\underline{\text{各列相加}}}(x-5)(x-1)^2$，极小式为$(x-5)(x-1)$或$(x-5)(x-1)^2$

计算$(A-5I)(A-I)=0\Rightarrow$极小式 $g(x)=(x-5)(x-1)$

令 $g_1(x)=(x-5)$，$g_2(x)=(x-1)$

$\Rightarrow A=1\cdot\left(\dfrac{g_1(A)}{1-5}\right)+5\cdot\left(\dfrac{g_2(A)}{5-1}\right)$

Ex.令 $B=A^T=\begin{pmatrix}2&1&1\\2&3&2\\1&1&2\end{pmatrix}$求 B 的极小式与谱分解

换位公式： 令 $A=A_{m\times n}$，$B=B_{n\times m}$

则$|xI_m-AB|=x^{m-n}|xI_n-BA|$

Pf：令 $P=\begin{pmatrix}I_m&A\\0&I_n\end{pmatrix}$，$Q=\begin{pmatrix}AB&0\\B&0\end{pmatrix}$，$\widetilde{Q}=\begin{pmatrix}0&0\\B&AB\end{pmatrix}$

$$\Rightarrow P\widetilde{Q}=\begin{pmatrix} I_m & A \\ 0 & I_n \end{pmatrix}\begin{pmatrix} 0 & 0 \\ B & AB \end{pmatrix}=\begin{pmatrix} AB & ABA \\ B & BA \end{pmatrix}$$

$$QP=\begin{pmatrix} AB & 0 \\ B & 0 \end{pmatrix}\begin{pmatrix} I_m & A \\ 0 & I_n \end{pmatrix}=\begin{pmatrix} AB & ABA \\ B & BA \end{pmatrix}$$

$$\Rightarrow QP=P\widetilde{Q}\Rightarrow P^{-1}QP=\widetilde{Q}\Rightarrow Q\backsim\widetilde{Q}$$

$$\Rightarrow |xI-Q|=|xI-\widetilde{Q}|$$

$$\Rightarrow \begin{vmatrix} (xI_m-AB) & 0 \\ -B & xI_n \end{vmatrix}=\begin{vmatrix} xI_m & 0 \\ -B & xI_n-BA \end{vmatrix}$$

$$\Rightarrow |xI_m-AB|x^n=x^m|xI_n-BA|$$

$$\Rightarrow |xI_m-AB|=x^{m-n}|xI_n-BA|$$

常用换位公式：设 $A=A_{m\times n}$，$B=B_{n\times m}$

（1）$|xI_m-AB|=x^{m-n}|xI_n-BA|$

（2）$|I_m-AB|=|I_n-BA|$；用$(\pm kA)$代替 A 得$|I_m\pm kAB|=|I_n\pm kBA|$

（3）迹换位：$\mathrm{tr}(AB)=\mathrm{tr}(BA)$；$\begin{cases} X=\begin{pmatrix} x_{11} & & * \\ & x_{22} & \ddots \\ * & & \\ & & x_{nn} \end{pmatrix} \\ tr(X)=x_{11}+x_{22}+\cdots+x_{nn} \end{cases}$

Pf:（3）：写$|xI-AB|=(x-\lambda_1)(x-\lambda_2)\ldots(x-\lambda_n)$， $\lambda_1,\lambda_2,\ldots,\lambda_n$ 为 BA 的特征值

$$\Rightarrow (xI_m-AB)=x^{m-n}(x-\lambda_1)(x-\lambda_2)\cdots(x-\lambda_n)$$

$$\Rightarrow \{0,0,\cdots,0,\lambda_1,\lambda_2,\cdots,\lambda_n\}\text{为 }AB\text{ 的特征值}$$

$$\Rightarrow tr(AB)=0+0+\cdots+0+\lambda_1+\lambda_2,\cdots+\lambda_n=\lambda_1+\lambda_2,\cdots+\lambda_n=tr(BA)$$

eg. $\alpha=(a_1,a_2,\ldots,a_n)^T,\beta=(b_1,b_2,\ldots,b_n)^T$

$$tr(\alpha\beta^T)=tr(\beta^T\alpha)=tr\left((b_1,b_2,\cdots,b_n)\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}\right)=a_1b_1+a_2b_2+\cdots+a_nb_n$$

eg. $A=\alpha\beta^T=\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix}(b_1,b_2,\cdots,b_n)=\begin{pmatrix}a_1b_1&a_1b_2&\cdots&a_1b_n\\a_2b_1&a_2b_2&\cdots&a_2b_n\\\vdots&\vdots&\ddots&\vdots\\a_nb_1&a_nb_2&\cdots&a_nb_n\end{pmatrix}$

$|I_n-\alpha\beta^T|=|I_1-(\beta^T\alpha)|=1-\beta^T\alpha=1-(a_1b_1+a_2b_2++a_nb_n)$

$|xI_n-\alpha\beta^T|=x^{n-1}|(xI_1)-\beta^T\alpha|=x^{n-1}(x-\beta^T\alpha)=x^{n-1}(x-\mathrm{tr}(A))$

$\Rightarrow$谱 $\sigma(A)=\{0, 0, \ldots, 0, \mathrm{tr}(a)\}$

eg.镜面阵 $A=I_n-2\varepsilon\varepsilon^T$，（$|\varepsilon|^2=1=\varepsilon^T\varepsilon$），$\varepsilon=(a_1, a_2, \ldots, a_n)\in R^n$

$|xI_n-A|=|(x-1)I_n+2\varepsilon\varepsilon^T|=(x-1)^{n-1}|(x-1)I_1+2\varepsilon^T\varepsilon|=(x-1)^{n-1}(x+1)$

$\sigma(A)=\{1, 1, \ldots, 1, -1\}$

同理：$|I_n-2\varepsilon\varepsilon^T|=|I_1-2\varepsilon^T\varepsilon|=(1-2)=-1$

迹公式：（1）$tr(AB^T)=tr(B^TA)=\sum(a_{ij}b_{ij})$

（2）$tr(AB^H)=tr(B^HA)=\sum(a_{ij}\overline{b_{ij}})$

Pf：$A=\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&\ddots&\vdots\\a_{m1}&a_{m2}&\cdots&a_{mn}\end{pmatrix}$，$B^T=\begin{pmatrix}b_{11}&b_{21}&\cdots&b_{m1}\\b_{12}&b_{22}&\cdots&b_{m2}\\\vdots&\vdots&\ddots&\vdots\\b_{1n}&b_{2n}&\cdots&b_{mn}\end{pmatrix}$

$$\Rightarrow AB^T=\begin{pmatrix}\sum_{i=1}^n a_{1i}b_{1i}&&*\\&\ddots&\\*&&\sum_{i=1}^n a_{mi}b_{mi}\end{pmatrix}$$

$tr(AB^T)=\sum(a_{ij}b_{ij})$

（3）$tr(AA^H)=tr(A^HA)=\sum|a_{ij}|^2\ge 0$

迹公式应用在内积定义中：

C^n中标准内积为$(x,y)=tr(x\bullet y^H)=tr(y^Hx)=\sum_{k=1}^n x_k\overline{y_k}$

$C^{m\times n}$中标准内积为$(A,B)=tr(A\bullet B^H)=\sum(a_{ij}\overline{b_{ij}})$

推论：若 $A=A_{m\times n}$ 则 AA^H 与 A^HA 的正特征值相同

$\because |xI_m-AA^H|=x^{m-n}|xI_n-A^HA|$

正交补与正交子空间

定义： 设$W_1,W_2 \subset C^n\left(R^n\right)$，若$\forall x \in W_1, y \in W_2$，有 $x \perp y$，即$(x, y) = 0$

称 W_1 与 W_2 正交，记为 $W_1 \perp W_2$

定义： $x \perp W \Leftrightarrow x \perp y\left(\forall y \in W\right)$

引理： $W_1 \perp W_2 \Rightarrow W_1 \cap W_2 = \left\{\vec{0}\right\} \Rightarrow W_1 + W_2 = W_1 \oplus W_2$（直和）

$$\Leftrightarrow \dim\left(W_1 + W_2\right) = \dim W_1 + \dim W_2$$

定义： 若 2 个子空间，$W_1,W_2 \subset C^n\left(R^n\right)$，适合 $W_1 \perp W_2$，$\dim W_1 + \dim W_2 = n$

则称 W_1，W_2 互为正交补

W_1，W_2 互为正交补 $W_1 + W_2 = C^n$ 且 $W_1 \perp W_2$

一般子空间 W 的正交补记为 $W^{\perp}$（唯一的），$W \oplus W^{\perp} = C^n$

引理： $A = A_{m\times n} \in C^{m\times n}$，则（1）$\mathscr{N}\left(A^H\right) = \left\{x \middle| A^H x = 0\right\}$与 $\mathscr{R}(A)$正交

$$\mathscr{R}(A) \perp \mathscr{N}\left(A^H\right) \Rightarrow \mathscr{R}(A) \oplus \mathscr{N}(A^H) = C^m$$

（2）（同理）$\mathscr{N}(A) \perp \mathscr{R}(A^H)$

Pf: $\forall x \in \mathrm{N}\left(A^H\right)$，$\forall y \in \mathrm{R}\left(A\right)$，即 $y = AZ$

$$\Rightarrow \left(x, y\right) = y^H x = \left(AZ\right)^H x = Z^H\left(A^H x\right) = 0 \Rightarrow \mathscr{N}(A^H) \perp \mathscr{R}(A)$$

应用： 在实矩阵条件下：$A = A_{m\times n} \in R^{m\times n}$

有（1）$\mathscr{R}(A) \perp \mathscr{N}(A^T)$，$\mathscr{R}(A) \oplus \mathscr{N}(A^T) = R^m$，即 $\mathscr{R}(A)^{\perp} = \mathscr{N}(A^T)$

（2）$\mathscr{R}(A^T) \perp \mathscr{N}(A)$，$\mathscr{R}(A^T) \oplus \mathscr{N}(A) = R^n$，即 $\mathscr{R}(A^T)^{\perp} = \mathscr{N}(A)$

Ex.求$W = span\left(\alpha_1, \alpha_2, \cdots, \alpha_s\right) \subset R^n$的正交补 $W^{\perp}$

解：令 $A = (\alpha_1, \alpha_2, \ldots, \alpha_s)_{n\times s}$ 可知 $\mathscr{R}(A) = \mathrm{span}(\alpha_1, \alpha_2, \ldots, \alpha_s)$

求解空间 $\mathscr{N}\left(A^T\right) = \left\{x \middle| A^T x = 0\right\} = span\left(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_{n-r}\right)$，$r = \mathrm{rank}(A)$

$\Rightarrow \mathscr{N}(A^T) = \mathscr{R}(A) = \mathrm{span}(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{n-r})$

注： 幂等（投影）阵性质

幂等条件：$A^2 = A$，$A \in C^{n\times n}$ $A \in C^{n\times n}$

则有（1） $A^2=A\Leftrightarrow A(I-A)=(I-A)A=0$

（2）A 的特征值只可能为 0 或 1，且 A 可对角化（∵0 化式 x^2-x 无重根）

（3）$\mathscr{R}(A)\oplus\mathscr{N}(A)=C^n$

∵ $\forall x=Ax+(x-Ax)=(Ax)+(I-A)x$，且 $A(I-A)x=(A-A^2)x=0x=0$

即 $Ax\in\mathscr{R}(A)$，$(I-A)x\in\mathscr{N}(A)\Rightarrow C^n=\mathscr{R}(A)\oplus\mathscr{N}(A)$

且 $\dim\mathscr{R}(A)\oplus\dim\mathscr{N}(A)=r(A)+n-r(A)=n=\dim C^n$

$C^n=\mathscr{R}(A)+\mathscr{N}(A)$

（4）$\mathscr{N}(A)=\mathscr{R}(I-A)$

∵ $\forall y\in\mathscr{R}(I-A)$，$y=(I-A)x$

必有：$Ay=A(I-A)x=0\Rightarrow y\in\mathscr{N}(A)$

同样：$\forall y\in\mathscr{N}(A)\Rightarrow Ay=0\Rightarrow y=y-Ay=(I-A)y\in\mathscr{R}(I-A)$

eg.若 $A^2=A=A^H$（叫正交投影）

则 $\mathscr{N}(A^H)\perp\mathscr{R}(A)$，即 $\mathscr{N}(A)\perp\mathscr{R}(A)$，且 $\mathscr{R}(A)^{\perp}=\mathscr{N}(A)$，$\mathscr{R}(A)\oplus\mathscr{N}(A)=C^n$

§4　常用矩阵分解

已知分解公式：QR 分解；许尔公式 $A=Q\begin{pmatrix}\lambda_1 & & * \\ & \lambda_2 & \\ O & & \ddots \\ & & & \lambda_n\end{pmatrix}Q^H$，（$Q$ 为 $\mathscr{U}$阵）；正规分解；谱分解

秩 1 分解

若 $\operatorname{rank}(A)=1$，则 $A=\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix}(b_1,b_2,\cdots,b_n)=\alpha\beta^T$

Pf：∵ $A=A_{m\times n}$，$\operatorname{rank}(A)=1$

$\Rightarrow$有一个非 0 列 $\alpha=\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix}$，其它列都是 α 的倍数

可写 $\alpha_1 = b_1\alpha, \alpha_2 = b_2\alpha, \dots, \alpha_n = b_n\alpha$

$$\Rightarrow A = (\alpha_1, \alpha_2, \cdots, \alpha_n)_{m\times n} = (b_1\alpha, b_2\alpha, \cdots, b_n\alpha) = \alpha(b_1, b_2, \cdots, b_n) = \alpha\beta^T$$

eg. $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}(1,1,1)$

秩 r 分解（也叫满秩分解或高低分解）

条件：$A = A_{m\times n}$，$\text{rank}(A) = r$

则有 $A = B_{m\times r}\, C_{r\times n}$

eg. $A = \begin{pmatrix} 1 & 1 & 2 \\ 0 & 2 & 2 \\ 1 & 1 & 2 \end{pmatrix}$，取 $\beta_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$，$\beta_2 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$

则：$\alpha_1 = \beta_1, \alpha_2 = \beta_2, \alpha_1 = \beta_1 + \beta_2$

$$\Rightarrow A = (\alpha_1, \alpha_2, \alpha_3) = (\beta_1, \beta_2)\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$$ （广阵格式）

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

eg. $A = \begin{pmatrix} 2 & 4 & 1 & 1 \\ 1 & 2 & -1 & 2 \\ -1 & 2 & -2 & 1 \end{pmatrix}_{3\times 4}$ 取 $\beta_1 = \begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix}$，$\beta_2 = \begin{pmatrix} 1 \\ -1 \\ -2 \end{pmatrix}$

写 $\alpha_1 = \beta_1, \alpha_2 = 2\beta_1, \alpha_3 = \beta_2, 4 = \beta_1 - \beta_2$

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

$$A = (\alpha_1, \alpha_2, \alpha_3, \alpha_4)\underline{\underline{(\text{广阵})}}(\beta_1, \beta_2)\begin{pmatrix} 1 & 2 & 0 & 1 \\ 0 & 0 & 1 & -1 \end{pmatrix}$$

方法： $A \xrightarrow{\text{行变}} \widetilde{C} = \begin{pmatrix} I_r & * & \cdots & * \\ 0 & * & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & * & \cdots & * \end{pmatrix}$ （行最简形）

I_r 表示 A 中前 r 列为极大无关组

则有 $A = (\beta_1, \beta_2, \ldots, \beta_r)(I_r, *, \ldots, *)_{r\times n}$

注： 若 $\widetilde{C}$ 中的 I_r 分布在其它列，则有相应公式

解法：$A \xrightarrow{\text{行变}} \begin{pmatrix} 1 & 2 & 0 & 1 \\ 0 & 0 & 1 & -1 \\ \hline 0 & 0 & 0 & 0 \end{pmatrix}$，取 $\beta_1 = \begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix}$，$\beta_2 = \begin{pmatrix} 1 \\ -1 \\ -2 \end{pmatrix}$

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

Ex.《矩阵分析》P211　1（1）（2）（3）

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

秩 r 分解（满秩分解） $A = A_{m\times n}$，$\mathrm{rank}(A) = r$

有 $A = B_{m\times r}\, C_{r\times n}$, $B = B_{m\times r}$ 为列满秩（高阵）

$C = C_{r\times n}$ 为行满秩（低阵）

引理： 设 $A \xrightarrow{\text{行变}} \begin{pmatrix} I_r & D \\ 0 & 0 \end{pmatrix}$　（行最简形）

则 A 中前 r 列 $\alpha_1, \alpha_2, \ldots, \alpha_r$ 为无关组

且 $A = (\alpha_1, \alpha_2, \ldots, \alpha_r)(I_r \quad D) = BC$

Pf: 由条件 $A = P\begin{pmatrix} I_r & D \\ 0 & 0 \end{pmatrix}$　（P 为可逆阵）

$$\because \begin{pmatrix} I_r & D \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} I_r \\ 0 \end{pmatrix}\begin{pmatrix} I_r & D \end{pmatrix}$$

令 $P\begin{pmatrix} I_r \\ 0 \end{pmatrix} = B = B_{m\times r} \Rightarrow A = P\begin{pmatrix} I_r \\ 0 \end{pmatrix}\begin{pmatrix} I_r & D \end{pmatrix} = B\begin{pmatrix} I_r & D \end{pmatrix}$

$\because A = (\alpha_1, \alpha_2, \ldots, \alpha_r, \ldots)$　　$B(I_r\, D) = (BI_r \quad BD) = (B, \ldots)$

$$(\alpha_1, \alpha_2, \cdots, \alpha_r | \cdots) = (B | \cdots) \Rightarrow (\alpha_1, \alpha_2, \cdots, \alpha_r) = B$$

$A = (\alpha_1, \alpha_2, \ldots, \alpha_r)(I_r\ D)$

同样： 当 I_r 的列分布在其它位置，也有相应结论 $A = (\beta_1, \beta_2, \ldots, \beta_r)C$

eg. $A=\begin{pmatrix}0&1&-1&-1&1\\0&-2&2&-2&6\\0&1&-1&-2&3\end{pmatrix}_{3\times5}$

解：（行变法）$A\xrightarrow{\text{行变}}\begin{pmatrix}0&\begin{pmatrix}1\\0\end{pmatrix}&\begin{matrix}-1\\0\end{matrix}&\begin{pmatrix}0\\1\end{pmatrix}&\begin{matrix}-1\\-2\end{matrix}\\\overline{0}&\overline{0}&\overline{0}&\overline{0}&\overline{0}\end{pmatrix}$

取 $B=(\beta_1,\beta_2)=\begin{pmatrix}1&-1\\-2&-2\\1&-2\end{pmatrix}\Rightarrow A=\begin{pmatrix}1&-1\\-2&-2\\1&-2\end{pmatrix}\begin{pmatrix}0&1&-1&0&-1\\0&0&0&1&-2\end{pmatrix}$

列满（高）阵性质

引理：（1）若 $B=B_{m\times r}$ 为列满的，则它有左逆阵 B_L 使得 $B_LB=I_r$

（2）若 $C=C_{r\times n}$ 为行满的，则它有右逆阵 C_R 使得 $CC_R=I_r$

其中可取 $B_L=(B^HB)^{-1}B^H \qquad C_R=C^H(CC^H)^{-1}$

验证：（1）若 B 为列满的，则左消法成立：$BX=BY\Leftrightarrow X=Y$

（2）若 C 为行满的，则右消法成立：$XC=YC\Leftrightarrow X=Y$

引理：（1）$y^Hy=|y|^2=0\Leftrightarrow y=0$，$y\in C^n$

（2）$rank(A)=rank(A^T)=rank(\overline{A})=rank(A^H)$

（3）$\mathrm{rank}(A^HA)=\mathrm{rank}(AA^H)=\mathrm{rank}(A) \qquad A=A_{m\times n}$

Pf：$\because$ 先证明 $A^HAx=0$ 与 $Ax=0$ 同解

$\because$ 由 $A^HAx=0\Rightarrow x^HA^HAx=0\Rightarrow (Ax)^H(Ax)=0$ （$y^Hy=0$）

$\Rightarrow Ax=0$ （同解）

$\Rightarrow rank(A^HA)=rank(A)$

（4）$A^HA=0\Leftrightarrow A=0$

Pf：$\because \mathrm{rank}(A)=\mathrm{rank}(A^HA)=0$

$rank(A)=0\Rightarrow A=0$

注： 利用 $tr(A^HA)=tr(AA^H)=\sum|a_{ij}|^2$

（5）$A^HAx=0\Leftrightarrow Ax=0$ （同解）

（6）A^HA、AA^H都是 Hermite 半正定阵，且有相同的正特征值

∵（$|xI_m - AA^H| = x^{m-n}|xI_n - A^HA|$）

定义：设 $A = A_{m\times n}$，A^HA 得特征值（谱）为 $\lambda_1 \ge 0, \lambda_2 \ge 0, \ldots, \lambda_n \ge 0$

称$\sqrt{\lambda_1}, \sqrt{\lambda_2}, \cdots, \sqrt{\lambda_n}$为 A 的奇异值

若 $\mathrm{rank}(A^HA) = \mathrm{rank}(A) = r$，则恰有 r 个正根 $\lambda_1 \ge 0, \lambda_2 \ge 0, \ldots, \lambda_r \ge 0$，$\lambda_{r+1} = \lambda_{r+2} = \ldots = \lambda_n = 0$，称$\sqrt{\lambda_1}, \sqrt{\lambda_2}, \cdots, \sqrt{\lambda_r}$为 A 的正奇异值

（同样：A^HA 也有 r 个正根 $\lambda_1 \ge 0, \lambda_2 \ge 0, \ldots, \lambda_r \ge 0$）

第 1 奇异分解公式（短分解）：设 $A = A_{m\times n}$ 的正奇异值为$\sqrt{\lambda_1}, \sqrt{\lambda_2}, \cdots, \sqrt{\lambda_r}$

$r = \mathrm{rank}(A)$则有 2 个次 $\mathscr{U}$阵 P_1 与 Q_1

使得 $A = P_1 \Delta Q_1^H$ $\Delta = \begin{pmatrix} \sqrt{\lambda_1} & & & O \\ & \sqrt{\lambda_2} & & \\ & & \ddots & \\ O & & & \sqrt{\lambda_r} \end{pmatrix}$ （$P_1^H P_1 = I_r$， $Q_1^H Q_1 = I_r$）

Pf：∵ $Q^H\left(A^H A\right)Q = \begin{pmatrix} \lambda_1 & & & O \\ & \lambda_2 & & \\ & & \ddots & \\ O & & & \lambda_n \end{pmatrix}$ （正规分解） Q 为 $\mathscr{U}$阵

且 $\lambda_1 \ge \lambda_2 \ge \ldots \ge \lambda_r > 0, \lambda_{r+1} = \lambda_{r+2} = \ldots = \lambda_n = 0$ $r = \mathrm{rank}(A)$

写 $Q = (q_1, q_2, \ldots, q_n)_{n\times n}$

$\Rightarrow \left(A^H A\right)q_1 = \lambda_1 q_1, \left(A^H A\right)q_2 = \lambda_2 q_2, \cdots, \left(A^H A\right)q_r = \lambda_r q_r$

$\left(A^H A\right)q_{r+1} = \lambda_{r+1} q_{r+1} = 0, \cdots, \left(A^H A\right)q_n = \lambda_n q_n = 0$

令$Q_1 = \left(\frac{q_1}{|q_1|}, \frac{q_2}{|q_2|}, \cdots, \frac{q_r}{|q_r|}\right)_{n\times r}$ （次酉）

$P_1 = \left(\frac{Aq_1}{|Aq_1|}, \frac{Aq_2}{|Aq_2|}, \cdots, \frac{Aq_r}{|Aq_r|}\right)_{m\times r}$ 也为次酉

∵ $(Aq_1, Aq_2) = (Aq_2)^H (Aq_1) = q_2^H \left(A^H A\right)q_1 = \lambda_1\left(q_2^H q_1\right) = 0$ （$q_1 \perp q_2$）

∴P_1 为次酉

又知：$|Aq_1|^2 = (Aq_1)^H (Aq_1) = q_1^H \left(A^H A\right)q_1 = \lambda_1\left(q_1^H q_1\right) = \lambda_1 |q_1|^2 = \lambda_1$

$\Rightarrow |Aq_1| = \sqrt{\lambda_1}$，同理$|Aq_2| = \sqrt{\lambda_2}, \cdots, |Aq_r| = \sqrt{\lambda_r} \ge 0$

$$\Rightarrow P_1=\left(\frac{Aq_1}{|Aq_1|},\frac{Aq_2}{|Aq_2|},\cdots,\frac{Aq_r}{|Aq_r|}\right)=\left(\frac{Aq_1}{\sqrt{\lambda_1}},\frac{Aq_2}{\sqrt{\lambda_2}},\cdots,\frac{Aq_r}{\sqrt{\lambda_r}}\right)_{m\times r}$$ （次 U）

注： 由 $A^HAx=0\Leftrightarrow Ax=0$

$$\therefore A^HAq_{r+1}=A^HAq_{r+2}=\cdots=A^HAq_n=0\Rightarrow Aq_{r+1}=Aq_{r+2}=\cdots=Aq_n=0$$

$$\Rightarrow A\left(q_1q_1^H+q_2q_2^H+\cdots+q_rq_r^H\right)=A\left(q_1q_1^H+q_2q_2^H+\cdots+q_rq_r^H+q_{r+1}q_{r+1}^H+\cdots+q_nq_n^H\right)$$

且 $q_1q_1^H+q_2q_2^H+\cdots+q_nq_n^H=(q_1,q_2,\cdots,q_n)\begin{pmatrix}q_1^H\\q_2^H\\\vdots\\q_n^H\end{pmatrix}=QQ^H=I_n$

$$\Rightarrow A\left(q_1q_1^H+q_2q_2^H+\cdots+q_nq_n^H\right)=AI_n=A$$

计算：$P_1\Delta Q_1^H=\left(\frac{Aq_1}{\sqrt{\lambda_1}},\frac{Aq_2}{\sqrt{\lambda_2}},\cdots,\frac{Aq_r}{\sqrt{\lambda_r}}\right)\begin{pmatrix}\sqrt{\lambda_1}&&&O\\&\sqrt{\lambda_2}&&\\&&\ddots&\\O&&&\sqrt{\lambda_r}\end{pmatrix}\begin{pmatrix}q_1^H\\q_2^H\\\vdots\\q_n^H\end{pmatrix}$

$$=(Aq_1,Aq_2,\cdots,Aq_r)\begin{pmatrix}q_1^H\\q_2^H\\\vdots\\q_n^H\end{pmatrix}$$

$$=A\left(q_1q_1^H+q_2q_2^H+\cdots+q_rq_r^H\right)=AI_n=A$$

第 2 奇异分解公式： 设 $A=A_{m\times n}$，奇异值为 $\sqrt{\lambda_1},\sqrt{\lambda_2},\cdots,\sqrt{\lambda_r}>0$ 则存在 2 个次 U 阵

$P=P_{m\times m}$，$Q=Q_{n\times n}$ 使得 $A=P(\sum)Q^H$ $\quad\Sigma=\begin{pmatrix}\Delta&0\\0&0\end{pmatrix}_{m\times n}$

Pf：把 P_1、Q_1 扩充为 U 阵

$P=(P_1,P_2)_{m\times m}$ （不唯一） $Q=(Q_1,Q_2)_{n\times n}$ （不唯一）

计算 $P(\Sigma)Q^H=(P_1,P_2)\begin{pmatrix}\Delta&0\\0&0\end{pmatrix}(Q_1,Q_2)=P_1\Delta Q_1^H=A$

分解方法： 1.求(A^HA)得特征值 $\lambda_1\ge\lambda_2\ge\ldots\ge\lambda_r>0$ $\quad r=\mathrm{rank}(A)$

正奇异值为 $\sqrt{\lambda_1},\sqrt{\lambda_2},\cdots,\sqrt{\lambda_r}$

2.求 $\lambda_1,\lambda_2,\ldots,\lambda_r$ 对应的正交特征向量：$q_1,q_2,\ldots,q_r$（不必单位化）

3.令次酉阵 $P_1=\left(\frac{Aq_1}{|Aq_1|},\frac{Aq_2}{|Aq_2|},\cdots,\frac{Aq_r}{|Aq_r|}\right)$，$Q_1=\left(\frac{q_1}{|q_1|},\frac{q_2}{|q_2|},\cdots,\frac{q_r}{|q_r|}\right)$

则有 $A=P_1\Delta Q_1^H$ $\Delta=\begin{pmatrix}\sqrt{\lambda_1} & & & O\\ & \sqrt{\lambda_2} & & \\ & & \ddots & \\ O & & & \sqrt{\lambda_r}\end{pmatrix}$

4.可用观察扩充法求 2 个 $\mathscr{U}$阵 $P=(P_1,P_2)$，$Q=(Q_1,Q_2)$

则有 $A=P(\sum)Q^H$ $\Sigma=\begin{pmatrix}\Delta & 0\\ 0 & 0\end{pmatrix}_{m\times n}$

注： 可求 $A^Hx=0$ 得 P_2，可求 $Ax=0$ 得 Q_2 （$\because \begin{matrix}Ax=0\Leftrightarrow A^HAx=0\\ AA^H=0\Leftrightarrow AA^Hx=0\end{matrix}$ ）

eg.（1）$A=\begin{pmatrix}1 & 1\\ 0 & 0\\ 1 & 1\end{pmatrix}_{3\times 2}$ （2）$C=\begin{pmatrix}1 & 0 & 1\\ 1 & 0 & 1\end{pmatrix}_{2\times 3}$ 求奇异分解

解：（1）$A^HA=\begin{pmatrix}2 & 2\\ 2 & 2\end{pmatrix}$，谱 $\sigma(A^HA)=\{4,0\}$，奇异值 $\sqrt{4}=2$

$\lambda=4$ 的特征向量：$q_1=\begin{pmatrix}1\\1\end{pmatrix}$，令 $Q_1=\left(\frac{q_1}{|q_1|}\right)=\begin{pmatrix}\frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}}\end{pmatrix}$，令 $P_1=\left(\frac{Aq_1}{|Aq_1|}\right)=\begin{pmatrix}\frac{1}{\sqrt{2}}\\ 0\\ \frac{1}{\sqrt{2}}\end{pmatrix}$

$\Rightarrow$（短分解）$A=P_1\Delta Q_1^H=\begin{pmatrix}\frac{1}{\sqrt{2}}\\ 0\\ \frac{1}{\sqrt{2}}\end{pmatrix}(2)\left(\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}}\right)$

扩充为 $\mathscr{U}$阵 $P=(P_1,P_2)=\begin{pmatrix}\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}}\\ 0 & 1 & 0\\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}}\end{pmatrix}$ （$\mathscr{U}$阵） （$P^H=P$）

$$Q=\left(Q_1,Q_2\right)=\begin{pmatrix}\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}}\end{pmatrix}$$ （$\mathscr{U}$阵） （$Q^H=Q$）

$$\Rightarrow A=P\Sigma Q^H=\begin{pmatrix}\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}}\\ 0 & 1 & 0\\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}}\end{pmatrix}\begin{pmatrix}2 & 0\\ 0 & 0\\ 0 & 0\end{pmatrix}\begin{pmatrix}\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}}\end{pmatrix}$$

（2）同理可求 C^HC 的特征值$\{4, 0, 0\}$

或用转置公式：$C^H=A=P\Sigma Q^H\Rightarrow C=A^H=\left(P\Sigma Q^H\right)^H=Q\Sigma P^H=\cdots$

Ex.求奇异分解

1.（1）$A=\begin{pmatrix}1 & 0\\ 0 & 1\\ 1 & 1\end{pmatrix}$ （2）$C=\begin{pmatrix}1 & 0 & 1\\ 0 & 1 & 1\end{pmatrix}$

2.（1）$A=\begin{pmatrix}2 & 3\\ 0 & 0\\ 0 & 0\end{pmatrix}$ （2）$C=\begin{pmatrix}2 & 0 & 0\\ 3 & 0 & 0\end{pmatrix}$

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

§5 范数与级数

eg. C^n 中向量 $x=(x_1, x_2, \ldots, x_n)^T$ 长度

$$|x|=\sqrt{(x,x)}=\sqrt{x^Hx}=\sqrt{tr\left(xx^H\right)}=\sqrt{|x_1|^2+|x_2|^2+\cdots+|x_n|^2}$$

有 3 条性质：

（1）正性：$|x|>0$（$x\neq\vec{0}$） $|0|=0$ （$|x|=0\Leftrightarrow x=\vec{0}$）

（2）齐性：$|kx|=|k||x|$

（3）三角性：$|x+y|\le|x|+|y|$

推论：（1）$|-x|=|x|$，（2）$||x|-|y||\le|x-y|$

注： 任一内积空间 V 中，可引入长度（模）：$|\alpha|=\sqrt{(\alpha,\alpha)}$ $\alpha\in V$

有舒瓦兹不等式：$|(\alpha,\beta)| \le \sqrt{(\alpha,\alpha)}\sqrt{(\beta,\beta)} = |\alpha||\beta| \Rightarrow$ 三角性：$|\alpha + \beta| \le |\alpha| + |\beta|$

且有（1）正性，（2）齐性

范数定义：若空间 V 中有一个实函数 $\varphi(x) = ||x||$ 适合

（1）正性：$\varphi(x) > 0$（$x \neq \vec{0}$） $\varphi(0) = 0$

（2）齐性：$\varphi(kx) = |k|\varphi(x)$

（3）三角式：$\varphi(x + y) \le \varphi(x) + \varphi(y)$

则称：$\varphi(x) = ||x||$ 为 V 上一个范数

推论：（1）$||-x|| = ||x||$，（2）$|||x|| - ||y||| \le ||x - y||$

C^n 中常用 3 种范数：$x = (x_1, x_2, \dots, x_n)^T$

（1）∞范数：$||x||_\infty = \max(|x_1|, |x_2|, \dots, |x_n|)$

（2）1 范数：$||x||_1 = \sum|x_j| = |x_1| + |x_2| + \dots + |x_n|$

（3）2 范数：$\|x\|_2 = |x| = \sqrt{(x,x)} = \sqrt{x^H x} = \sqrt{tr(xx^H)} = \sqrt{|x_1|^2 + |x_2|^2 + \cdots + |x_n|^2}$（长度）

都具有（1）（2）（3）

eg. $V = C^{n\times n}$ $A = (a_{ij}), B = (b_{ij}) \in V$

利用内积：$(A,B) = tr(AB^H) = tr(B^H A) = \sum a_{ij}\overline{b_{ij}}$

长度（模）$\|A\| = \sqrt{(A,A)} = \sqrt{tr(AA^H)} = \sqrt{\sum |a_{ij}|^2}$

具有（1）正性（2）齐性（3）$||A + B|| \le ||A|| + ||B||$

应用（收敛性）

范数等价定理：任 2 个范围 $||x||_a$、$||x||_b$ 都等价（略证）

即 $\exists k_1 > 0, k_2 > 0$ 使得 $k_1||x||_b \le ||x||_a \le k_2||x||_b$ 对一切 x 成立

收敛定义：设 $x^{(k)} = \left(x_1^{(k)}, x_2^{(k)}, \cdots, x_n^{(k)}\right)^T$（$k = 1, 2, 3, \dots$）为向量列，$\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)^T$

若 $x_1^{(k)} \to a_1, x_2^{(k)} \to a_2, \cdots, x_n^{(k)} \to a_n$（$k \to \infty$）

称 $x^{(k)} \to \alpha$ 或 $\lim x^{(k)} = \alpha$

收敛引理：$x^{(k)} \to \alpha \Leftrightarrow \left\|x^{(k)} - \alpha\right\| \to 0$ （$||x||$ 为任取范数）

（可用 $||x||_1$ 证明，再用等价性）

矩阵范数定义：设 $C^{n\times n}$ 上有实函数 $\varphi(x)=||x||\ x\in C^{n\times n}$

适合：（1）正性 $\varphi(x)>0, \varphi(0)=0$

（2）齐性：$\varphi(kx)=|k|\varphi(x)$

（3）三角式：$\varphi(x+y)\le\varphi(x)+\varphi(y)$

（4）相容性：$\varphi(xy)\le\varphi(x)\varphi(y)\quad ||xy||\le||x||\,||y||$

称 $\varphi(x)=||x||$ 为 $C^{n\times n}$ 上的矩阵范数（方阵范数）

eg. m_1 范数：任取 $A=(a_{ij})\ \in C^{n\times n}$，规定 $\varphi(x)=\|x\|_{m_1}\triangleq\sum|a_{ij}|$ （总合）

则有：（1）正性；（2）齐性；（3）三角性；（4）相容 $||AB||_m\le||A||_m||B||_m$

eg. F 范数：任取 $A=(a_{ij})_{n\times n}$，$B=(b_{ij})_{n\times n}$，规定 $\|A\|_F=\sqrt{tr(A^HA)}=\sqrt{\sum|a_{ij}|^2}$

则有：则有：（1）正性；（2）齐性；（3）三角性；（4）相容 $||AB||_F\le||A||_F||B||_F$

$C^{n\times n}$ **也有 3 种常用范数**：$A=(a_{ij})_{n\times n}=(\alpha_1,\alpha_2,\cdots,\alpha_n)=\begin{pmatrix}A_1\\A_2\\\vdots\\A_n\end{pmatrix}$（行）

∞范数：$||A||_\infty=\max(||A_1||_1, ||A_2||_1, \ldots, ||A_n||_1)$（行范数）

1 范数：$||A||_1=\max(||\alpha_1||_1, ||\alpha_2||_1, \ldots, ||\alpha_n||_1)$（列范数）

2 范数：$\|A\|_2=\sqrt{\lambda_1}=\sqrt{(A^HA)}$的最特征值（最大奇异值）

都有（1）（2）（3）（4）相容性 $||AB||\le||A||\,||B||$

矩阵与向量的相容条件：$||Ax||\le||A||\,||x||$，$A\in C^{n\times n}$，$x\in C^n$

设 $||A||_m$ 为矩阵范数，$||x||_v$ 为向量范数

相容条件为：$||Ax||_v\le||A||_m||x||_v$

定理 1：每种向量范数 $||x||_v$ 都对应一种矩阵范数 $||A||_m$ 使得 $||Ax||_v\le||A||_m||x||_v\ A\in C^{n\times n}, x\in C^n$

注：这种与 $||x||_v$ 对应的 $||A||_m$ 叫做 $||x||_v$ 的诱导范数

常用诱导范数

（1）∞范数 $||x||_\infty$ 诱导 $||A||_\infty$（行范）$||Ax||_\infty\le||A||_\infty||x||_\infty$

（2）1 范数 $||x||_1$ 诱导 $||A||_1$（列范）$||Ax||_1\le||A||_1||x||_1$

（3）2 范数 $||x||_2=|x|$ 诱导 $||A||_2$（根范）$||Ax||_2\le||A||_2||x||_2$

其它相容公式：$\begin{cases}\|Ax\|_2 \le \|A\|_F \|x\|_2 \\ \|Ax\|_1 \le \|A\|_{m_1} \|x\|_1\end{cases}$

定理 3： 任一矩阵范数：$\|A\|$都对应（诱导）一个向量范数$\|x\|$

使得$\|Ax\| \le \|A\|\|x\|$　$\forall A \in C^{n\times n}$，$x\in C^n$

Pf：固定一个向量$\alpha = (a_1, a_2, \cdots, a_n)^T \ne \vec{0}$

规定$\|x\| \triangleq \|(x\alpha^T)\|$（有定义！）　$x\alpha^T \in C^{n\times n}$

check（1）$\|x\| > 0$　（$x \ne \vec{0}$）　（2）$\|kx\| = |k|\|x\|$　（3）三角

（4）$\|Ax\| \triangleq \|(Ax)\alpha^T\| = \|A(x\alpha^T)\| \le \|A\|\|x\alpha^T\| = \|A\|\|x\|$

eg.取$\|A\| = \|A\|_F$，取 $\alpha = (1, 0, \ldots, 0)^T$

规定$\|x\| \triangleq \|x\alpha^T\|_F = \left\|\begin{pmatrix} x_1 & 0 & \cdots & 0 \\ x_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ x_n & 0 & \cdots & 0 \end{pmatrix}\right\|_F = \sqrt{|x_1|^2 + |x_2|^2 + \cdots + |x_n|^2} = \|x\|_2$

$\Rightarrow \|Ax\|_2 \le \|A\|_F \|x\|_2$

引理： $A\in C^{n\times n}$，$\|A\|$为任一范数

则：（1）$\|A^k\| \le \|A\|^k$（$\because \|A^2\| \le \|A\|^2$，…）

（2）谱半径 $\rho(A) \le \|A\|$，其中 $\rho(A) = \max(|\lambda_1|, |\lambda_2|, \ldots, |\lambda_n|)$

（3）A 为正规$\Rightarrow \rho(A) = \|A\|_2$

引理： 固定 $Q = Q_{n\times n}$（可逆）$\|X\|$为矩阵范数，$X\in C^{n\times n}$

令 $\varphi(X) = \|Q^{-1}XQ\|$，则 $\varphi(X)$也是矩阵范数

$\because$（1）（2）（3）

（4）：$\varphi(AB) = \|Q^{-1}(AB)Q\| = \|(Q^{-1}AQ)(Q^{-1}BQ)\| \le \|Q^{-1}AQ\|\|Q^{-1}BQ\| = \varphi(A)\,\varphi(B)$

小范数公式： 设 ε 为任取小正数，则某一范数$\|x\|$

$A = A_{n\times n}$ 为已知，使得（1）$\|A\| \le \rho(A) + \varepsilon$；（2）若 $\rho(A) < 1$ 则有$\|A\| < 1$

Pf：$\because A \backsim J$（Jordan 形）

可写 $J=\begin{pmatrix}\lambda_1 & (*) & & \\ & \lambda_2 & \ddots & \\ & & \ddots & (*) \\ & & & \lambda_n\end{pmatrix}$ (*)为 0 或 1

$\exists$ 可逆 $P=P_{n\times n}$ 使得 $P^{-1}AP=J=\begin{pmatrix}\lambda_1 & (*) & & \\ & \lambda_2 & \ddots & \\ & & \ddots & (*) \\ & & & \lambda_n\end{pmatrix}$ $|*|\le 1$

取 $D=\begin{pmatrix}\varepsilon & & & \\ & \varepsilon^2 & & \\ & & \ddots & \\ & & & \varepsilon^n\end{pmatrix}\Rightarrow D^{-1}\left(P^{-1}AP\right)D=D^{-1}JD=\begin{pmatrix}\lambda_1 & (*)\varepsilon & & & \\ & \lambda_2 & (*)\varepsilon & & \\ & & \ddots & \ddots & \\ & & & \lambda_{n-1} & (*)\varepsilon \\ & & & & \lambda_n\end{pmatrix}$

引用∞（行）范数$\|X\|_\infty$；$X\in C^{n\times n}$

令 $Q=PD$，则 $\varphi(X)=\|Q^{-1}XQ\|_\infty$为新范数

$$\Rightarrow \varphi(A)=\left\|Q^{-1}AQ\right\|_\infty=\left\|\begin{pmatrix}\lambda_1 & (*)\varepsilon & & & \\ & \lambda_2 & (*)\varepsilon & & \\ & & \ddots & \ddots & \\ & & & \lambda_{n-1} & (*)\varepsilon \\ & & & & \lambda_n\end{pmatrix}\right\|_\infty$$

$\because |\lambda_1|+|(*)\varepsilon|\le|\lambda_1|+\varepsilon, |\lambda_2|+|(*)\varepsilon|\le|\lambda_2|+\varepsilon, \ldots$

$\Rightarrow \varphi(A)=\left\|Q^{-1}AQ\right\|_\infty\le\rho(A)+\varepsilon$

（2） $\because\rho(A)<1$，取 $\varepsilon>0$ 很小，$\rho(A)+\varepsilon<1$

则$\|A\|\le\rho(A)+\varepsilon<1$

常用范数公式：

小范数引理： 固定任一方阵 A；$\forall\varepsilon>0$，则有某个范数$\|\cdot\|_A$ 使得：

（1）$\|A\|_A\le\rho(A)+\varepsilon$

（2）若 $\rho(A)<1$，也有$\|A\|_A<1$

注： 这个$\|\cdot\|_A$与 A 有关，对另外矩阵 B，$\|B\|_A\le\rho(B)+\varepsilon$ 不一定成立

引理：（1） $\rho(A)<1\Rightarrow\left\|A^k\right\|\to 0$ （$k\to\infty$）

（2）某一范数$\|A\|<1 \Rightarrow \|A^k\| \to 0 \quad (k \to \infty)$

Pf:（2）若$\|A\|<1 \Rightarrow \|A^k\| \le \|A\|^k \to 0 \Rightarrow \|A^k\| \to 0$

（1）若$\rho(A)<1 \Rightarrow \exists$某范数$\|A\|<1 \overset{(2)}{\Rightarrow} \|A^k\| \to 0$

若$\|A^k\| \to 0$且$\rho(A^k) \le \|A^k\| \to 0 \Rightarrow \rho(A^k) \to 0$

又$\rho(A^k) = (\rho(A))^k \to 0$

∴（1）$\Rightarrow \lim\limits_{k\to\infty} A^k = 0$

推论： $\forall \varepsilon > 0$，$\dfrac{\|A^k\|}{(\rho(A)+\varepsilon)^k} \to 0 \quad (k \to \infty) \Rightarrow \|A^k\| \le M(\rho(A)+\varepsilon)^k$（$k$很大）

Pf: 令$B = \dfrac{A}{\rho(A)+\varepsilon} \Rightarrow \rho(B) = \dfrac{\rho(A)}{\rho(A)+\varepsilon} < 1 \Rightarrow B^k \to 0 \quad (k \to \infty)$

注： $\sigma(A) = \{\lambda_1, \lambda_2, \ldots, \lambda_n\}$（$A$的谱）；$\rho(A) = \max\{|\lambda_1|, |\lambda_2|, \ldots, |\lambda_n|\}$

引理：（1）若$\sigma(A) = \{\lambda_1, \lambda_2, \ldots, \lambda_n\}$，则$\sigma(A^k) = \{\lambda_1^k, \lambda_2^k, \cdots, \lambda_n^k\}$

（2）对任一多项式$f(x)$有$\sigma(f(A)) = \{f(\lambda_1), f(\lambda_2), \ldots, f(\lambda_n)\}$

（3）$\rho(A^k) = \max\{|\lambda_1|^k, |\lambda_2|^k, \ldots, |\lambda_n|^k\} = \rho^k(A)$

定义： 级数$\sum\limits_{k=0}^{\infty} A_k = A_0 + A_1 + \cdots + A_k + \cdots$

（$A_k \in C^{n\times n}$）它收敛于$A \Leftrightarrow \lim\limits_{k\to\infty}(A_0 + A_1 + \cdots + A_k) = A$

引理： 绝对收敛，必收敛

即：$\sum\limits_{k=0}^{\infty}\|A_k\| = \|A_0\| + \|A_1\| + \cdots + \|A_k\| + \cdots$收敛$\Rightarrow \sum\limits_{k=0}^{\infty} A_k$收敛

eg.（1）某个$\|A\|<1 \Rightarrow \sum\limits_{k=0}^{\infty} A^k = I + A + \cdots + A^k + \cdots$绝对收敛

（2）$\rho(A)<1 \Rightarrow \sum\limits_{k=0}^{\infty} A^k$绝对收敛

Pf:（1）$\|A\|<1 \Rightarrow \sum\limits_{k=0}^{\infty}\|A^k\| \le \sum\limits_{k=0}^{\infty}\|A\|^k$　（因为$\|A^k\| \le \|A\|^k$）

且$\sum_{k=0}^{\infty}\|A\|^k=\dfrac{\|I\|}{1-\|A\|}$　（绝对收敛）

（2）$\rho(A)<1\Rightarrow$某$\|A\|<1\Rightarrow\sum_{k=0}^{\infty}A^k$ 收敛

引理： 若$\rho(A)<1$，则$I+A+\ldots+A^k+\ldots=(I-A)^{-1}$

Pf：$\because(I+A+\ldots+A^k+\ldots)$收敛

$$\Rightarrow(I-A)(I+A+\cdots+A^k+\cdots)=I(I+A+\cdots+A^k+\cdots)-A(I+A+\cdots+A^k+\cdots)$$

$$=(I+A+\cdots+A^k+\cdots)-(A+A^2+\cdots+A^{k+1}+\cdots)=I$$

$$\Rightarrow(I-A)^{-1}=I+A+\cdots+A^k+\cdots$$

eg. $\rho(A)<1$（或$\|A\|<1$），则$(I-A)$可逆且$(I-A)^{-1}=I+A+\ldots+A^k+\ldots$

Pf：写$\sigma(A)=\{\lambda_1,\lambda_2,\ldots,\lambda_n\}$则$\sigma(\mathrm{I}-A)=\{1-\lambda_1,1-\lambda_2,\ldots,1-\lambda_n\}$　（取$f(x)=1-x$）

且：$|\lambda_1|\le\rho(A)<1,|\lambda_2|\le\rho(A)<1,\ldots,|\lambda_n|\le\rho(A)<1$

$|1-\lambda_1|\ge1-|\lambda_1|>0,|1-\lambda_2|\ge1-|\lambda_2|>0,\ldots,|1-\lambda_n|\ge1-|\lambda_n|>0$

$\Rightarrow\det(I-A)=|I-A|=(1-\lambda_1)(1-\lambda_2)\cdots(1-\lambda_n)\neq0\Rightarrow(I-A)$可逆

注： 幂级数$\sum_{k=0}^{\infty}C_kx^k=C_0+C_1x+\cdots+C_kx^k+\cdots$

规定$A=A_{n\times n}$的幂级数为$\sum_{k=0}^{\infty}C_kA^k=C_0I+C_1A+\cdots+C_kA^k+\cdots$

收敛定理： 设$\sum_{k=0}^{\infty}C_kx^k$ 的半径为 R

（1）$\rho(A)<R\Rightarrow\sum_{k=0}^{\infty}C_kA^k$ 绝对收敛

（2）$\rho(A)>R\Rightarrow\sum_{k=0}^{\infty}C_kA^k$ 发散（无意义）

常用级数公式：

由$f(x)=\sum_{k=0}^{\infty}C_kx^k$　　$|x|<R$（收敛半径）

$$\Rightarrow f(A) \triangleq \sum_{k=0}^{\infty} C_k A^k \quad \rho(A) < R \text{（或某}\|A\| < R\text{）}$$

（1） $f(x) = e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$ （$R = +\infty$）

$$f(A) = e^A \triangleq \sum_{k=0}^{\infty} \frac{A^k}{k!} \quad \text{（}A\text{ 为任一方阵）}$$

（2） $f(x) = \sin x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!}$ （$R = +\infty$）

$$f(A) = \sin A \triangleq \sum_{k=0}^{\infty} (-1)^k \frac{A^{2k+1}}{(2k+1)!} \quad \text{（}A\text{ 为任一方阵）}$$

（3） $f(x) = \cos x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!}$ （$R = +\infty$）

$$f(A) = \cos A \triangleq I - \frac{A^2}{2!} + \frac{A^4}{4!} - \frac{A^6}{6!} + \cdots + (-1)^k \frac{A^{2k}}{(2k)!} + \cdots \quad \text{（}A\text{ 为任一方阵）}$$

（4） $f(x) = (1-x)^{-1} = \sum_{k=0}^{\infty} x^k$ （$|x| < 1$）

$$f(A) = (I - A)^{-1} = I + A + \ldots + A^k + \ldots \quad \text{（}\rho(A) < 1\text{）}$$

（5） $f(x) = \ln(1+x) = \sum_{k=1}^{\infty} (-1)^{k+1} \frac{x^k}{k}$ （$|x| < 1$）

$$f(A) = \ln(I + A) = \sum_{k=1}^{\infty} (-1)^{k+1} \frac{A^k}{k} \quad \text{（}\rho(A) < 1\text{，或}\|A\| < 1\text{）}$$

eg. $A = \begin{pmatrix} \varepsilon & b \\ 0 & \varepsilon \end{pmatrix}$，$|\varepsilon| < 1$，求 $\sum_{k=0}^{\infty} A^k$

解： $\rho(A) = |\varepsilon| < 1 \Rightarrow \sum_{k=0}^{\infty} A^k$ 收敛

由公式 $\sum_{k=0}^{\infty} A^k = (I - A)^{-1} = \begin{pmatrix} 1-\varepsilon & -b \\ 0 & 1-\varepsilon \end{pmatrix}^{-1} = \frac{1}{(1-\varepsilon)^2} \begin{pmatrix} 1-\varepsilon & b \\ 0 & 1-\varepsilon \end{pmatrix}$

eg. $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$，求 $e^A = \sum_{k=0}^{\infty} \frac{(At)^k}{k!}$

解：$A^0 = I, A^1 = A, A^2 = AA = -I, A^3 = A(A^2) = -A, A^4 = A^2A^2 = I, \ldots$

$$e^{A}=e^{\begin{pmatrix}0 & t\\ -t & 0\end{pmatrix}}=\sum_{k=0}^{\infty}\frac{(At)^k}{k!}$$

$$=\left(I+\frac{A^2t^2}{2!}+\frac{A^4t^4}{4!}+\frac{A^6t^6}{6!}+\cdots\right)+\left(\frac{At}{1!}+\frac{A^3t^3}{3!}+\frac{A^5t^5}{5!}+\cdots\right)$$

$$=\left(I-\frac{t^2}{2!}I+\frac{t^4}{4!}I-\frac{t^6}{6!}I+\cdots\right)+\left(tA-\frac{t^3}{3!}A+\frac{t^5}{5!}A-\frac{t^7}{7!}A+\cdots\right)$$

$$=\left(i-\frac{t^2}{2!}+\frac{t^4}{4!}-\frac{t^6}{6!}+\cdots\right)I+\left(t-\frac{t^3}{3!}+\frac{t^5}{5!}-\frac{t^7}{7!}+\cdots\right)A$$

$$=(\cos t)I+(\sin t)A=\begin{pmatrix}\cos t & 0\\ 0 & \cos t\end{pmatrix}+\begin{pmatrix}0 & \sin t\\ -\sin t & 0\end{pmatrix}$$

$$\therefore e^{A}=\begin{pmatrix}\cos t & \sin t\\ -\sin t & \cos t\end{pmatrix}\to \text{正交阵}$$

引理： 设 $f(x)=\sum_{k=0}^{\infty}C_kx^k$，$|x|<R$

若 $\sigma(A)=\{\lambda_1,\lambda_2,\ldots,\lambda_n\}$全体特征值

则 $f(A)$的全体特征值为 $\sigma(f(A))=\{f(\lambda_1),f(\lambda_2),\ldots,f(\lambda_n)\}$

且行列式：$\det(f(A))=|f(A)|=f(\lambda_1)f(\lambda_2)\ldots f(\lambda_n)$

Pf：由许尔公式：$P^{-1}AP=\begin{pmatrix}\lambda_1 & & & *\\ & \lambda_2 & & \\ & & \ddots & \\ O & & & \lambda_n\end{pmatrix}$

$$C_k\left(P^{-1}AP\right)^k=P^{-1}C_kA^kP=\begin{pmatrix}C_k\lambda_1^k & & & *'\\ & C_k\lambda_2^k & & \\ & & \ddots & \\ O & & & C_k\lambda_n^k\end{pmatrix}\quad k=0,1,2,\cdots$$

相加：$\sum_{k=0}^{\infty}C_k\left(P^{-1}AP\right)^k=P^{-1}\left(\sum_{k=0}^{\infty}C_kA^k\right)P=\begin{pmatrix}\sum_{k=0}^{\infty}C_k\lambda_1^k & & & *''\\ & \sum_{k=0}^{\infty}C_k\lambda_2^k & & \\ & & \ddots & \\ O & & & \sum_{k=0}^{\infty}C_k\lambda_n^k\end{pmatrix}$

$$\Rightarrow f\left(P^{-1}AP\right)=P^{-1}f(A)P=\begin{pmatrix} f(\lambda_1) & & & (*) \\ & f(\lambda_2) & & \\ & & \ddots & \\ O & & & f(\lambda_n) \end{pmatrix}\rightarrow \text{上三角阵}$$

$\Rightarrow f(A)$的全体特征值为 $\sigma(f(A)) = \{f(\lambda_1), f(\lambda_2), \ldots, f(\lambda_n)\}$

推论：（1）$f(x) = e^x$，$\sigma(A) = \{\lambda_1, \lambda_2, \ldots, \lambda_n\}$

e^A的谱：$\sigma\left(e^A\right)=\left\{e^{\lambda_1},e^{\lambda_2},\cdots,e^{\lambda_n}\right\}$

且 $\det\left(e^A\right)=\left|e^A\right|=e^{\lambda_1}e^{\lambda_2}\cdots e^{\lambda_n}=e^{\lambda_1+\lambda_2+\cdots+\lambda_n}=e^{tr(A)}=e^{a_{11}+a_{22}+\cdots+a_{nn}}\neq 0$

且 e^A 可逆，$(e^A)^{-1} = e^{-A}$

（2）$f(x)=\sin x\Rightarrow\sigma(\sin A)=\{\sin\lambda_1,\sin\lambda_2,\cdots,\sin\lambda_n\}$

且 $\det(\sin A) = (\sin\lambda_1)(\sin\lambda_2)\ldots(\sin\lambda_n)$

可对角阵（单纯阵）计算公式：

设 $A = A_{n\times n}$ 可对角（极小式无重根），则有：

$f(A) = f(\lambda_1)G_1+f(\lambda_2)G_2+\ldots+f(\lambda_s)G_s$ （$f(x)=\sum_{k=0}^{\infty}C_k x^k$，$x < R$）

其中 $\lambda_1, \lambda_2, \ldots, \lambda_s$ 为 A 互异特征值，$G_1, G_2, \ldots, G_s$ 为投影阵（谱阵），$G_i=\dfrac{g_i(A)}{g_i(\lambda_i)}$

Ex.利用投影公式再解 P208 例 5 并求 e^A、P206 例 3 并求 e^A

Ex.P2476、8

Ex. $A=\begin{pmatrix}0&1&0\\0&0&1\\0&0&0\end{pmatrix}$用定义求 e^A

求 $f(A)$： 1.L-S 插值公式

2.待定矩阵法

A 的特征多项式：$c(x) = |\lambda I - A|$ 极小式：$g(x)$

1. $g(x)$有重根

（1）一个 k 重单根 $g(x) = (x - b)^k$

公式：$f(A) = f(b)G+f'(b)F_1+f''(b)F_2+\ldots+f^{(k-1)}(b)F_{k-1}$

其中：$G, F_1, F_2, \ldots, F_{k-1}$ 固定矩阵（待定），$f(x)$为任意解析式

（2）$g(x) = (x-a)(x-b)^2$

公式：$f(A) = f(a)G_1 + f(b)G_2 + f'(b)F_1$

（3）$g(x) = (x-a)^2(x-b)^2$

公式：$f(A) = f(a)G_1 + f'(a)F_1 + f(b)G_2 + f'(b)F_2$

（4）一般地 $g(x) = (x-\lambda_1)^{k_1}(x-\lambda_2)^{k_2}\cdots(x-\lambda_s)^{k_s}$

2. $g(x)$无重根

$g(x) = (\mathrm{x}-\lambda_1)(\mathrm{x}-\lambda_2)\ldots(\mathrm{x}-\lambda_s)$

公式：$f(A) = f(\lambda_1)G_1 + f(\lambda_2)G_2 + \ldots + f(\lambda_s)G_s$

例 1：$A = \begin{pmatrix} 2 & 0 & 0 \\ 1 & -1 & 1 \\ 1 & -1 & 3 \end{pmatrix}$求 e^{At}

解：$|\lambda I - A| = (x-2)^3$ 极小多项式：$g(x) = (x-2)^2$

公式：$f(A) = f(2)G + f'(2)F$ 其中 G、F 为待定阵

令 $f(x) = x-2$ $f'(x) = 1$，则 $f(A) = A-2I$ $f(2) = 0$ $f'(2) = 1$

$\therefore F = A - 2I$

令 $f(x) = 1$ $f'(x) = 0$，则 $f(A) = I$ $f(2) = 1$ $f'(2) = 0$

$\therefore G = I$

$\therefore f(A) = f(2)I + f'(2)(A-2I)$

当 $f(x) = e^{xt}$ 时，$f'(x) = te^{xt}$ $f(2) = e^{2t}$ $f'(2) = te^{2t}$

$$e^{At} = e^{2t}I + te^{2t}(A-2I) = e^{2t}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + te^{2t}\begin{pmatrix} 0 & 0 & 0 \\ 1 & -1 & 1 \\ 1 & -1 & 1 \end{pmatrix} = e^{2t}\begin{pmatrix} 1 & 0 & 0 \\ t & 1-t & 1 \\ t & -t & 1+t \end{pmatrix}$$

例 2：$A = \begin{pmatrix} 2 & 1 & 4 \\ 0 & 2 & 0 \\ 0 & 3 & 1 \end{pmatrix}$求 e^{At}

解：特征多项式：$|\lambda I - A| = (x-1)(x-2)^2$

极小多项式：$g(x) = (x-1)(x-2)^2$

公式 $f(A) = f(1)G_1 + f(2)G_2 + f'(2)F$

令 $f(x) = (x-2)^2$，则 $f'(x) = 2(x-2)$ $f(1) = 1$ $f(2) = 0$ $f'(2) = 0$

$\therefore G_1 = (A-2I)^2$

令 $f(x) = (x-1)(x-2)$，则 $f'(x) = 2x-3$　$f(1) = 0$　　$f(2) = 0$　　$f'(2) = 1$

$\therefore F = (A-I)(A-2I)$

令 $f(x) = (x-1)$，则 $f'(x) = 1$　　$f(1) = 0$　　$f(2) = 1$　　$f'(2) = 1$

$\therefore A - I = G_2 + F \Rightarrow G_2 = (A-I)(3I-A)$

$\therefore f(A) = f(1)(A-2I)^2 + f(2)(A-I)(3I-A) + f'(2)(A-I)(A-2I)$

当 $f(x) = e^{xt}$ 时，$f'(x) = te^{xt}$　$f(1) = e^t$ $f(2) = e^{2t}$　$f'(2) = te^{2t}$

$\therefore e^{At} = e^t(A-2I)^2 + e^{2t}(A-I)(3I-A) + te^{2t}(A-I)(A-2I)$

$$= e^t\begin{pmatrix} 0 & 1 & 4 \\ 0 & 0 & 0 \\ 0 & 3 & -1 \end{pmatrix}^2 + e^{2t}\begin{pmatrix} 1 & 1 & 4 \\ 0 & 1 & 0 \\ 0 & 3 & 0 \end{pmatrix}\begin{pmatrix} 1 & -1 & -4 \\ 0 & 1 & 0 \\ 0 & -3 & 2 \end{pmatrix} + te^{2t}\begin{pmatrix} 1 & 1 & 4 \\ 0 & 0 & 0 \\ 0 & 3 & -1 \end{pmatrix}\begin{pmatrix} 0 & 1 & 4 \\ 0 & 0 & 0 \\ 0 & 3 & -1 \end{pmatrix}$$

$$= e^t\begin{pmatrix} 0 & 12 & -4 \\ 0 & 0 & 0 \\ 0 & -3 & -1 \end{pmatrix} + e^{2t}\begin{pmatrix} 1 & -12 & 4 \\ 0 & 1 & 0 \\ 0 & 3 & 0 \end{pmatrix} + te^{2t}\begin{pmatrix} 0 & 13 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

例：$A = \begin{pmatrix} 2 & 2 & 0 \\ -1 & 1 & 0 \\ -1 & -2 & 1 \end{pmatrix}$ 求 e^{At}

解：特征多项式：$|\lambda I - A| = x(x-1)^2$

极小多项式：$g(x) = x(x-1)$

公式：$f(A) = f(0)G_1 + f(1)G_2$

令 $f(x) = x-1$，则 $f(0) = -1$　$f(1) = 0$

$\therefore G_1 = I - A$

令 $f(x) = x$，则 $f(0) = 0$　$f(1) = 1$

$\therefore G_2 = A$

$\therefore f(A) = f(0)(I-A) + f(1)A$

当 $f(x) = e^{xt}$ 时，$f(0) = 1$　$f(1) = e^t$

$\therefore f(A) = (I-A) + e^tA$

习题：1. $A = \begin{pmatrix} -1 & 1 & 0 \\ -4 & 3 & 0 \\ 1 & 0 & 2 \end{pmatrix}$，求 $f(A)$的公式，再求 e^{At}

2. $A=\begin{pmatrix}3 & 0 & 8\\ 3 & -1 & 6\\ -2 & 0 & -5\end{pmatrix}$，求 $f(A)$的公式，再求 e^{At}

3. $A=\begin{pmatrix}2 & 1 & 1\\ 2 & 3 & 2\\ 1 & 1 & 2\end{pmatrix}$，求 $f(A)$的公式，再求 e^{At}

4. $A=\begin{pmatrix}\begin{pmatrix}-1 & 1\\ 0 & -2\end{pmatrix} & \begin{matrix}0 & 0\\ 0 & 0\end{matrix}\\ \begin{matrix}0 & 0\\ 0 & 0\end{matrix} & \begin{pmatrix}-1 & 1\\ 0 & -2\end{pmatrix}\end{pmatrix}$，求 $f(A)$

例：设 A 有极小多项式 $g(x)=(x-b)^k$　（$k\le n$）

求公式 $f(A)=f(b)G+f'(b)F_1+f''(b)F_2+\ldots+f^{(k-1)}(b)F_{k-1}$ 中

待定矩阵 $G, F_1, F_2, \ldots, F_{k-1}$

解：$f(x)\equiv 1$　　$f'(b)=f''(b)=\quad=f^{(k-1)}(b)=0$

$\therefore G=I$

$f(x)=(x-b)$　　$f'(b)=1$　　$f(b)=f''(b)=\quad=f^{(k-1)}(b)=0$

$\therefore F_1=A-bI$

$f(x)=(x-b)^2$　　$f''(b)=2$　　$f(b)=f'(b)=\quad=f^{(k-1)}(b)=0$

$$\therefore (A-bI)^2=2F_2\Rightarrow F_2=\frac{(A-bI)^2}{2!}$$

以此类推：$F_{k-1}=\dfrac{(A-bI)^{k-1}}{(k-1)!}$

$$\therefore f(A)=f(b)I+f'(b)(A-bI)+f''(b)\frac{(A-bI)^2}{2!}+\cdots+f^{(k-1)}(b)\frac{(A-bI)^{k-1}}{(k-1)!}$$

当 $A=J=\begin{pmatrix}b & 1 & & & \\ & b & 1 & & \\ & & b & \ddots & \\ & & & \ddots & 1\\ & & & & b\end{pmatrix}_{n\times n}$ 时，极小式：$g(x)=(x-b)^n$

$$D=A-bI=\begin{pmatrix}0&1&&&\\&0&1&&\\&&0&\ddots&\\&&&\ddots&1\\&&&&0\end{pmatrix}\Rightarrow D^2=\begin{pmatrix}0&0&1&&\\&0&0&\ddots&\\&&0&\ddots&1\\&&&\ddots&0\\&&&&0\end{pmatrix}\Rightarrow D^{n-1}=\begin{pmatrix}0&0&0&\cdots&1\\&0&0&\ddots&\vdots\\&&0&\ddots&0\\&&&\ddots&0\\&&&&0\end{pmatrix}$$

$$f(A)=f(b)I+f'(b)D+f''(b)\frac{D^2}{2!}+\cdots+f^{(k-1)}(b)\frac{D^{k-1}}{(k-1)!}$$

$$=\begin{pmatrix}f(b)&f'(b)&\frac{f''(b)}{2!}&\cdots&\frac{f^{(k-1)}(b)}{(k-1)!}\\&f(b)&f'(b)&\ddots&\vdots\\&&f(b)&\ddots&\frac{f''(b)}{2!}\\&&&\ddots&f'(b)\\&&&&f(b)\end{pmatrix}$$

引理：$\|A\|<1$ 或 $\rho(A)<1$，则$(I-A)$可逆

令 $B=I-A$，（$A=I-B$）

引理：若$\|I-B\|<1$，则 B 可逆

积的求导公式：

$$\frac{d}{dt}\big(A(t)B(t)C(t)\big)=\frac{dA(t)}{dt}B(t)C(t)+\frac{dB(t)}{dt}A(t)C(t)+\frac{dC(t)}{dt}A(t)B(t)$$

指数求导公式：$\dfrac{de^{At}}{dt}=e^{At}A=Ae^{At}$

Pf：$e^{At}=I+At+\dfrac{A^2t^2}{2!}+\dfrac{A^3t^3}{3!}+\cdots$

$$\frac{d}{dt}\left(e^{At}\right)=0+A+\frac{A^2t}{1!}+\frac{A^3t^2}{2!}+\frac{A^4t^3}{3!}\cdots$$

$$=A\left(I+At+\frac{A^2t^2}{2!}+\frac{A^3t^3}{3!}+\cdots\right)=\left(I+At+\frac{A^2t^2}{2!}+\frac{A^3t^3}{3!}+\cdots\right)A$$

$$=Ae^{At}=e^{At}A$$

Euler 公式：（$i=\sqrt{-1}$）

$e^{iA}=\cos A+i\sin A$　　（利用 $e^{iA}=\sum\limits_{k=0}^{\infty}\dfrac{(iA)^k}{k!}$）

用$-A$代替A：$e^{iA}=\cos A-i\sin A$　两式相加减

$$\cos A=\frac{e^{iA}+e^{-iA}}{2} \qquad \sin A=\frac{e^{iA}-e^{-iA}}{2}$$

$$\frac{d\cos At}{dt}=-(\sin At)A=-A(\sin At) \qquad \frac{d\sin At}{dt}=(\cos At)A=A(\cos At)$$

***f*(*A*)的算法：**

1. A可对角化，极小式$g(x)=(x-\lambda_1)(x-\lambda_2)\ldots(x-\lambda_s)$无重根

$\Rightarrow f(A)=f(\lambda_1)G_1+f(\lambda_2)G_2+\cdots+f(\lambda_s)G_s$

$G_1, G_2, \ldots, G_s$（固定）可用公式$G_i=\dfrac{g_i(A)}{g_i(\lambda_i)}$求出；$f(x)$为任一解析式

2. A不可对角化，极小式$g(x)$有重根，可用“待定矩阵法”（广义谱分解公式）

例如：A的极小式$g(x)=(x-b)^2$

可令公式：$f(A)=f(b)G+f'(b)F$；G、F固定（待定）$f(x)$为任一解析式

分别令$f(x)\equiv 1$、$f(x)=(x-b)$代入$\Rightarrow G=I$、$F=(A-bI)$

$\Rightarrow f(A)=f(b)I+f'(b)(A-bI)$

再令$f(x)=e^{tx}$　$f'(x)=te^{tx}\Rightarrow f(A)=e^{At}=\cdots$

对角公式：设$D=\begin{pmatrix} b_1 & & O \\ & b_2 & & \\ & & \ddots & \\ O & & & b_n \end{pmatrix}$，则$f(D)=\begin{pmatrix} f(b_1) & & O \\ & f(b_2) & & \\ & & \ddots & \\ O & & & f(b_n) \end{pmatrix}$

Jordan块公式$D=\begin{pmatrix} b & 1 & & \\ & b & \ddots & \\ & & \ddots & 1 \\ & & & b \end{pmatrix}_{p\times p}$　（极小式$g(x)=(x-b)^p$）

$$\text{则 } f(D)=\begin{pmatrix} f(b) & f'(b) & \frac{f''(b)}{2!} & \cdots & \frac{f^{(p-1)}(b)}{(p-1)!} \\ & f(b) & f'(b) & \ddots & \vdots \\ & & f(b) & \ddots & \frac{f''(b)}{2!} \\ & & & \ddots & f'(b) \\ & & & & f(b) \end{pmatrix}$$

eg. $B=\begin{pmatrix}\begin{pmatrix}2&1\\0&2\end{pmatrix}&\begin{matrix}0\\0\end{matrix}\\\begin{matrix}0&0\end{matrix}&(0)\end{pmatrix}$，求 e^B，令 $f(x)=e^x$，$f'(x)=e^x$

$$e^B=\begin{pmatrix}\begin{pmatrix}e^2&e\\0&e^2\end{pmatrix}&\begin{matrix}0\\0\end{matrix}\\\begin{matrix}0&0\end{matrix}&(1)\end{pmatrix}$$

引理：（1）若 $AB=BA$，则 $e^Ae^B=e^Be^A=e^{A+B}$

（2）$e^{-A}e^A=e^Ae^{-A}=e^0=I$

（3）$AB=BA\Rightarrow e^{At}e^{Bt}=e^{Bt}e^{At}=e^{(A+B)t}$

线性微分方程：

引理： 若 $\dfrac{dA(t)}{dt}=A'(t)\equiv 0$，则 $A(t)=C$（常值矩阵）

1. $\dfrac{dx}{dt}=Ax$，$x(0)=C$，$x=(x_1(t),x_2(t),\quad\ldots,x_n(t))^T$，$A=A_{n\times n}$
2. $\dfrac{dY}{dt}=AY$，$Y(0)=C_{n\times n}$，$Y=(y_{ij}(t))$
3. $\dfrac{dY}{dt}=YA$，$Y(0)=C_{n\times n}$，$Y=(y_{ij}(t))$
4. $\dfrac{dY}{dt}=AY+YB$，$Y(0)=F_{n\times n}$

引理： $\dfrac{dY}{dt}=AY+YB$，$Y(0)=F$ 的解公式为：$Y=e^{At}Fe^{Bt}$

Pf：利用 $\dfrac{de^{-At}Y}{dt}=-Ae^{-At}Y+e^{-At}\dfrac{dY}{dt}$，$\dfrac{dYe^{-Bt}}{dt}=\dfrac{dY}{dt}e^{-Bt}-Ye^{-Bt}B$

若 Y 为 $\dfrac{dY}{dt}=AY+YB$ 的解

$$\Rightarrow e^{-At}\frac{dY}{dt}e^{-Bt}=Ae^{-At}Ye^{-Bt}+e^{-At}Ye^{-Bt}B$$

$$\Rightarrow e^{-At}\frac{dY}{dt}e^{-Bt}-\left(Ae^{-At}\right)Ye^{-Bt}-e^{-At}Y\left(e^{-Bt}B\right)=0$$

$$\Rightarrow\frac{d}{dt}\left(e^{-At}Ye^{-Bt}\right)\equiv 0\Rightarrow e^{-At}Ye^{-Bt}\equiv C$$

令 $t=0\Rightarrow e^0Y(0)e^0\equiv C\Rightarrow IY(0)I=C$

$\Rightarrow C=Y(0)=F\Rightarrow Y=e^{At}Ce^{Bt}=e^{At}Fe^{Bt}$（有唯一解）

再验证：$Y=e^{At}Fe^{Bt}$适合$\dfrac{dY}{dt}=AY+YB$ （且 $Y(0)=F$）

特别： $B=0\,B=0$时，$\dfrac{dY}{dt}=AY$有解 $Y=e^{At}F$

$A=0$ 时，$\dfrac{dY}{dt}=AY$有解 $Y=e^{At}F$

同理：可知$\dfrac{dx}{dt}=Ax$，$x(0)=C$ ，$x=(x_1,x_2,\quad \dots,x_n)\in R^n$，有唯一解$x=e^{At}\vec{C}$

Ex.1. $A=\begin{pmatrix}2&0&0\\1&1&1\\1&-1&3\end{pmatrix}$，求解齐次方程$\dfrac{dx}{dt}=Ax$，$x(0)=\begin{pmatrix}1\\1\\1\end{pmatrix}$

2. $A=\begin{pmatrix}2&2&1\\-2&6&1\\0&0&4\end{pmatrix}$，求解齐次方程$\dfrac{dx}{dt}=Ax$，$x(0)=\begin{pmatrix}1\\0\\0\end{pmatrix}$

§6 广义逆A^+

A^+定义： 若$A=A_{m\times n}$，若有$X=X_{n\times m}$使得：

（1）$AXA=A$ （2）$XAX=X$ （3）$(AX)^H=AX$ （4）$(XA)^H=XA$

（(3)、(4）为 Hermite 条件）

则称X为A的一个“+号”广义逆，记为A^+

特别： $A=A_{n\times n}$为可逆时，$A^+=A^{-1}$

$A=0_{n\times n}$时，$A^+=0_{n\times n}$

唯一定理： 若$A=A_{m\times n}$则，A^+是唯一的

（若有 2 个A^+：X与Y适合（1）（2）（3）（4）$\Rightarrow X=Y$）

A^+存在公式： 设短奇异分解：$A=P_1\Delta Q_1^H$ $\Delta=\begin{pmatrix}\sqrt{\lambda_1}&&&O\\&\sqrt{\lambda_2}&&\\&&\ddots&\\O&&&\sqrt{\lambda_r}\end{pmatrix}$

$r=\operatorname{rank}(A)$ $P_1^HP_1=I_r$ $Q_1^HQ_1=I_r$

则$A^+=Q_1\Delta^{-1}P_1^H$

Check： （1）$AA^+A=\left(P_1\Delta Q_1^HQ_1\Delta^{-1}P_1^H\right)P_1\Delta Q_1^H=P_1\Delta Q_1^H=A$

（2）$A^+AA^+=A^+$

（3） $AA^+ = P_1P_1^H$ （Hermite）

（4） $A^+A = Q_1Q_1^H$ （Hermite）

公式：若 $A = P_1\Delta Q_1^H$，则 $A^+ = Q_1\Delta^{-1}P_1^H$ （短奇异分解）

Ex. $A = \begin{pmatrix} 1 & 1 \\ 2 & 2 \\ 0 & 0 \end{pmatrix}$，利用奇异（短）分解求 A^+

（M-B）广义逆 A+

（Moore，Benrose）A^+的 4 个条件：$A = A_{m\times n}$ $A^+ = (\ldots)_{n\times m}$

（1）$AA^+A = A$ （2）$A^+AA^+ = A^+$ （3）$(AA^+) = (AA^+)^H$（4）$(A^+A) = (A^+A)^H$

注：A^+是唯一的

A^+奇异分解公式：设 $A = P_1\Delta Q_1^H$ （短分解），则 $A^+ = Q_1\Delta^{-1}P_1^H$

Eg. $A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \\ 1 & 1 \end{pmatrix}$ $\lambda_1 = 4$，$\sqrt{\lambda_1} = 2$ $\Delta = (2)$

$$A = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{pmatrix}(2)\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) \Rightarrow A^+ = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}(2)^{-1}\left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right) = \frac{1}{4}\begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{4} & 0 & \frac{1}{4} \\ \frac{1}{4} & 0 & \frac{1}{4} \end{pmatrix}$$

注：奇异值对误差不敏感（很稳定）

引理：齐次方程 $Ax = 0$ 的通解为：$\xi = (I_n - A^+A)y$，$\forall y \in C^n$

Pf：若 $\xi = (I_n - A^+A)y \Rightarrow A\xi = A(I_n - A^+A)y = (A - AA^+A)y = (A - A)y = 0$

另外：若 ξ 为 $Ax = 0$ 的解，$A\xi = 0$

写 $\xi = \xi - 0 = I_n\xi - A^+A\xi \Rightarrow \xi = (I_n - A^+A)\xi$ （适合公式）

引理：非齐次方程 $Ax = b$（若有解）通解为：

$$x = (A^+b) + (I_n - A^+A)y \quad \forall y$$

Pf：只须证 $x_0 = A^+b$ 为 $Ax = b$ 的解

$\because Ax_0 = A(A^+b)$ $\because Ax = b$ 有解

可写 $b = A\xi$ 代入

$\Rightarrow Ax_0 = A(A^+b) = AA^+(A\xi) = (AA^+A)\xi = b$

注： $AA^+b = b \Leftrightarrow Ax = b$ 有解

问： 若 $Ax = b$ 有解，哪个解的长度最佳（最小）

$x = (x_1, x_2, \ldots, x_n)^T \in C^n$ $|x|^2 = x^H x = |x_1|^2 + |x_2|^2 + \ldots + |x_n|^2$

引理： 令 $x_0 = A^+b$，则 $x_0 \perp y$ $y \in \mathscr{N}(A) = \{x|Ax = 0\}$

Pf：$\because$ 内积 $(y, x_0) = x_0^H y = (A^+b)^H y = [((A^+A)A^+)b]^H y$

$= (A^+b)^H (A^+A)^H y = (A^+b)^H (A^+A) y = 0$

$\Rightarrow x_0 \perp y$ （$x_0 \perp \mathscr{N}(A)$）

最小长度解公式：

若 $Ax = b$ 有解，则 $x_0 = A^+b$ 为最小长度解

Pf：用 $Ax = b$ 通解 $x = x_0 + y$ $\forall y \in \mathscr{N}(A)$

$\because x_0 \perp y$ 用勾股定理：$|x|^2 = |x_0 + y|^2 = |x_0|^2 + |y|^2 \geq |x_0|^2$

$\therefore |x_0|^2$ 为最小

Q：若 $Ax = b$ 无解（矛盾方程）

求一个 x_0 使得 $|Ax_0 - b|^2$ 为最小，即 $|Ax_0 - b|^2 = \min\{|Ax - b|^2 \,|\, x \in C^n\}$

令值域（相空间）$\mathscr{R}(A) = \{Ay|y \in C^n\}$（为线形空间）$\subset C^m$

引理： $\forall A = A_{m\times n}$，$\forall b \in C^m$，则 $x_0 = A^+b$ 适合 $A^H x_0 - A^H b = 0$

（即 $A^H Ax = A^H b$ 有一个解 $x_0 = A^+b$）

Pf：$A^H Ax_0 = A^H AA^+b = A^H (AA^+)^H b = (AA^+A)^H b = (A)^H b \Rightarrow A^H Ax_0 = A^H b$

推论： $\forall A = A_{m\times n}$，$\forall b \in C^m \Rightarrow A^H Ax = A^H b$ 必有解（$x_0 = A^+b$）

极小二乘解定理： $x_0 = A^+b$ 为 $Ax = b$ 的一个极小二乘解

引理： 设 $x_0 = A^+b$，则 $(Ax_0 - b) \perp Ay \ \forall y$

Pf：内积：$(Ax_0-b, Ay)=(Ay)^H(Ax_0-b)=y^HA^H(Ax_0-b)=y^H(A^HAx_0-A^Hb)=0$

定理 Pf： $\because \mathrm{Ax-b}=(Ax_0-b)+(Ax-Ax_0)=(Ax_0-b)+A(x-x_0)$

$\because (Ax_0-b)\perp A(x-x_0)$ 勾股定理

$\Rightarrow |Ax-b|^2=|Ax_0-b|^2+|A(x_0-x)|^2\geq|Ax_0-b|^2$（为极小值）

“小二解”公式： $Ax=b$ 的全体小二解为 $x=x_0+y \quad y\in \mathcal{N}(A)$

或 $x=x_0+(I_n-A^+A)y \quad \forall y$ （$Ay=0$）

$\because Ax-b=A(x_0+y)-b=Ax_0+Ay-b=Ax_0-b\Rightarrow|Ax-b|^2=|Ax_0-b|^2$ （极小）

最佳小二解（唯一）： $x_0=A^+b$ 是 $Ax=b$ 的最小长度“小二解”

Pf：$\because$全体小二解为 $x=x_0+y$ （$Ay=0$）

利用 $x_0\perp y\Rightarrow|x|^2=|x_0|^2+|y|^2\geq|x_0|^2$ （最佳）

注： 若 A 为高阵 $\mathrm{rank}(A)=n$，则只有解小二解

Ex.P3152、3、4（6）

Ex.（1） $A=\begin{pmatrix}1&2\\0&0\\2&4\end{pmatrix}$，$b=\begin{pmatrix}0\\1\\0\end{pmatrix}$，求 $Ax=b$ 的最佳小二解

（2）A 同上，$b=\begin{pmatrix}1\\0\\2\end{pmatrix}$，用 A^+表示 $Ax=b$ 的通解

A^+性质与计算（$P_1^HP_1=I_r=Q_1^HQ_1$）

1. **A^+短分解公式：** 若 $A=P_1\Delta Q_1^H$，$\Delta=\begin{pmatrix}\sqrt{\lambda_1}&&&O\\&\sqrt{\lambda_2}&&\\&&\ddots&\\O&&&\sqrt{\lambda_r}\end{pmatrix}$，则 $A^+=Q_1\Delta^{-1}P_1^H$

满分解公式： $A=BC$（满分解）$\Rightarrow A^+=C^+B^+$ （$\begin{matrix}C^+=C_R=C^H(CC^H)^{-1}\\B^+=B_L=(B^HB)^{-1}B^H\end{matrix}$）

性质：（4 个条件：$AA^+A=A \quad A^+AA^+=A^+ \quad AA^+=(AA^+)^H \quad A^+A=(A^+A)^H$）

$\mathrm{rank}(A^+)=\mathrm{rank}(A)=\mathrm{rank}(A^H) \quad \mathrm{rank}(AA^+)=\mathrm{rank}(A^+A)=\mathrm{rank}(A)$

Pf：$A^+=A^+AA^+\Rightarrow rank(A^+)\leq rank(A)$，同理：$\mathrm{rank}(A)\leqslant\mathrm{rank}(A^+)$

A^+A、AA^+都半正定

Pf: $\because \left(AA^+\right)^2 = AA^+ \Rightarrow f(x) = x^2 - x$ 为 AA^+的 0 化式

AA^+的特征值只有 1 或 0（含重复）

$$\Rightarrow AA^+ \backsim \begin{pmatrix} r\text{个}\left\{\begin{matrix} 1 & & \\ & \ddots & \\ & & 1 \end{matrix}\right. & & & O \\ & 0 & & \\ O & & \ddots & \\ & & & 0 \end{pmatrix}$$，且 AA^+为 Hermite$\Rightarrow AA^+$ 为半正定

$(A^+)^+ = \mathrm{A}$

$(kA)^+ = \dfrac{1}{k}A^+ \qquad (k \neq 0)$

$(D,0)^+ = \begin{pmatrix} D^+ \\ 0 \end{pmatrix} \quad \begin{pmatrix} D \\ 0 \end{pmatrix}^+ = \left(D^+,0\right)$

规定： 数 k 的 $k^+ \overset{\Delta}{=} \begin{cases} 0 & k=0 \\ \dfrac{1}{k} & k \neq 0 \end{cases}$ （特别：$0^+ = 0$）

对角公式： $\begin{pmatrix} k_1 & & & O \\ & k_2 & & \\ & & \ddots & \\ O & & & k_n \end{pmatrix}^+ = \begin{pmatrix} k_1^+ & & & O \\ & k_2^+ & & \\ & & \ddots & \\ O & & & k_n^+ \end{pmatrix}$

令 $D = \begin{pmatrix} k_1 & & & O \\ & k_2 & & \\ & & \ddots & \\ O & & & k_n \end{pmatrix} \Rightarrow D^+D = DD^+$

准对角公式： $\begin{pmatrix} A_1 & & & O \\ & A_2 & & \\ & & \ddots & \\ O & & & A_s \end{pmatrix}_{m\times n}^+ = \begin{pmatrix} A_1^+ & & & O \\ & A_2^+ & & \\ & & \ddots & \\ O & & & A_s^+ \end{pmatrix}_{n\times m}$

$\mathscr{U}$分解公式： 设 PQ 为 $\mathscr{U}$阵，则$(PAQ)^+ = Q^+A^+P^+ = Q^{-1}A^+P^{-1}$

QR 高阵分解公式： 设 $A = A_{m\times r}$ 为高阵，且 $A = QR$ （$R = R_{r\times r}$，Q 为次 $\mathscr{U}$）

则 $A^+ = R^+Q^+ = R^{-1}Q^H$

秩 1 公式： 若 $rank(A) = 1$，则 $A^+ = \dfrac{1}{tr\left(A^HA\right)}A^H = \dfrac{1}{\sum\left|a_{ij}\right|^2}A^H$

引理：$A=(a_{ij})_{m\times n}$，则

（1）$tr\left(A^HA\right)=tr\left(AA^H\right)=\sum\left|a_{ij}\right|^2=\|A\|_F^2$

（2）$\mathrm{rank}(A^HA)=\mathrm{rank}(AA^H)=\mathrm{rank}(A)$

（3）A^HA，AA^H 半正定

设 $A^HA\backsim\begin{pmatrix}\lambda_1 & & O\\ & \lambda_2 & & \\ & & \ddots & \\ O & & & \lambda_n\end{pmatrix}$　　$\lambda_1, \lambda_2, \ldots, \lambda_n\geq 0$

$\mathrm{tr}(A^HA)=\lambda_1+\lambda_2+\ldots+\lambda_n$

Pf：$rank(A)=1\Rightarrow A^HA\backsim\begin{pmatrix}\lambda_1 & & O\\ & 0 & & \\ & & \ddots & \\ O & & & 0\end{pmatrix}$　$\lambda_1>0$（$\sqrt{\lambda_1}$ 为奇异值）

$\Rightarrow\lambda_1=\lambda_1+0+\cdots+0=tr\left(A^HA\right)=\sum\left|a_{ij}\right|^2$

由短分解：$A=\begin{pmatrix}a_1\\ a_2\\ \vdots\\ a_m\end{pmatrix}\left(\sqrt{\lambda_1}\right)(b_1,b_2,\cdots,b_n)$　$p_1=\begin{pmatrix}a_1\\ a_2\\ \vdots\\ a_m\end{pmatrix}$　$q^H=(b_1, b_2, \ldots, b_n)$为次 $\mathscr{U}$

（$p_1^Hp_1=1=q_1^Hq_1$）　写 $A=p_1\left(\sqrt{\lambda_1}\right)q_1^H\Rightarrow A^H=q_1\left(\sqrt{\lambda_1}\right)^{-1}p_1^H$

且 $A^{-1}=q_1\left(\sqrt{\lambda_1}\right)^{-1}p_1^H=\dfrac{1}{\lambda_1}\left(q_1\left(\sqrt{\lambda_1}\right)p_1^H\right)$

$$A^+=\frac{1}{\lambda_1}A^H=\frac{1}{tr\left(A^HA\right)}A^H=\frac{1}{\sum\left|a_{ij}\right|^2}A^H$$

Eg. $A=\begin{pmatrix}1 & 2\\ 0 & 0\\ -1 & -2\end{pmatrix}$　$rank(A)=A^+=\dfrac{1}{1^2+2^2+1^2+2^2}A^H\Rightarrow A^+=\dfrac{1}{10}\begin{pmatrix}1 & 0 & -1\\ 2 & 0 & -2\end{pmatrix}$

$$\text{Eg. } A=\begin{pmatrix}\begin{pmatrix}1&1&0\\1&1&0\end{pmatrix} & O\\ O & \begin{pmatrix}1&2&-1\\-1&-2&1\end{pmatrix}\end{pmatrix}_{4\times 6} \Rightarrow A^+=\begin{pmatrix}\frac{1}{4}\begin{pmatrix}1&1\\1&1\\0&0\end{pmatrix} & O\\ O & \frac{1}{12}\begin{pmatrix}1&1\\2&-2\\-1&1\end{pmatrix}\end{pmatrix}_{6\times 4}$$

Ex.A 为正规$\Rightarrow A^+A=AA^+$，且$(A^+)^k=(A^k)^+\left(A^+\right)^k=\left(A^k\right)^+$

$$\because A=QDQ^H=Q\begin{pmatrix}\lambda_1&&&O\\&\lambda_2&&\\&&\ddots&\\O&&&\lambda_n\end{pmatrix}Q^H\Rightarrow A^+=QD^+Q^H=Q\begin{pmatrix}\lambda_1^+&&&O\\&\lambda_2^+&&\\&&\ddots&\\O&&&\lambda_n^+\end{pmatrix}Q^H$$

Eg. $A=\begin{pmatrix}1&0\\0&1\\1&1\end{pmatrix}$ $\operatorname{rank}(A)=2$，求奇异值、A^+

$A^H=\begin{pmatrix}1&0&1\\0&1&1\end{pmatrix}$ $A^HA=\begin{pmatrix}2&1\\1&2\end{pmatrix}$ $\lambda_1=3$，$\lambda_2=1$ 奇异值为$\sqrt{3}$，1

$\lambda_1=3$ 的特征向量为$\begin{pmatrix}1\\1\end{pmatrix}$ $\lambda_2=1$ 的特征向量为$\begin{pmatrix}1\\-1\end{pmatrix}$

短奇异分解：$A=\begin{pmatrix}\frac{1}{\sqrt6}&\frac{1}{\sqrt2}\\\frac{1}{\sqrt6}&-\frac{1}{\sqrt2}\\\frac{2}{\sqrt6}&0\end{pmatrix}\begin{pmatrix}\sqrt3&0\\0&1\end{pmatrix}\begin{pmatrix}\frac{1}{\sqrt2}&\frac{1}{\sqrt2}\\\frac{1}{\sqrt2}&-\frac{1}{\sqrt2}\end{pmatrix}$

$$\Rightarrow A^+=\begin{pmatrix}\frac{1}{\sqrt2}&\frac{1}{\sqrt2}\\\frac{1}{\sqrt2}&-\frac{1}{\sqrt2}\end{pmatrix}\begin{pmatrix}\left(\sqrt3\right)^{-1}&0\\0&1\end{pmatrix}\begin{pmatrix}\frac{1}{\sqrt6}&\frac{1}{\sqrt2}\\\frac{1}{\sqrt6}&-\frac{1}{\sqrt2}\\\frac{2}{\sqrt6}&0\end{pmatrix}=\frac{1}{\sqrt2}\begin{pmatrix}1&1\\1&-1\end{pmatrix}\begin{pmatrix}\frac{1}{\sqrt{18}}&\frac{1}{\sqrt{18}}&\frac{2}{\sqrt{18}}\\\frac{1}{\sqrt2}&-\frac{1}{\sqrt2}&0\end{pmatrix}$$

$$=\frac{1}{6}\begin{pmatrix}1&1\\1&-1\end{pmatrix}\begin{pmatrix}1&1&2\\3&-3&0\end{pmatrix}=\frac{1}{6}\begin{pmatrix}4&-2&2\\-2&4&2\end{pmatrix}$$

解法 2：用满分解 $A=\begin{pmatrix}1&0\\0&1\\1&1\end{pmatrix}(I_r)\Rightarrow A^+=A_L=\left(A^HA\right)^{-1}A^H$

$$A^+=\begin{pmatrix}2&1\\1&2\end{pmatrix}^{-1}A^H=\frac{1}{3}\begin{pmatrix}2&-1\\-1&2\end{pmatrix}\begin{pmatrix}1&0&1\\0&1&1\end{pmatrix}=\frac{1}{3}\begin{pmatrix}2&-1&1\\-1&2&1\end{pmatrix}$$

$$\begin{pmatrix}a&b\\c&d\end{pmatrix}^{-1}=\frac{1}{ad-bc}\begin{pmatrix}d&-b\\-c&a\end{pmatrix}$$

Eg. $A=\begin{pmatrix}1&3\\0&0\\2&1\end{pmatrix}$，$\text{rank}(A)\neq 1$

$$\Rightarrow A=QR=\frac{1}{\sqrt{5}}\begin{pmatrix}1&3\\0&0\\2&1\end{pmatrix}R\Rightarrow R=Q^TA=\begin{pmatrix}\sqrt{5}&\sqrt{5}\\0&\sqrt{5}\end{pmatrix}=\sqrt{5}\begin{pmatrix}1&1\\0&1\end{pmatrix}$$

（Q 为次 $\mathscr{U}$，$Q^HQ=I_r$）

$$\Rightarrow A^+=Q^+R^+=Q^{-1}R^H=\left(\frac{1}{\sqrt{5}}\right)\begin{pmatrix}1&1\\0&1\end{pmatrix}^{-1}\frac{1}{\sqrt{5}}\begin{pmatrix}1&0&2\\2&0&-1\end{pmatrix}=\frac{1}{5}\begin{pmatrix}-1&0&3\\2&0&-1\end{pmatrix}$$

公式：$A^+=(A^HA)^+A^H$，$A^+=A^H(AA^H)^+(A^H)^+=(A^+)^H$

Pf：用短分解：$A=P_1\Delta Q_1^H\quad A^H=Q_1(\Delta)P_1^H\quad A^+=Q_1(\Delta)^{-1}P_1^H$

$$A^HA=\left(Q_1\Delta P_1^H\right)\left(P_1\Delta Q_1^H\right)=Q_1\Delta^2Q_1^H$$

$$\Rightarrow\left(A^HA\right)^{-1}=Q_1\Delta^{-2}Q_1^H\Rightarrow\left(A^HA\right)^+A^H=Q_1\Delta^{-2}Q_1^H\left(Q_1\Delta P_1^H\right)$$

应用：$\because A^HA$ 为 Hermite 阵

$A^HA=\lambda_1G_1+\lambda_2G_2+\ldots+\lambda_sG_s$（$\lambda_1,\lambda_2,\ldots,\lambda_s$ 互异）

（$G_1^H=G_1,G_2^H=G_2,\cdots,G_s^H=G_s$）

$$\Rightarrow\left(A^HA\right)^+=\lambda_1^+G_1+\lambda_2^+G_2+\cdots+\lambda_s^+G_s$$
$$\Rightarrow A^+=\left(A^HA\right)^+A^H=\left(\lambda_1^+G_1+\lambda_2^+G_2+\cdots+\lambda_s^+G_s\right)A^H$$

Ex. 若 A 为正规（或 Hermite）且 有谱分解 $A=\lambda_1G_1+\lambda_2G_2+\ldots+\lambda_sG_s$

（$G_1^H=G_1,G_2^H=G_2,\cdots,G_s^H=G_s$）

则 $A^+=\lambda_1^+G_1+\lambda_2^+G_2+\cdots+\lambda_s^+G_s$

但： 对一般可对角阵 $A=\lambda_1G_1+\lambda_2G_2+\ldots+\lambda_sG_s$ ， $A^+\neq\lambda_1^+G_1+\lambda_2^+G_2+\cdots+\lambda_s^+G_s$

Ex. $A=(1,1)$， $B=\begin{pmatrix}1\\0\end{pmatrix}$，验证$(AB)^+\neq B^+A^+$

但对满分解 $A=BC\Rightarrow A^+=(BC)^+=C^+B^+$

§7 直积拉直及应用

定义： 设 $A=(a_{ij})_{m\times n}$ $B=(b_{ij})_{p\times q}$

规定 $A\otimes B\triangleq\begin{pmatrix}a_{11}B & a_{12}B & \cdots & a_{1n}B\\ a_{21}B & a_{22}B & \cdots & a_{2n}B\\ \vdots & \vdots & \ddots & \vdots\\ a_{n1}B & a_{n2}B & \cdots & a_{nn}B\end{pmatrix}_{(mp)\times(nq)}$ 叫 A 与 B 的直积（张量积）

注： $A\otimes B\underline{\underline{记为}}(a_{ij}B)$

Eg. $A=\begin{pmatrix}a & b\\ c & d\end{pmatrix}$ $B=\begin{pmatrix}2\\3\end{pmatrix}$

$$A\otimes B=\begin{pmatrix}aB & bB\\ cB & dB\end{pmatrix}=\begin{pmatrix}2a & 2b\\ 3a & 3b\\ 2c & 2d\\ 3c & 3d\end{pmatrix}\quad B\otimes A=\begin{pmatrix}2A\\ 3A\end{pmatrix}=\begin{pmatrix}2a & 2b\\ 2c & 2d\\ 3a & 3b\\ 3c & 3d\end{pmatrix}$$

$A\otimes B\neq B\otimes A$

引理： $I_m\otimes I_n=I_{mn}=I_n\otimes I_m$

$$I_m=\begin{pmatrix}1 & & & O\\ & 1 & & \\ & & \ddots & \\ O & & & 1\end{pmatrix}_{m\times m}\Rightarrow I_m\otimes I_n=\begin{pmatrix}I_n & & & O\\ & I_n & & \\ & & \ddots & \\ O & & & I_n\end{pmatrix}$$

性质：

定理： （1）两个上三角阵直积也是上三角阵

（2）对角阵直积也是对角阵

（3） $I_m\otimes I_n=I_{mn}=I_n\otimes I_m$

Pf:（1） $A=\begin{pmatrix} a_1 & & & * \\ & a_2 & & \\ & & \ddots & \\ O & & & a_n \end{pmatrix}_{n\times n}$ $B=\begin{pmatrix} b_1 & & & * \\ & b_2 & & \\ & & \ddots & \\ O & & & b_p \end{pmatrix}_{p\times p}$

$$A\otimes B=\begin{pmatrix} a_1B & & & (*) \\ & a_2B & & \\ & & \ddots & \\ O & & & a_nB \end{pmatrix}=\begin{pmatrix} \begin{pmatrix} a_1b_1 & & \\ & \ddots & \\ & & a_1b_p \end{pmatrix} & & (*) \\ & \ddots & \\ O & & \begin{pmatrix} a_nb_1 & & \\ & \ddots & \\ & & a_nb_p \end{pmatrix} \end{pmatrix}$$

引理（分块公式）：$\begin{pmatrix} A & B \\ C & D \end{pmatrix}\otimes F=\begin{pmatrix} A\otimes F & B\otimes F \\ C\otimes F & D\otimes F \end{pmatrix}$

Pf：由定义：$\begin{pmatrix} A & B \\ C & D \end{pmatrix}\otimes F=\begin{pmatrix} (a_{ij}) & (b_{ij}) \\ (c_{ij}) & (d_{ij}) \end{pmatrix}\otimes F=\begin{pmatrix} (a_{ij}F) & (b_{ij}F) \\ (c_{ij}F) & (d_{ij}F) \end{pmatrix}=\begin{pmatrix} A\otimes F & B\otimes F \\ C\otimes F & D\otimes F \end{pmatrix}$

但是 $A\otimes(B_1,B_2)\neq(A\otimes B_1, A\otimes B_2)$

定理（转置公式）：（1）$(A\otimes B)^T=A^T\otimes B^T$ （2）$(A\otimes B)^H=A^H\otimes B^H$

Pf:（1）$\because (A\otimes B)^T=\begin{pmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}B & a_{n2}B & \cdots & a_{nn}B \end{pmatrix}^T=\begin{pmatrix} a_{11}B^T & a_{21}B^T & \cdots & a_{n1}B^T \\ a_{12}B^T & a_{22}B^T & \cdots & a_{n2}B^T \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n}B^T & a_{2n}B^T & \cdots & a_{nn}B^T \end{pmatrix}$

$$=(A^T)\otimes(B^T)=A^T\otimes B^T$$

Eg. $A=\begin{pmatrix} a & b \\ c & d \end{pmatrix}$，$B=\begin{pmatrix} 2 \\ 3 \end{pmatrix}\Rightarrow\begin{cases} A\otimes B=\begin{pmatrix} 2a & 2b \\ 3a & 3b \\ 2c & 2d \\ 3c & 3d \end{pmatrix} \\ (A\otimes B)^T=\begin{pmatrix} 2a & 3a & 2c & 3c \\ 2b & 3b & 2d & 3d \end{pmatrix} \end{cases}$

又 $A^T\otimes B^T=\begin{pmatrix} a & b \\ c & d \end{pmatrix}\otimes(2,3)=\begin{pmatrix} (2a,3a) & (2c,3c) \\ (2b,3b) & (2d,3d) \end{pmatrix}$

比较 $(AB)^T=B^TA^T$ $(AB)^H=B^HA^H$

吸收公式：$(A\otimes B)(C\otimes D)=(AC)\otimes(BD)$ （若 AC、BD 有意义）

Pf：$\left(A\otimes B\right)\left(C\otimes D\right)=\begin{pmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}B & a_{n2}B & \cdots & a_{nn}B \end{pmatrix}\begin{pmatrix} c_{11}D & c_{12}D & \cdots & c_{1n}D \\ c_{21}D & c_{22}D & \cdots & c_{2n}D \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1}D & c_{n2}D & \cdots & c_{nn}D \end{pmatrix}$

$=\left(\left(a_{i1}B\right)\left(c_{1j}D\right)+\left(a_{i2}B\right)\left(c_{2j}D\right)+\cdots+\left(a_{in}B\right)\left(c_{nj}D\right)\right)$

$=\left(\sum_{k=1}^{n}\left(a_{ik}Bc_{kj}D\right)\right)=\left(\sum_{k=1}^{n}\left(a_{ik}c_{kj}\right)\left(BD\right)\right)=\left(\sum_{k=1}^{n}\left(a_{ik}c_{kj}\right)\right)\left(BD\right)$

$\underline{\underline{\text{由定义}}}\left(AC\right)\otimes\left(BD\right)$

推论：（1）$\left(A_1\otimes B_1\right)\left(A_2\otimes B_2\right)\left(A_3\otimes B_3\right)=\left(A_1A_2A_3\right)\otimes\left(B_1B_2B_3\right)$

（1）’ $\left(A_1\otimes B_1\right)\left(A_2\otimes B_2\right)\cdots\left(A_k\otimes B_k\right)=\left(A_1A_2\cdots A_k\right)\otimes\left(B_1B_2\cdots B_k\right)$

（2）A、B 为方阵$\Rightarrow\left(A\otimes B\right)^k=A^k\otimes B^k$

$\because\left(A\otimes B\right)^k=\left(A\otimes B\right)\left(A\otimes B\right)\cdots\left(A\otimes B\right)=A^k\otimes B^k$

（3）$A=A_{m\times m}$、$B=B_{n\times n}\ \Rightarrow\left(A\otimes I_n\right)\left(I_m\otimes B\right)=A\otimes B=\left(I_m\otimes B\right)\left(A\otimes I_n\right)$

注：$\left(A\otimes I_n\right)$与$\left(I_m\otimes B\right)$可交换

逆公式：A、B 可逆$\Rightarrow\left(A\otimes B\right)^{-1}=A^{-1}\otimes B^{-1}$

Pf：$\because\left(A\otimes B\right)\left(A^{-1}\otimes B^{-1}\right)=\left(AA^{-1}\right)\otimes\left(BB^{-1}\right)=\left(I_m\right)\otimes\left(I_n\right)=I_{mn}\Rightarrow\left(A\otimes B\right)^{-1}=A^{-1}\otimes B^{-1}$

注：$\begin{cases}\left(A\otimes B\otimes C\right)^T=A^T\otimes B^T\otimes C^T \\ \left(A\otimes B\otimes C\right)^H=A^H\otimes B^H\otimes C^H \\ \left(A\otimes B\otimes C\right)^{-1}=A^{-1}\otimes B^{-1}\otimes C^{-1}\end{cases}$ （若 A、B、C 都可逆）

推论：若 A、B 为 $\mathscr{U}$ 阵，则$\left(A\otimes B\right)$也为 $\mathscr{U}$ 阵，且$\left(A\otimes B\right)^H=\left(A\otimes B\right)^{-1}$

$\because A^H=A^{-1}$、$B^H=B^{-1}$ （$\mathscr{U}$阵条件）

$\Rightarrow\left(A\otimes B\right)^H=A^H\otimes B^H=A^{-1}\otimes B^{-1}=\left(A\otimes B\right)^{-1}\Rightarrow\left(A\otimes B\right)$为 $\mathscr{U}$ 阵

或$\left(A\otimes B\right)^H\left(A\otimes B\right)=\left(A^H\otimes B^H\right)\left(A\otimes B\right)=\left(A^HA\right)\otimes\left(B^HB\right)=I_m\otimes I_n=I_{mn}$

比较**“穿脱公式”**：$\begin{cases}\left(ABC\right)^T=C^TB^TA^T \\ \left(ABC\right)^H=C^HB^HA^H \\ \left(ABC\right)^{-1}=C^{-1}B^{-1}A^{-1}\end{cases}$ （A、B、C 都可逆）

Eg.证明：（1）$e^{A\otimes I}=e^A\otimes I$ （2）$e^{I\otimes B}=I\otimes e^B$ （3）$e^{\left(A\otimes I+I\otimes B\right)}=e^A\otimes e^B$

Pf:（2）$e^{I\otimes B}\underset{=\!=}{\text{定义}}\sum_{k=0}^{\infty}\frac{(I\otimes B)^k}{k!}=\sum_{k=0}^{\infty}\frac{I^k\otimes B^k}{k!}=I\otimes\sum_{k=0}^{\infty}\frac{B^k}{k!}=I\otimes e^B$

（3）$\because XY=YX$（可交换）$\Rightarrow e^{X+Y}=e^Xe^Y=e^Ye^X$

$\because (A\otimes I_n)(I_m\otimes B)$ 可交换

$\Rightarrow e^{(A\otimes I+I\otimes B)}=e^{A\otimes I}e^{I\otimes B}=(e^A\otimes I)(I\otimes e^B)\underset{=\!=}{\text{吸收}}(e^A)\otimes(e^B)=e^A\otimes e^B$

秩公式： $rank(A\otimes B)=rank(A)\cdot rank(B)\quad A=A_{m\times n}$、$B=B_{p\times q}$

Pf：设 rank(A) = r、rank(B) = s

由标准形公式 $\Rightarrow\begin{cases}PAQ=\begin{pmatrix}I_r&0\\0&0\end{pmatrix}\underset{=\!=}{\text{记为}}A_1\\ \widetilde{P}B\widetilde{Q}=\begin{pmatrix}I_s&0\\0&0\end{pmatrix}=B_1\end{cases}$ P、Q、$\widetilde{P}$、$\widetilde{Q}$ 都可逆

$\Rightarrow (P\otimes\widetilde{P})(A\otimes B)(Q\otimes\widetilde{Q})=(PAQ)\otimes(\widetilde{P}B\widetilde{Q})=A_1\otimes B_1$

$$=\begin{pmatrix}I_r\otimes B_1&0\otimes B_1\\0\otimes B_1&0\otimes B_1\end{pmatrix}=\begin{pmatrix}\begin{pmatrix}r\text{个}\begin{cases}B_1&&&O\\&B_1&&\\&&\ddots&\\O&&&B_1\end{cases}\end{pmatrix}&O\\O&O\end{pmatrix}$$

$\Rightarrow rank(A_1\otimes B_1)=rs\Rightarrow rank(A\otimes B)=rank(A)\cdot rank(B)$

比较 rank(AB) $\le$ min{rank(A), rank(B)}

引理： 若 $A\backsim A_1$（相似），$B\backsim B_1$（相似），则 $A\otimes B\backsim A_1\otimes B_1$（相似）

$\because P^{-1}AP=A_1$，$Q^{-1}BQ=B_1\Rightarrow(P\otimes Q)^{-1}(A\otimes B)(P\otimes Q)=A_1\otimes B_1$

用许尔公式：$P^{-1}AP=A_1=\begin{pmatrix}\lambda_1&&&(*)\\&\lambda_2&&\\&&\ddots&\\O&&&\lambda_m\end{pmatrix}$（上三角阵）

$Q^{-1}BQ=B_1=\begin{pmatrix}t_1&&&(*)\\&t_2&&\\&&\ddots&\\O&&&t_n\end{pmatrix}$ 特征值为 $\sigma(B)=\{t_1,t_2,\ldots,t_n\}$

$$\Rightarrow A\otimes B \backsim A_1\otimes B_1=\begin{pmatrix} \lambda_1 B_1 & & & (*) \\ & \lambda_2 B_1 & & \\ & & \ddots & \\ O & & & \lambda_m B_1 \end{pmatrix}$$

$$\text{且 } A_1\otimes B_1=\begin{pmatrix} \begin{pmatrix} \lambda_1 t_1 & & * \\ & \ddots & \\ & & \lambda_1 t_n \end{pmatrix} & & * \\ & \ddots & \\ & & \begin{pmatrix} \lambda_m t_1 & & * \\ & \ddots & \\ & & \lambda_m t_n \end{pmatrix} \end{pmatrix}$$

引理： 设 $A = A_{m\times m}$ $\sigma(A) = \{\lambda_1, \lambda_2, \ldots, \lambda_m\}$

$B = B_{n\times n}$ $\sigma(B) = \{t_1, t_2, \ldots, t_n\}$

（1） $A\otimes B$ 的全体特征值为 mn 个$\{\lambda_i t_j\}$ $i = 1, 2, \ldots, m$ $j = 1, 2, \ldots, n$

（2） $A\otimes I_n + I_m\otimes B$ 的全体特征值为 mn 个$\{\lambda_i + t_j\}$ $i = 1, 2, \ldots, m$ $j = 1, 2, \ldots, n$

（3） $A\otimes I_n - I_m\otimes B$ 的全体特征值为 mn 个$\{\lambda_i - t_j\}$ $i = 1, 2, \ldots, m$ $j = 1, 2, \ldots, n$

Pf:（3）$\because (P\otimes Q)^{-1}(A\otimes I_n)(P\otimes Q)=(P^{-1}AP)\otimes(Q^{-1}I_nQ)=A_1\otimes I_n$

$(P\otimes Q)^{-1}(I_m\otimes B)(P\otimes Q)=(P^{-1}I_mP)\otimes(Q^{-1}BQ)=I_m\otimes B_1$

$\Rightarrow (P\otimes Q)^{-1}(A\otimes I_n+I_m\otimes B)(P\otimes Q)=A_1\otimes I_n+I_m\otimes B_1$

$\Rightarrow A\otimes I_n + I\otimes B_1 \backsim A_1\otimes I_n + I_m\otimes B_1$

$$A_1 \otimes I_n + I_m \otimes B_1$$

$$= \begin{pmatrix} \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ O & & \lambda_1 \end{pmatrix} & & * \\ & \ddots & \\ O & & \begin{pmatrix} \lambda_m & & * \\ & \ddots & \\ O & & \lambda_m \end{pmatrix} \end{pmatrix} + \begin{pmatrix} \begin{pmatrix} t_1 & & * \\ & \ddots & \\ O & & t_n \end{pmatrix} & & * \\ & \ddots & \\ O & & \begin{pmatrix} t_1 & & * \\ & \ddots & \\ O & & t_n \end{pmatrix} \end{pmatrix}$$

$$= \begin{pmatrix} \begin{pmatrix} \lambda_1 + t_1 & & * \\ & \ddots & \\ O & & \lambda_1 + t_n \end{pmatrix} & & * \\ & \ddots & \\ O & & \begin{pmatrix} \lambda_m + t_1 & & * \\ & \ddots & \\ O & & \lambda_m + t_n \end{pmatrix} \end{pmatrix}$$

同理：$A \otimes I_n - I_m \otimes B$ 的特征值为 $\{\lambda_i - t_j\}$

注： $\because B^T$ 与 B 有相同的特征值 $\{t_1, t_2, \ldots, t_n\}$

推论： （1）$A^T \otimes I_n + I_m \otimes B^T$ 的全体特征值为 mn 个 $\{\lambda_i + t_j\}$

（2）$A^T \otimes I_n - I_m \otimes B^T$ 的全体特征值为 mn 个 $\{\lambda_i - t_j\}$

直积的特征值： 设方阵 $A = A_{m\times m}$、$B = B_{n\times n}$

令 A 的 m 个特征值（谱）$\sigma(A) = \{\lambda_1, \lambda_2, \ldots, \lambda_m\}$

B 的 n 个特征值（谱）$\sigma(B) = \{t_1, t_2, \ldots, t_n\}$

则 （1） $A \otimes B$ 的 mn 个特征值 $\{\lambda_i t_j\}$ $i = 1, 2, \ldots, m$ $j = 1, 2, \ldots, n$

（2） $A \otimes I_n + I_m \otimes B$ 的 mn 个特征值 $\{\lambda_i + t_j\}$

（3） $A \otimes I_n - I_m \otimes B$ 的 mn 个特征值 $\{\lambda_i - t_j\}$

（2）' $A^T \otimes I_n + I_m \otimes B^T$ 的 mn 个特征值 $\{\lambda_i + t_j\}$

（3）' $A^T \otimes I_n - I_m \otimes B^T$ 的 mn 个特征值 $\{\lambda_i - t_j\}$

推论： $|A \otimes B| = |A|^n |B|^m$

Pf: $\because |A \otimes B| = (\lambda_1 t_1 \cdot \lambda_1 t_2 \cdots\cdots \lambda_1 t_n)(\lambda_2 t_1 \cdot \lambda_2 t_2 \cdots\cdots \lambda_2 t_n)\cdots(\lambda_m t_1 \cdot \lambda_m t_2 \cdots\cdots \lambda_m t_n)$

$$=\left(\lambda_1^n\cdot\lambda_2^n\cdot\dots\cdot\lambda_m^n\right)\left(t_1\cdot t_2\cdot\dots\cdot t_n\right)^m=\left(\lambda_1\cdot\lambda_2\cdot\dots\cdot\lambda_m\right)^n\left(t_1\cdot t_2\cdot\dots\cdot t_n\right)^m=|A|^n|B|^m$$

Pf1：用许尔公式$\Rightarrow P^{-1}AP=\begin{pmatrix}\lambda_1 & & * \\ & \lambda_2 & \\ & & \ddots & \\ O & & & \lambda_m\end{pmatrix}\underline{\underline{记}}A_1$ （上三角阵）

$$\Rightarrow\left(P\otimes I_n\right)^{-1}\left(A\otimes B\right)\left(P\otimes I_n\right)=\left(P^{-1}AP\right)\otimes B=A_1\otimes B$$

$$=\begin{pmatrix}\lambda_1 B & & *B \\ & \lambda_2 B & \\ & & \ddots & \\ O & & & \lambda_m B\end{pmatrix}\Rightarrow|A\otimes B|=|\lambda_1 B||\lambda_2 B|\cdots|\lambda_m B|=\left(\lambda_1^n|B|\right)\left(\lambda_2^n|B|\right)\cdots\left(\lambda_m^n|B|\right)$$

拉直定义：$A=(a_{ij})_{m\times n}$

规定$\vec{A}=\left(a_{11},a_{12},\cdots,a_{1n},a_{21},a_{22},\cdots,a_{2n},\cdots,a_{m1},a_{m2},\cdots,a_{mn}\right)^T$ （列向量）

例：$\overrightarrow{\begin{pmatrix}1 & 2\\ 3 & 4\end{pmatrix}}=(1,2,3,4)^T=\begin{pmatrix}1\\2\\3\\4\end{pmatrix}$

特别：（1）$\alpha=(a_1,a_2,\dots,a_n)^T$为列向量，则$\vec{\alpha}=\alpha=\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix}$

（2）若 A$=(\alpha_1,\alpha_2,\dots,\alpha_n)$为行向量，$\vec{A}=A^T$

设$A=\begin{pmatrix}A_1\\A_2\\\vdots\\A_n\end{pmatrix}$（按行）$\Rightarrow\vec{A}=\left(A_1,A_2,\cdots,A_n\right)^T=\begin{pmatrix}A_1^T\\A_2^T\\\vdots\\A_n^T\end{pmatrix}$

注：拉直是$C^{m\times n}$与C^{mn}（列空间）之间的同构（一一对应）

性质：（1）$\overrightarrow{A+B}=\vec{A}+\vec{B}$ **注**：$\overrightarrow{A_{m\times n}}=\overrightarrow{B_{m\times n}}\Leftrightarrow A_{m\times n}=B_{m\times n}$

（2）$\overrightarrow{kA}=k\vec{A}$

（3）$\overrightarrow{\dfrac{dA(t)}{dt}}=\dfrac{d\overrightarrow{A(t)}}{dt}$

拉直公式：$\overrightarrow{(ABC)}=\left(A\otimes C^T\right)\vec{B}$

Pf：$A=(a_{ij})_{m\times n}$、$B=(b_{ij})_{n\times p}$、$C=(c_{ij})_{p\times q}$

按行写 $B=\begin{pmatrix}B_1\\B_2\\\vdots\\B_n\end{pmatrix}\Rightarrow\vec{B}=\begin{pmatrix}B_1^T\\B_2^T\\\vdots\\B_n^T\end{pmatrix}\in C^{np}$

$$\overrightarrow{(ABC)}=\overrightarrow{\begin{pmatrix}a_{11}B_1+a_{12}B_2+\cdots+a_{1n}B_n\\a_{21}B_1+a_{22}B_2+\cdots+a_{2n}B_n\\\vdots\\a_{m1}B_1+a_{m2}B_2+\cdots+a_{mn}B_n\end{pmatrix}C}=\overrightarrow{\begin{pmatrix}(a_{11}B_1+a_{12}B_2+\cdots+a_{1n}B_n)C\\(a_{21}B_1+a_{22}B_2+\cdots+a_{2n}B_n)C\\\vdots\\(a_{m1}B_1+a_{m2}B_2+\cdots+a_{mn}B_n)C\end{pmatrix}}$$

$$=\begin{pmatrix}C^T\left(a_{11}B_1^T+a_{12}B_2^T+\cdots+a_{1n}B_n^T\right)\\C^T\left(a_{21}B_1^T+a_{22}B_2^T+\cdots+a_{2n}B_n^T\right)\\\vdots\\C^T\left(a_{m1}B_1^T+a_{m2}B_2^T+\cdots+a_{mn}B_n^T\right)\end{pmatrix}=\begin{pmatrix}a_{11}C^T&a_{12}C^T&a_{1n}C^T\\a_{21}C^T&a_{22}C^T&a_{2n}C^T\\\\a_{m1}C^T&a_{m2}C^T&a_{mn}C^T\end{pmatrix}\begin{pmatrix}B_1^T\\B_2^T\\\vdots\\B_n^T\end{pmatrix}$$

$$=\left(A\otimes C^T\right)\vec{B}$$

特例：$\overrightarrow{AY}=\overrightarrow{AYI}=\left(A\otimes I\right)\vec{Y}$

$\overrightarrow{YB}=\overrightarrow{IYB}=\left(I\otimes B^T\right)\vec{Y}$

$\overrightarrow{AY\pm YB}=\left(A\otimes I\pm I\otimes B^T\right)\vec{Y}$

$\overrightarrow{AYB+A^2YB^2}=\left(A\otimes B^T+A^2\otimes\left(B^2\right)^T\right)\vec{Y}=\left(A\otimes B^T+\left(A\otimes B^T\right)^2\right)\vec{Y}$

应用：求矩阵方程：$AYB=F$ 的解

方程：先拉直 $\left(A\otimes B^T\right)\vec{Y}=\vec{F}$ 再求解

用增广阵 $\left(A\otimes B^T\middle|\vec{F}\right)\xrightarrow{\text{行变}}\cdots$

补充公式：$\left(A\otimes B\right)^+=A^+\otimes B^+$ （只须验证 4 个条件）

结论：$AYB=F$ 的最佳小二解为 $Y_0=A^+FB^+$

Pf：$\because\left(A\otimes B^T\right)\vec{Y}=\vec{F}$ 最佳小二解为 $\overrightarrow{Y_0}=\left(A\otimes B^T\right)^+\vec{F}$

$\overrightarrow{Y_0}=\left(A\otimes B^T\right)^+\vec{F}=\left(A^+\otimes\left(B^T\right)^+\right)\vec{F}=\overrightarrow{\left(A^+FB^+\right)}$

引理：$A=A_{m\times m}$、$B=B_{n\times n}$

（1）$\left(A\otimes I_n - I_m\otimes B\right)$可逆$\Leftrightarrow A$、$B$没有公共特征值（$\lambda_i \neq t_j$）

或$\left(A\otimes I_n - I_m\otimes B^T\right)$可逆$\Leftrightarrow A$、$B$没有公共特征值

（2）$\left(A\otimes I_n + I_m\otimes B\right)$可逆$\Leftrightarrow A$、$(-B)$没有公共特征值（$\lambda_i \neq -t_j$）

应用：$A = A_{m\times m}$、$B = B_{n\times n}$

1. $AY - YB = F$ 有唯一解$\Leftrightarrow \left(A\otimes I_n - I_m\otimes B\right)$可逆$\Leftrightarrow A$、$B$无公共特征值

∵拉直：$\left(A\otimes I_n - I_m\otimes B^T\right)\vec{Y} = \vec{F}$

2. $AY - YB = F$ 有唯一解$\Leftrightarrow A$与$(-B)$无公共特征值

Eg.若A、B无公共特征值，则$\begin{pmatrix} A & F \\ 0 & B \end{pmatrix}$相似于$\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$

Pf：令$P = \begin{pmatrix} I_m & Y \\ 0 & I_n \end{pmatrix}$ （可逆） $P^{-1} = \begin{pmatrix} I_m & (-Y) \\ 0 & I_n \end{pmatrix}$

$\Rightarrow P\begin{pmatrix} A & F \\ 0 & B \end{pmatrix}P^{-1} = \begin{pmatrix} A & F + YB - AY \\ 0 & B \end{pmatrix}$

再求解方程：$F + YB - AY = 0 \Leftrightarrow AY - YB = F$

∵A、B无公共特征值$\Rightarrow$有唯一解F使得$F + YB - AY = 0$

对于这个Y：必有$P\begin{pmatrix} A & F \\ 0 & B \end{pmatrix}P^{-1} = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$

推广：若A与B、C无公共特征值

则$\begin{pmatrix} A & F_1 & F_2 \\ & B & 0 \\ & & C \end{pmatrix} \backsim \begin{pmatrix} A & & \\ & B & \\ & & C \end{pmatrix}$

Ex.讲义 P3112、9

验证$\left(A\otimes B\right)^+ = A^+ \otimes B^+$