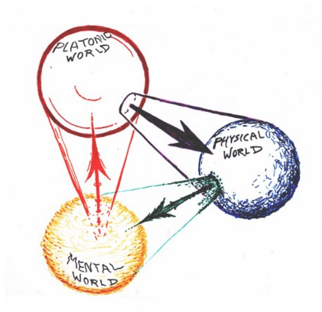
高等代数讲义

作者：Hongxin Yang 南风

时间：August 7, 2025



目录

第 1 章 相关定理环境等的标识

Proof 使用方法:begin+ proof+ end

注 使用方法:begin+ remark+ end

Example 1.1 使用方法；begin+ example + end

练习 **1.1** 使用方法:begin+ exercise +end

性质 使用方法:begin+ property + end

Problem 1.1 使用方法:begin + problem +end

结论 使用方法:begin+ conclusion +end

笔记 使用方法；begin + note + end

Definition 1.1

使用方法:begin+ definition+ end

Proposition 1.1

使用方法:begin+ proposition+ end

Lemma 1.1

使用方法:begin+ lemma+ end

Corollary 1.1

使用方法:begin + corollary + end

公理 1.1

使用方法:begin+ axiom + end

公设 1.1

使用方法:begin + postulate + end

第 2 章 行列式

2.1 行列式重要模板

Theorem 2.1.1

重要模板一：A, B为n阶矩阵。

$$|A+B| = |A| + |B| + \sum_{1\leqslant k\leqslant n-1}\left(\sum_{\substack{1\leqslant i_1<i_2<\cdots<i_k\leqslant n\\ 1\leqslant j_1<j_2<\cdots<j_k\leqslant n}} A\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{pmatrix}\widehat{B}\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{pmatrix}\right)$$

Proof 设$|A| = |\alpha_1, \alpha_2 \cdots \alpha_n|$ $|B| = |\beta_1, \beta_2 \cdots \beta_n|$ 对$|A+B|$应用拆分法 st.每一列仅含有α_i 或者β_i 共2^n 个行列式
对每个行列式用$Laplace$定理按照含有A的那些列向量展开即可 (常用于拆分后对角阵 (因为这样只有主子式非0)或秩比较小的矩阵)

Theorem 2.1.2

重要模板二：爪型行列式

$$求\begin{vmatrix} \alpha+\beta & \alpha\beta & 0 & \cdots & 0 & 0 \\ 1 & \alpha+\beta & \alpha\beta & \cdots & 0 & 0 \\ 0 & 1 & \alpha+\beta & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & \alpha+\beta \end{vmatrix} = \frac{\alpha^{n+1}-\beta^{n+1}}{\alpha-\beta}(\alpha\neq\beta) = \alpha^n + \alpha^{n-1}\beta + \cdots + \alpha\beta^{n-1} + \beta^n(\alpha=\beta)$$

Proof 法一：第一行展开，$\Rightarrow D_n = (\alpha+\beta)D_{n-1} - \alpha\beta D_{n-2}$ 。利用第二数学归纳法证明即可

法二：将第一列拆成$\begin{pmatrix}\alpha\\1\\\vdots\\0\\0\end{pmatrix} + \begin{pmatrix}\beta\\0\\\vdots\\0\\0\end{pmatrix}$，其中$\begin{pmatrix}\alpha\\1\\\vdots\\0\\0\end{pmatrix}$从第一列起的 $-\beta$ 倍加到后一列。 $\Rightarrow D_n = \alpha^n + \beta D_{n-1}$ 同理$D_n = \beta^n + \alpha D_{n-1}$

$\Rightarrow (\alpha-\beta)D_n = \alpha^{n+1} - \beta^{n+1} \Rightarrow \dfrac{\alpha^{n+1}-\beta^{n+1}}{\alpha-\beta}(\alpha\neq\beta)$

$when\ \ \alpha=\beta$时

$when\alpha\neq 0$时候。$D_n = \alpha^n + \beta D_{n-1} \Leftrightarrow D_n = \alpha^n + \alpha D_{n-1} \Leftrightarrow \dfrac{D_n}{\alpha^n} - \dfrac{D_{n-1}}{\alpha^{n-1}} = 1$

$\Rightarrow$ (等差数列)$D_n = (n+1)\alpha^n = \alpha^n + \alpha^{n-1}\beta + \cdots + \alpha\beta^{n-1} + \beta^n(\alpha=\beta)$

$when\alpha=0$时候。无论是从$D_n = \alpha^n + \beta D_{n-1}$还是从最原始的表达式都知道$D_n = 0 = \alpha^n + \alpha^{n-1}\beta + \cdots + \alpha\beta^{n-1} + \beta^n(\alpha=\beta)$

法三：从$D_n = (\alpha+\beta)D_{n-1} - \alpha\beta D_{n-2}$

引入特征根方程：$x^2 - (\alpha+\beta)x + \alpha\beta = 0$的两根为$a, b$

$\Rightarrow a+b = \alpha+\beta\ \ ande\ \ ab = \alpha\beta$(不难看出根$a, b$为$\alpha\beta$)and$D_1 = \alpha+\beta = a+b \quad D_2 = (\alpha+\beta)^2 - \alpha\beta = (a+b)^2 - ab$

$$\begin{cases} D_n - aD_{n-1} = b(D_{n-1} - aD_{n-2}) \\ D_n - bD_{n-1} = a(D_{n-1} - bD_{n-1}) \end{cases} \Rightarrow 由递推可以知道\begin{cases} D_n - aD_{n-1} = b^{n-2}(D_2 - aD_1) \\ D_n - bD_{n-1} = a^{n-2}(D_2 - bD_1) \end{cases} \Rightarrow D_2 - aD_1 = b^2 \text{ and} D_2 - bD_1 = a^2$$

$$\Rightarrow \begin{cases} D_n - aD_{n-1} = b^n \\ D_n - bD_{n-1} = a^n \end{cases} \Rightarrow (b-a)D_n = b^{n+1} - a^{n+1}$$

所以当$a\neq b \Leftrightarrow \Delta\neq 0 \Leftrightarrow \alpha\neq\beta \Rightarrow D_n = \dfrac{\alpha^{n+1}-\beta^{n+1}}{\alpha-\beta}(\alpha\neq\beta)$

当$a=b \Leftrightarrow \Delta=0 \Leftrightarrow \alpha=\beta$(此时$a=b=\alpha=\beta$) $\quad D_n-aD_{n-1}=a^n$

$\Rightarrow$ when$a=0 \Leftrightarrow a=b=\alpha=\beta=0 \Rightarrow$ 易知$D_n=0$

$\Rightarrow$ when$a\neq 0 \Leftrightarrow \frac{D_n}{a^n}-\frac{D_{n-1}}{a^{n-1}}=1\left(\text{且}\frac{D_1}{a}=\frac{a+b}{a}=2\right) \Rightarrow \frac{D_n}{a^n}=(n-1)\times 1+2 \Rightarrow D_n=(n+1)a^n=(n+1)\alpha^n$

Theorem 2.1.3

求三线性行列式$\begin{vmatrix} a & b & & & & \\ c & a & b & & & \\ & c & a & \ddots & & \\ & & \ddots & \ddots & b & \\ & & & c & a & b \\ & & & & c & a \end{vmatrix}$

Proof $prove: D_n=aD_{n-1}-bcD_{n-2}$

引入特征根方程$x^2-ax+bc=0$的两根为α and$\beta \Rightarrow \alpha+\beta=a$ and$\alpha\beta=bc$

$D_1=a=\alpha+\beta$ and$D_2=a^2-bc=(\alpha+\beta)^2-\alpha\beta$

$$\Rightarrow \begin{cases} D_n-\alpha D_{n-1}=\beta(D_{n-1}-\alpha D_{n-2}) \\ D_n-\beta D_{n-1}=\alpha(D_{n-1}-\beta D_{n-2}) \end{cases} \Rightarrow \begin{cases} D_n-\alpha D_{n-1}=\beta^{n-2}(D_2-\alpha D_1) \\ D_n-\beta D_{n-1}=\alpha^{n-2}(D_2-\beta D_1) \end{cases} \Rightarrow \begin{cases} D_n-\alpha D_{n-1}=\beta^n \\ D_n-\beta D_{n-1}=\alpha^n \end{cases} \Rightarrow (\alpha-\beta)D_n=\alpha^{n+1}-\beta^{n+1}$$

$(1): when\ \alpha\neq\beta \Leftrightarrow \Delta\neq 0 \Leftrightarrow a^2\neq 4bc \quad D_n=\frac{\alpha^{n+1}-\beta^{n+1}}{\alpha-\beta}$

$(2): when\ \alpha=\beta \Leftrightarrow \Delta=0 \Leftrightarrow a^2=4bc$且$\alpha=\beta=\frac{a}{2}$此时$D_n-\alpha D_{n-1}=\alpha^n$

$1^\circ: when\ \alpha=0 \Rightarrow a=0 and\ bc=0$(所以$bc$之间必定有一个为0所以原行列式第一行或者第一列为0) $\Rightarrow D_n=0$

$2^\circ: when\alpha\neq 0 \Rightarrow \frac{D_n}{\alpha^n}-\frac{D_{n-1}}{\alpha^{n-1}}=1 \Rightarrow$ 此时$\frac{D_1}{\alpha}=\frac{\alpha+\beta}{\alpha}=2 \Rightarrow \frac{D_n}{\alpha^n}=(n-1)\times 1+2 \Rightarrow D_n=(n+1)\alpha^n \Rightarrow D_n=(n+1)\left(\frac{a}{2}\right)^n$

Theorem 2.1.4

三线型行列式爪型

$$D_n=\begin{vmatrix} a_1 & b_2 & b_3 & \cdots & b_{n-1} & b_n \\ c_2 & a_2 & & & & \\ c_3 & & a_3 & & & \\ \vdots & & & \ddots & & \\ c_{n-1} & & & & a_{n-1} & \\ c_n & & & & & a_n \end{vmatrix}$$

Proof 最后一列展开$D_n=a_nD_{n-1}-b_nc_na_2a_3\cdots a_{n-1}$ 并且$D_1=a_1 D_2=a_1a_2-b_2c_2$

$D_n-a_nD_{n-1}=-b_nc_na_2a_3\cdots a_{n-1}$(利用差分方程法) $\Rightarrow \frac{D_n}{a_2a_3\cdots a_n}-\frac{D_{n-1}}{a_2a_3\cdots a_{n-1}}=-\frac{b_nc_n}{a_n}$

利用递推法知道$\frac{D_n}{a_2a_3\cdots a_n}=a_1-\sum_{k=2}^{n}\frac{b_kc_k}{a_k}$

$D_n=\prod_{i=1}^{n}a_i-a_2a_3\cdots a_n\sum_{k=2}^{n}\frac{b_kc_k}{a_k}\left(D_n=\prod_{i=1}^{n}a_i-\sum_{k=2}^{n}a_2\cdots\widehat{a_k}\cdots a_nb_kc_k\right)$

$tip:$ 当$a_i=0$进行展开即可观察可得也可以由多项式性质

Theorem 2.1.5

Hessenberg型行列式

$$D_n=\begin{vmatrix} la_1 & a_2 & a_3 & \cdots & a_{n-1} & a_n \\ c_1 & b_1 & & & & \\ & c_2 & b_2 & & & \\ & & \ddots & \ddots & & \\ & & & c_{n-2} & b_{n-2} & \\ & & & & c_{n-1} & b_{n-1} \end{vmatrix}$$ (三头蛇头比心住部位指标大)(否则可以行列互换)

Proof 按照最后一列展开$D_n=(-1)^{n+1}a_nc_1c_2\cdots c_{n-1}+b_{n-1}D_{n-1}$ and$D_1=a_1$ and$D_2=a_1b_1-c_1a_2$

$D_n-b_{n-1}D_{n-1}=(-1)^{n+1}a_nc_1c_2\cdots c_{n-1}$(利用差分方程法) $\Rightarrow \dfrac{D_n}{b_1b_2\cdots b_{n-1}}-\dfrac{D_{n-1}}{b_1b_2\cdots b_{n-2}}=(-1)^{n+1}\dfrac{c_1c_2\cdots c_{n-1}a_n}{b_1b_2\cdots b_{n-1}}$

递推得到：$D_n=a_1b_1b_2\cdots b_{n-1}+\sum\limits_{k=2}^{n}(-1)^{k+1}c_1\cdots c_{k-1}a_kb_k\cdots b_{n-1}$

Theorem 2.1.6 (行列式求导)

证明：$$\frac{d}{dt}\begin{vmatrix} a_{11}(t) & a_{12}(t) & \cdots & a_{1n}(t) \\ a_{21}(t) & a_{22}(t) & \cdots & a_{2n}(t) \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}(t) & a_{n2}(t) & \cdots & a_{nn}(t) \end{vmatrix}=\sum_{j=1}^{n}\begin{vmatrix} a_{11}(t) & \cdots & \frac{d}{dt}a_{1j}(t) & \cdots & a_{1n}(t) \\ a_{21}(t) & \cdots & \frac{d}{dt}a_{2j}(t) & \cdots & a_{2n}(t) \\ \vdots & & \vdots & & \vdots \\ a_{n1}(t) & \cdots & \frac{d}{dt}a_{nj}(t) & \cdots & a_{nn}(t) \end{vmatrix}$$

Proof *prove :*

$$LHS=\frac{d}{dt}\sum_{i_1i_2\cdots i_n}(-1)^{\tau(i_1i_2\cdots i_n)}a_{i_11}(t)\cdots a_{i_jj}(t)\cdots a_{i_nn}(t)=\sum_{i_1i_2\cdots i_n}(-1)^{\tau(i_1i_2\cdots i_n)}\cdot\left(\sum_{j=1}^{n}a_{i_11}(t)\cdots\frac{d}{dt}a_{i_jj}(t)\cdots a_{i_nn}(t)\right)$$
$$=\sum_{j=1}^{n}\sum_{i_1i_2\cdots i_n}(-1)^{\tau(i_1i_2\cdots i_n)}a_{i_11}(t)\cdots\frac{d}{dt}a_{i_jj}(t)\cdots a_{i_nn}(t)=RHS$$

Theorem 2.1.7 (对分行列式)

求 $$\begin{vmatrix} x_1 & y & y & \cdots & y & y \\ z & x_2 & y & \cdots & y & y \\ z & z & x_3 & \cdots & y & y \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ z & z & z & \cdots & x_{n-1} & y \\ z & z & z & \cdots & z & x_n \end{vmatrix}$$

Proof 类似的，$C_n=C_n-C_{n-1}\Rightarrow$ 最后一列展开，$D_n=(x_n-z)D_{n-1}+z\prod\limits_{i=1}^{n-1}(x_i-y)$ 同理转置得到：$D_n=(x_n-y)D_{n-1}+y\prod\limits_{i=1}^{n-1}(x_i-z)$

$\Rightarrow (z-y)D_n=z\prod\limits_{1}^{n}(x_i-y)-y\prod\limits_{1}^{n}(x_i-z)$

$1°$：when$z\neq y$ $D_n=\dfrac{z\prod\limits_{i=1}^{n}(x_i-y)-y\prod\limits_{i=1}^{n}(x_i-z)}{z-y}$

2° : when$z=y$ 令$E_n=\dfrac{D_n}{\prod\limits_{i=1}^{n}(x_i-y)}$. $\Rightarrow E_n=E_{n-1}+\dfrac{y}{x_n-y}\Rightarrow E_n=1+\sum\limits_{i=1}^{n}\dfrac{y}{x_i-y}\Rightarrow D_n=\cdots$

由递推式子知道为$D_n=(x_n-y)\,D_{n-1}+y\prod\limits_{i=1}^{n-1}(x_i-y)\Rightarrow D_n=\prod\limits_{i=1}^{n}(x_i-y)+y\sum\limits_{i=1}^{n}\left(\prod\limits_{j=1,j\neq i}^{n}(x_j-z)\right)$

Theorem 2.1.8

拆分法$|A|=\begin{vmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&\ddots&\vdots\\a_{n1}&a_{n2}&\cdots&a_{nn}\end{vmatrix}$

$prove:|A(t_1,t_2\cdots t_n)|=\begin{vmatrix}a_{11}+t_1&a_{12}+t_2&\cdots&a_{1n}+t_n\\a_{21}+t_1&a_{22}+t_2&\cdots&a_{2n}+t_n\\\vdots&\vdots&\ddots&\vdots\\a_{n1}+t_1&a_{n2}+t_2&\cdots&a_{nn}+t_n\end{vmatrix}=|A|+\sum\limits_{j=1}^{n}t_j\sum\limits_{i=1}^{n}A_{ij}$

Proof $prove:|A(t_1,t_2\cdots t_n)|=\begin{vmatrix}a_{11}+t_1&a_{12}+t_2&\cdots&a_{1n}+t_n\\a_{21}+t_1&a_{22}+t_2&\cdots&a_{2n}+t_n\\\vdots&\vdots&\ddots&\vdots\\a_{n1}+t_1&a_{n2}+t_2&\cdots&a_{nn}+t_n\end{vmatrix}=|A|+\sum\limits_{j=1}^{n}t_j\sum\limits_{i=1}^{n}A_{ij}$

$$\text{LHS}=\begin{vmatrix}a_{11}+t_1&a_{12}+t_2&\cdots&a_{1n}\\a_{21}+t_1&a_{22}+t_2&\cdots&a_{2n}\\\vdots&\vdots&\ddots&\vdots\\a_{n1}+t_1&a_{n2}+t_2&\cdots&a_{nn}\end{vmatrix}+\begin{vmatrix}a_{11}+t_1&a_{12}+t_2&\cdots&t_n\\a_{21}+t_1&a_{22}+t_2&\cdots&t_n\\\vdots&\vdots&\ddots&\vdots\\a_{n1}+t_1&a_{n2}+t_2&\cdots&t_n\end{vmatrix}=\begin{vmatrix}a_{11}+t_1&a_{12}+t_2&\cdots&a_{1n}\\a_{21}+t_1&a_{22}+t_2&\cdots&a_{2n}\\\vdots&\vdots&\ddots&\vdots\\a_{n1}+t_1&a_{n2}+t_2&\cdots&a_{nn}\end{vmatrix}+t_n\begin{vmatrix}a_{11}+t_1&a_{12}+t_2&\cdots&1\\a_{21}+t_1&a_{22}+t_2&\cdots&1\\\vdots&\vdots&\ddots&\vdots\\a_{n1}+t_1&a_{n2}+t_2&\cdots&1\end{vmatrix}$$

$$=\begin{vmatrix}a_{11}+t_1&a_{12}+t_2&\cdots&a_{1n}\\a_{21}+t_1&a_{22}+t_2&\cdots&a_{2n}\\\vdots&\vdots&\ddots&\vdots\\a_{n1}+t_1&a_{n2}+t_2&\cdots&a_{nn}\end{vmatrix}+t_n\sum_{i=1}^{n}A_{in}$$

同样的操作我们可以得到$|A(t_1,t_2\cdots t_n)|=\begin{vmatrix}a_{11}+t_1&a_{12}+t_2&\cdots&a_{1n}+t_n\\a_{21}+t_1&a_{22}+t_2&\cdots&a_{2n}+t_n\\\vdots&\vdots&\ddots&\vdots\\a_{n1}+t_1&a_{n2}+t_2&\cdots&a_{nn}+t_n\end{vmatrix}=|A|+\sum\limits_{j=1}^{n}t_j\sum\limits_{i=1}^{n}A_{ij}$

Problem 2.1 计算$|A|=\begin{vmatrix}a&b&\cdots&b\\c&a&\cdots&b\\c&c&\ddots&b\\c&c&\cdots&a\end{vmatrix}$

笔记 $solve\ \ let|A(t)|=\begin{vmatrix}a+t&b+t&\cdots&b+t\\c+t&a+t&\cdots&b+t\\c+t&c+t&\ddots&b+t\\c+t&c+t&\cdots&a+t\end{vmatrix}=|A|+tu,u=\sum\limits_{i,j=1}^{n}A_{ij}$

and observe that $\boldsymbol{u}$ *is uncorrelated with* $\boldsymbol{t}$ *when* $\boldsymbol{t}=-b$ *we have* $|A(-b)|=\begin{vmatrix} a-b & 0 & \cdots & 0 \\ c-b & a-b & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ c-b & c-b & \cdots & a-b \end{vmatrix}=|A|-bu=(a-b)^n$

in th esame way $|A(-c)|=|A|-cu=(a-c)^n$

so if $b\neq c\Rightarrow |A|=\dfrac{b(a-c)^n-c(a-b)^n}{b-c}$

so if $b\neq c\Rightarrow |A|=(a+(n-1)b)(a-b)^{n-1}$*(because it is a det that the sum of every row is equal)*

Theorem 2.1.9 (添加一行一列型)

设$|A|=\left|a_{ij}\right|$是一个n阶行列式, A_{ij}是第(i,j)元素的代数余子式, M_{ij}为余子式

求证:
$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} & x_1 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & x_n \\ y_1 & y_2 & \cdots & y_n & 1 \end{vmatrix}_{n+1}=|A|-\sum_{i=1}^{n}\sum_{j=1}^{n}A_{ij}x_iy_j$$

Proof 再按照最后一行展开通项为$(-1)^{n+j}y_jM_{ij}\Rightarrow\sum_{j=1}^{n}(-1)^{n+j}y_jM_{ij}=\sum_{j=1}^{n}(-1)^{i+j}(-1)^{n-i}y_jM_{ij}=(-1)^{n-i}\sum_{j=1}^{n}y_jA_{ij}$

$\Rightarrow$ 第i 项为$(-1)^{i+n+1}x_i\times(-1)^{n-i}\sum_{j=1}^{n}y_jA_{ij}=-x_i\sum_{j=1}^{n}y_jA_{ij}=-\sum_{j=1}^{n}x_iy_jA_{ij}\quad\Rightarrow LHS=-\sum_{i=1}^{n}\sum_{j=1}^{n}A_{ij}x_iy_j+|A|$

特别的当右下角1变为z时，我们最终答案$|A|\to z\,|A|$也可以用摄动法证明在摄动法一节

Theorem 2.1.10 (类范德蒙德行列式)

设$f_k(x)=x^k+a_{k1}x^{k-1}+a_{k2}x^{k-2}+\cdots+a_{kk}$

求行列式
$$\begin{vmatrix} 1 & f_1(x_1) & f_2(x_1) & \cdots & f_{n-1}(x_1) \\ 1 & f_1(x_2) & f_2(x_2) & \cdots & f_{n-1}(x_2) \\ \vdots & \vdots & \vdots & & \\ 1 & f_1(x_n) & f_2(x_n) & \cdots & f_{n-1}(x_n) \end{vmatrix}$$

Proof *solve* $\Rightarrow\begin{vmatrix} 1 & x_1+a_{11} & x_1^2+a_{21}x_1+a_{22} & \cdots & f_{n-1}(x_1) \\ 1 & x_2+a_{11} & x_2^2+a_{21}x_2+a_{22} & \cdots & f_{n-1}(x_2) \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n+a_{11} & x_n^2+a_{21}x_n+a_{22} & \cdots & f_{n-1}(x_n) \end{vmatrix}$得到一个$VanderMonde$行列式 $\Rightarrow\prod_{1\leqslant i<j\leqslant n}^{n}(x_j-x_n)$

Problem 2.2 $|A|=\begin{vmatrix} 1 & \cos\theta_1 & \cos 2\theta_1 & \cdots & \cos(n-1)\theta_1 \\ 1 & \cos\theta_2 & \cos 2\theta_2 & \cdots & \cos(n-1)\theta_2 \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \cos\theta_n & \cos 2\theta_n & \cdots & \cos(n-1)\theta_n \end{vmatrix}$

笔记 *From DeMoivre Formula we know* $\cos n\theta+i\sin n\theta=(\cos\theta+i\sin\theta)^n=\sum_{k=0}^{n}C_n^k i^k\sin^k\theta\cos^{n-k}\theta$

contrast the former and the latter wecan get only when k is even number without loss of generality assume

$k=2t$ *so* $\sum_{t=1}^{2t\leqslant n}C_n^{2t}(-1)^t\left(1-\cos^2\theta\right)^t\cos^{n-2t}\theta$ *so unwind this formula we can get a about* $\cos\theta$ *ploynomial and the highest*

$dergee\ term\ is \cos^n\theta\ and\ the\ coefficient\ is\ C_n^{2t}\ so\ |A| every\ column\ can\ extract\ 2\prime spow\ (1+C_n^2+C_n^4+\cdots=2^{n-1})$

$intotal\ 2^{(1-1)+(2-1)+\cdots+(n-2)}=2^{\frac{(n-1)(n-2)}{2}}\ and\ use\ the\ the\ former\ question\ so\ \Rightarrow |A|=2^{\frac{(n-1)(n-2)}{2}}\prod_{i\leqslant i<j\leqslant n}^{n}(\cos\theta_j-\cos\theta_i)$

Problem 2.3 $|A|=\begin{vmatrix}\sin\theta_1 & \sin 2\theta_1 & \cdots & \sin n\theta_1\\ \sin\theta_2 & \sin 2\theta_2 & \cdots & \sin n\theta_2\\ \vdots & \vdots & & \vdots\\ \sin\theta_n & \sin 2\theta_n & \cdots & \sin n\theta_n\end{vmatrix}$

笔记 $method1also\ from\ De\ Moivre\ Formula\ we\ can\ know\ (\cos n\theta+i\sin n\theta)=(\cos\theta+i\sin\theta)^n=\sum_{k=0}^{n}C_n^k i^k\sin^k\theta\cos^{n-k}\theta$

$contrast\ the\ former\ and\ the\ latter\ we\ can\ get\ only\ when\ k\ is\ odd\ number\ without\ loss\ of\ generality$

$assume\ k=2t+1\ \sum_{t=0}^{2t+1\leqslant n}C_n^{2t+1}(-1)^t i\sin^{2t+1}\theta\cos^{n-2t-1}\theta=\sum_{t=0}^{2t+1\leqslant n}C_n^{2t+1}(-1)^t i\sin\theta\left(1-\cos^2\theta\right)^t\cos^{n-2t-1}\theta$

$so\ every\ row\ can\ contract\ \sin\theta_i\ and\ unwind\ this\ formula\ the\ highest\ degree\ term\ is \cos^{n-1}\theta$

$and\ the\ coefficient\ is\ 2\prime spow\ (C_n^1+C_n^3+\cdots=2^{n-1})\ now\ it\ becomes\ the\ former\ question$

$$so|A|=2^{\frac{(n-1)(n-2)}{2}}\times\prod_{i=1}^{n}\sin\theta_i\times\prod_{i\leqslant i<j\leqslant n}^{n}(\cos\theta_j-\cos\theta_i)$$

Theorem 2.1.11 (除了对角线元素不同其他都一样)

$$D_n=\begin{vmatrix}x_1-a_1 & x_2 & x_3 & \cdots & x_n\\ x_1 & x_2-a_2 & x_3 & \cdots & x_n\\ x_1 & x_2 & x_3-a_3 & \cdots & x_n\\ \vdots & \vdots & \vdots & & \vdots\\ x_1 & x_2 & x_3 & \cdots & x_n-a_n\end{vmatrix}$$

Proof

最快的方法是利用打洞原理将原矩阵拆开，

$method1$

$$|A|=\begin{vmatrix}x_1-a_1 & x_2 & x_3 & \cdots & x_n\\ a_1 & -a_2 & 0 & \cdots & 0\\ a_1 & 0 & -a_3 & \cdots & 0\\ \vdots & \vdots & \vdots & & \vdots\\ a_1 & 0 & 0 & \cdots & -a_n\end{vmatrix}=a_1\cdots a_n\begin{vmatrix}\frac{x_1}{a_1}-1 & \frac{x_2}{a_2} & \frac{x_3}{a_3} & \cdots & \frac{x_n}{a_n}\\ 1 & -1 & 0 & \cdots & 0\\ 1 & 0 & -1 & \cdots & 0\\ \vdots & \vdots & \vdots & & \vdots\\ 1 & 0 & 0 & \cdots & -1\end{vmatrix}\xrightarrow{C_1=C_1+C_2+\cdots C_n}\Rightarrow|A|=(-1)^{n-1}a_1\cdots a_n\left(\sum_{i=1}^{n}\frac{x_i}{a_i}-1\right)$$

$method2$

$use\ resolution;every\ colunm\ can\ be\ divided\ into\ two\ columns;one\ is\ all\ of\ x_j\ and\ another\ is\ -a_j\ and\ 0$

$so\ |A| at\ most\ has\ one\ column\ (x_j)\ otherwise\ det\ is 0.$

so$\det|A|=(-1)^n a_1\cdots a_n+\sum_{j=1}^{n}x_j(-1)^{n-1}a_1a_2\cdots a_{j-1}a_{j+1}\cdots a_n=(-1)^{n-1}a_1\cdots a_n\left(\sum_{i=1}^{n}\frac{x_i}{a_i}-1\right)$

$method3$

$$|A|=\begin{vmatrix} x_1-a_1 & x_2 & \cdots & x_n \\ x_1 & x_2-a_2 & \cdots & x_n \\ \vdots & \vdots & & \vdots \\ x_1 & x_2 & \cdots & x_n-a_n \end{vmatrix}=\begin{vmatrix} 1 & x_1 & x_2 & \cdots & x_n \\ 0 & x_1-a_1 & x_2 & \cdots & x_n \\ 0 & x_1 & x_2-a_2 & \cdots & x_n \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & x_1 & x_2 & \cdots & x_n-a_n \end{vmatrix}=\begin{vmatrix} 1 & x_1 & x_2 & \cdots & x_n \\ -1 & -a_1 & 0 & \cdots & 0 \\ -1 & 0 & -a_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ -1 & 0 & 0 & \cdots & -a_n \end{vmatrix}\text{利用爪型行列式即可}$$

$method4$

利用$Laplace|A+B|$方法来计算

$$LHS=\begin{pmatrix} x_1-a_1 & x_2 & \cdots & x_n \\ x_1 & x_2-a_2 & \cdots & x_n \\ \vdots & \vdots & & \vdots \\ x_1 & x_2 & \cdots & x_n-a_n \end{pmatrix}=\underbrace{\begin{pmatrix} x_1 & x_2 & \cdots & x_n \\ x_1 & x_2 & \cdots & x_n \\ \vdots & \vdots & & \vdots \\ x_1 & x_2 & \cdots & x_n \end{pmatrix}}_{A}+\underbrace{\begin{pmatrix} -a_1 & & & \\ & -a_2 & & \\ & & \ddots & \\ & & & -a_n \end{pmatrix}}_{B}$$

$$|LHS|=\begin{vmatrix} x_1 & x_2 & \cdots & x_n \\ x_1 & x_2 & \cdots & x_n \\ \vdots & \vdots & & \vdots \\ x_1 & x_2 & \cdots & x_n \end{vmatrix}+\begin{vmatrix} -a_1 & & & \\ & -a_2 & & \\ & & \ddots & \\ & & & -a_n \end{vmatrix}+\sum_{i=1}^{n}A\begin{pmatrix} i \\ i \end{pmatrix}\widehat{B}\begin{pmatrix} i \\ i \end{pmatrix}$$

(首先B 只能取主子式才不为0, 那么A 也只能取主子式但是为了不为0只能取一个元素)

$$=(-1)^n\prod_{i=1}^{n}a_i+\sum_{i=1}^{n}x_i(-1)^{n-1}a_1\cdots a_{i-1}a_{i+1}\cdots a_n=(-1)^{n-1}\prod_{i=1}^{n}a_i\left(\sum_{i=1}^{n}\frac{x_i}{a_i}-1\right)$$

$method5$

将最后一列拆分为$\begin{pmatrix} x_n \\ x_n \\ \vdots \\ x_n \end{pmatrix}+\begin{pmatrix} 0 \\ 0 \\ \vdots \\ -a_n \end{pmatrix}\Rightarrow D_n=-a_nD_{n-1}+x_n(-1)^{n-1}a_1a_2\cdots a_{n-1}\Rightarrow\frac{D_n}{(-1)^na_1\cdots a_n}-\frac{D_{n-1}}{(-1)^{n-1}a_1\cdots a_{n-1}}=-\frac{x_n}{a_n}$

$and\ D_1=\frac{x_1-a_1}{a_1}\Rightarrow D_n=(-1)^{n-1}\prod_{i=1}^{n}a_i\left(\sum_{i=1}^{n}\frac{x_i}{a_i}-1\right)$

Theorem 2.1.12 (柯西行列式)

$$|A|=\begin{vmatrix} (a_1+b_1)^{-1} & (a_1+b_2)^{-1} & \cdots & (a_1+b_n)^{-1} \\ (a_2+b_1)^{-1} & (a_2+b_2)^{-1} & \cdots & (a_2+b_n)^{-1} \\ \vdots & \vdots & & \vdots \\ (a_n+b_1)^{-1} & (a_n+b_2)^{-1} & \cdots & (a_n+b_n)^{-1} \end{vmatrix}$$

Proof $method1$

everyrow(i)extract $\prod_{j=1}^{n}(a_i+b_j)^{-1}\Rightarrow|A|=\prod_{i,j=1}^{n}(a_i+b_j)^{-1}\times|B|$

$|B|_{(i,j)}=\frac{\prod_{k=1}^{n}(a_i+b_k)}{a_i+b_j}=(a_i+b_1)(a_i+b_2)\cdots(a_i+b_{j-1})(a_i+b_{j+1})\cdots(a_i+b_n)$ now calculate$|B|$

$when\ a_i=a_j\ |B|=0 when\ b_i=b_j\ |B|=0 so\ |B| has\ factors\ (a_i-a_j)\ and\ (b_i-b_j)\ and\ a_i's\ degree\ is\ n-1$

$b_i\ is\ the\ same\ way\ so\ |B|=k\prod_{1\leqslant i<j\leqslant n}(a_i-a_j)(b_i-b_j)\ now\ we\ should\ confirm\ the\ \boldsymbol{k}\ let\ a_i=-b_i(i=1\cdots n)$

$|B| is$ diag $\{(a_1+b_2)\cdots(a_1+b_n),(a_2+b_1)(a_2+b_3)\cdots(a_2+b_n)\cdots\}$

so$|B|=\prod_{\substack{1\leqslant i,j\leqslant n\\ i\neq j}}(a_i-a_j)=\prod_{1\leqslant i<j\leqslant n}(a_i-a_j)(b_i-b_j)$ so$k=1$

$$\Rightarrow |A|=\frac{\prod_{1\leqslant i<j\leqslant n}(a_i-a_j)(b_i-b_j)}{\prod_{i,j=1}^{n}(a_i+b_j)}$$

method2

$$|A|=\begin{vmatrix}\frac{1}{a_1+b_1} & \frac{1}{a_1+b_2} & \cdots & \frac{1}{a_1+b_n}\\ \frac{1}{a_2+b_1} & \frac{1}{a_2+b_2} & \cdots & \frac{1}{a_2+b_n}\\ \vdots & \vdots & & \vdots\\ \frac{1}{a_n+b_1} & \frac{1}{a_n+b_2} & \cdots & \frac{1}{a_n+b_n}\end{vmatrix}$$

let$C_{1\sim n-1}=C_{1\sim n-1}-C_n$

$$|D_n|=\begin{vmatrix}\frac{b_n-b_1}{(a_1+b_1)(a_1+b_n)} & \cdots & \frac{b_n-b_{n-1}}{(a_1+b_{n-1})(a_1+b_n)} & \frac{1}{a_1+b_n}\\ \vdots & \vdots & & \\ \frac{b_n-b_1}{(a_{n-1}+b_1)(a_{n-1}+b_n)} & \cdots & \frac{b_n-b_{n-1}}{(a_{n-1}+b_{n-1})(a_{n-1}+b_n)} & \frac{1}{a_{n-1}+b_n}\\ \frac{b_n-b_1}{(a_n+b_1)(a_n+b_n)} & \cdots & \frac{b_n-b_{n-1}}{(a_n+b_{n-1})(a_n+b_n)} & \frac{1}{a_n+b_n}\end{vmatrix}=\frac{\prod_{i=1}^{n-1}(b_n-b_i)}{\prod_{j=1}^{n}(a_j+b_n)}\begin{vmatrix}\frac{1}{a_1+b_1} & \cdots & \frac{1}{a_1+b_{n-1}} & 1\\ \vdots & & \vdots & \vdots\\ \frac{1}{a_{n-1}+b_1} & \cdots & \frac{1}{a_{n-1}+b_{n-1}} & 1\\ \frac{1}{a_n+b_1} & \cdots & \frac{1}{a_n+b_{n-1}} & 1\end{vmatrix}$$

let$R_{1-n-1}=R_{1-n-1}-R_n$

$$\Rightarrow |D_n|=\frac{\prod_{i=1}^{n-1}(b_n-b_i)}{\prod_{j=1}^{n}(a_j+b_n)}\begin{vmatrix}\frac{a_n-a_1}{(a_1+b_1)(a_n+b_1)} & \cdots & \frac{a_n-a_1}{(a_1+b_{n-1})(a_n+b_{n-1})} & 0\\ \vdots & \vdots & & \\ \frac{a_n-a_{n-1}}{(a_{n-1}+b_1)(a_n+b_1)} & \cdots & \frac{a_n-a_{n-1}}{(a_{n-1}+b_{n-1})(a_n+b_{n-1})} & 0\\ \frac{1}{a_n+b_1} & \cdots & \frac{1}{a_n+b_{n-1}} & 1\end{vmatrix}=\frac{\prod_{i=1}^{n-1}(b_n-b_i)(a_n-a_i)}{\prod_{j=1}^{n}(a_j+b_n)\times\prod_{k=1}^{n-1}(a_n+b_k)}D_{n-1}$$

$$\Rightarrow D_n=\frac{\prod_{i=1}^{n-1}(b_n-b_i)(a_n-a_i)}{\prod_{j=1}^{n}(a_j+b_n)\times\prod_{k=1}^{n-1}(a_n+b_k)}D_{n-1}$$

由递推知道$D_n=\dfrac{\prod_{1\leqslant i<j\leqslant n}(a_j-a_i)(b_j-b_i)}{\prod_{i,j=1}^{n}(a_i+b_j)}$

Theorem 2.1.13 (循环矩阵行列式)

计算下列循环矩阵$\boldsymbol{A}$的行列式：$\boldsymbol{A}=\begin{pmatrix}a_1 & a_2 & a_3 & \cdots & a_n\\ a_n & a_1 & a_2 & \cdots & a_{n-1}\\ a_{n-1} & a_n & a_1 & \cdots & a_{n-2}\\ \vdots & \vdots & \vdots & & \vdots\\ a_2 & a_3 & a_4 & \cdots & a_1\end{pmatrix}$.

Proof 作多项式$f(x)=a_1+a_2x+a_3x^2+\cdots+a_nx^{n-1}$,

令$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$是1的所有$n$次方根.又令$\boldsymbol{V}=\begin{pmatrix}1&1&1&\cdots&1\\\varepsilon_1&\varepsilon_2&\varepsilon_3&\cdots&\varepsilon_n\\\varepsilon_1^2&\varepsilon_2^2&\varepsilon_3^2&\cdots&\varepsilon_n^2\\\vdots&\vdots&\vdots&&\vdots\\\varepsilon_1^{n-1}&\varepsilon_2^{n-1}&\varepsilon_3^{n-1}&\cdots&\varepsilon_n^{n-1}\end{pmatrix}$

则$\boldsymbol{AV}=\begin{pmatrix}f(\varepsilon_1)&f(\varepsilon_2)&f(\varepsilon_3)&\cdots&f(\varepsilon_n)\\\varepsilon_1f(\varepsilon_1)&\varepsilon_2f(\varepsilon_2)&\varepsilon_3f(\varepsilon_3)&\cdots&\varepsilon_nf(\varepsilon_n)\\\varepsilon_1^2f(\varepsilon_1)&\varepsilon_2^2f(\varepsilon_2)&\varepsilon_3^2f(\varepsilon_3)&\cdots&\varepsilon_n^2f(\varepsilon_n)\\\vdots&\vdots&\vdots&&\vdots\\\varepsilon_1^{n-1}f(\varepsilon_1)&\varepsilon_2^{n-1}f(\varepsilon_2)&\varepsilon_3^{n-1}f(\varepsilon_3)&\cdots&\varepsilon_n^{n-1}f(\varepsilon_n)\end{pmatrix}$因此$|\boldsymbol{A}\|\boldsymbol{V}|=|\boldsymbol{AV}|=f(\varepsilon_1)f(\varepsilon_2)\cdots f(\varepsilon_n)|\boldsymbol{V}|$.

因为ε_i互不相同,所以$|\boldsymbol{V}|\neq0$,从而$|\boldsymbol{A}|=f(\varepsilon_1)f(\varepsilon_2)\cdots f(\varepsilon_n)$.

Theorem 2.1.14

设$\boldsymbol{A}$为n阶方阵,$\boldsymbol{u},\boldsymbol{v}$为$n$维列向量,则

$\Longrightarrow\left|\boldsymbol{A}+\boldsymbol{uv}^{\mathrm{T}}\right|=|\boldsymbol{A}|+\boldsymbol{v}^{\mathrm{T}}\boldsymbol{A}^*\boldsymbol{u}$其中$\boldsymbol{A}^*$为$\boldsymbol{A}$的伴随矩阵

Proof (1)若$\boldsymbol{A}$可逆，由于$\begin{pmatrix}\boldsymbol{E}&-\boldsymbol{u}\\0&1\end{pmatrix}\begin{pmatrix}\boldsymbol{A}&\boldsymbol{u}\\-\boldsymbol{v}^{\mathrm{T}}&1\end{pmatrix}=\begin{pmatrix}\boldsymbol{A}+\boldsymbol{uv}^{\mathrm{T}}&0\\-\boldsymbol{v}^{\mathrm{T}}&1\end{pmatrix}$

$\begin{pmatrix}\boldsymbol{E}&0\\\boldsymbol{v}^{\mathrm{T}}\boldsymbol{A}^{-1}&1\end{pmatrix}\begin{pmatrix}\boldsymbol{A}&\boldsymbol{u}\\-\boldsymbol{v}^{\mathrm{T}}&1\end{pmatrix}=\begin{pmatrix}\boldsymbol{A}&\boldsymbol{u}\\0&1+\boldsymbol{v}^{\mathrm{T}}\boldsymbol{A}^{-1}\boldsymbol{u}\end{pmatrix}$上两式两边取行列式,可得

$\left|\boldsymbol{A}+\boldsymbol{uv}^{\mathrm{T}}\right|=|\boldsymbol{A}|\left(1+\boldsymbol{v}^{\mathrm{T}}\boldsymbol{A}^{-1}\boldsymbol{u}\right)=|\boldsymbol{A}|+\boldsymbol{v}^{\mathrm{T}}\boldsymbol{A}^*\boldsymbol{u}$

(2)若$\boldsymbol{A}$不可逆,则存在无穷多t的值使得$\boldsymbol{A}_1=\boldsymbol{A}+t\boldsymbol{E}$可逆.由(1)知$\begin{vmatrix}\boldsymbol{A}_1&\boldsymbol{u}\\-\boldsymbol{v}^{\mathrm{T}}&1\end{vmatrix}=|\boldsymbol{A}_1|+\boldsymbol{v}^{\mathrm{T}}\boldsymbol{A}_1^*\boldsymbol{u}$

上式两边均为t的多项式,且有无穷多t的值使得等式成立.故当$t=0$时,等式也成立.

Corollary 2.1

证明：$\begin{vmatrix}a_{11}+x&a_{12}+x&\cdots&a_{1n}+x\\a_{21}+x&a_{22}+x&\cdots&a_{2n}+x\\\vdots&\vdots&&\vdots\\a_{n1}+x&a_{n2}+x&\cdots&a_{nn}+x\end{vmatrix}=\left|a_{ij}\right|+x\sum_{i=1}^n\sum_{j=1}^nA_{ij}$,其中$A_{ij}$是$\left|a_{ij}\right|$中元素$a_{ij}$的代数余子式.

Proof 设$\boldsymbol{A}=\left(a_{ij}\right)$，则$D_n=\left|\boldsymbol{A}+\begin{pmatrix}x\\\vdots\\x\end{pmatrix}\begin{pmatrix}1,&\cdots&,1\end{pmatrix}\right|=|\boldsymbol{A}|+\begin{pmatrix}1,&\cdots&,1\end{pmatrix}\boldsymbol{A}^*\begin{pmatrix}x\\\vdots\\x\end{pmatrix}=\left|a_{ij}\right|+x\sum\limits_{i=1}^n\sum\limits_{j=1}^nA_{ij}$.故结论成立.

第 3 章 矩阵基础

3.1 矩阵的运算

Theorem 3.1.1

上(下)三角阵的和、差、数乘、乘积仍然是上(下)三角阵,且主对角元相对应

Proposition 3.1

设$\boldsymbol{A}$为实对称阵,若$\boldsymbol{A}^2=\boldsymbol{O}\Rightarrow\boldsymbol{A}=\boldsymbol{O}$

Proof 由题于$\boldsymbol{AA}'=\boldsymbol{O}\Rightarrow$所以$(\boldsymbol{AA}')_{ii}=a_{i1}^2+a_{i2}^2+\cdots+a_{in}^2=0\Rightarrow a_{i1}=a_{i2}=\cdots=a_{in}=0$
以上的i可取$1\sim n\Rightarrow\boldsymbol{A}=\boldsymbol{O}$

Theorem 3.1.2

与所有n阶对角阵乘法可交换的矩阵也必然是n阶对角阵
与所有n阶矩阵乘法可交换的矩阵为$k\boldsymbol{I}_n$纯量阵

Proof 设$A=\operatorname{diag}\{a_{11}\cdots a_{nn}\}\quad B=(b_{ij})_{n\times n}$
那么$(AB)_{ij}=a_{ii}b_{ij}=(BA)_{ij}=b_{ij}a_{jj}$当$i\neq j$时,但是$A$是任意的所以$a_{ii}$可以不等于$a_{jj}\Rightarrow b_{ij}=0$
所以仅$b_{ii}\neq0\Rightarrow B$是n阶可对角阵

(i) 用基础矩阵$\boldsymbol{E}_{ij}$即可.结合初等矩阵的知识
(ii) 由上题知至少为对角阵.再由第一类初等阵可交换得
(iii) 考虑第二类初等阵和第一类初等阵即可

Proof

设矩阵$\boldsymbol{A}=(a_{ij})$与所有n级矩阵可交换,则$\boldsymbol{A}$必为n级矩阵.
特别地,$\boldsymbol{A}$与n级基本矩阵$\boldsymbol{E}_{1j}(j=1,2,\cdots,n)$可交换,即$\boldsymbol{E}_{1j}\boldsymbol{A}=\boldsymbol{AE}_{1j}$

由此得出:$\begin{pmatrix}a_{j1}&a_{j2}&\cdots&a_{jj}&\cdots&a_{jn}\\0&0&\cdots&0&\cdots&0\\\vdots&\vdots&&\vdots&&\vdots\\0&0&\cdots&0&\cdots&0\end{pmatrix}=\begin{pmatrix}0&\cdots&0&a_{11}&0&\cdots&0\\0&\cdots&0&a_{21}&0&\cdots&0\\\vdots&&\vdots&\vdots&\vdots&&\vdots\\0&\cdots&0&a_{n1}&0&\cdots&0\end{pmatrix}$

于是$a_{j1}=0,\cdots,a_{j,j-1}=0,\quad a_{jj}=a_{11},\quad a_{j,j+1}=0,\cdots,a_{jn}=0$,由于$j$可取$1,2,\cdots,n$

因此$\boldsymbol{A}=\begin{pmatrix}a_{11}&0&0&\cdots&0&0\\0&a_{11}&0&\cdots&0&0\\\vdots&\vdots&\vdots&&\vdots&\vdots\\0&0&0&\cdots&0&a_{11}\end{pmatrix}$即$\boldsymbol{A}$是数量矩阵

笔记

由(i)知道,与所有奇异阵可交换的为纯量阵,(ii)所有初等阵可交换为纯量阵
(iii)所有正交阵可交换的为纯量阵

3.2 矩阵的逆阵

Definition 3.1 (逆阵定义)

设$\boldsymbol{A}$为n阶方阵, 若存在一个n阶方阵$\boldsymbol{B}$有$\boldsymbol{AB}=\boldsymbol{BA}=\boldsymbol{I}_n$

则称$\boldsymbol{B}$是$\boldsymbol{A}$的逆阵或者$\boldsymbol{A}$是$\boldsymbol{B}$的逆阵.记作$\boldsymbol{B}=\boldsymbol{A}^{-1}$

注 根据这个定义我们很容易的知道, 如果一个矩阵存在某一行或者某一列全部为 0 那么这个矩阵必定不可逆.

性质 1.逆阵唯一性

2.$\boldsymbol{A}$可逆则 $\left(\boldsymbol{A}^{-1}\right)^{-1}=\boldsymbol{A}$

3.$\boldsymbol{A},\boldsymbol{B}$均可逆则$\boldsymbol{AB}$可逆且$(\boldsymbol{AB})^{-1}=\boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$

4.$\boldsymbol{A}$可逆,c为一非零数,则$c\boldsymbol{A}$也可逆且$(c\boldsymbol{A})^{-1}=c^{-1}\boldsymbol{A}^{-1}$

5.若$\boldsymbol{A}$是可逆阵,则$\boldsymbol{A}'$也可逆且$(\boldsymbol{A}')^{-1}=\left(\boldsymbol{A}^{-1}\right)'$

Definition 3.2

设A是n阶方阵,A_{ij}是行列式$|A|$中第(i,j)元素a_{ij}的代数余子式

则称下列矩阵为矩阵A的伴随矩阵 : $\begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \cdots & \cdots & \cdots & \cdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix}$.$A$的伴随矩阵通常记为$A^*$.

Theorem 3.2.1

设$\boldsymbol{A}$为n阶方阵,A^*是A的伴随阵则$AA^*=A^*A=|A|\,I_n$

若$|A|\neq 0$，则A为可逆阵且$A^{-1}=\dfrac{1}{|A|}A^*$

3.3 初等变换与初等矩阵

Definition 3.3

下列3种矩阵变换分别称为矩阵的第一类、第二类、第三类初等行 (列) 变换 :

(1)对调矩阵中某两行 (列) 的位置;

(2)用一非零常数乘以矩阵的某一行 (列);

(3)将矩阵的某一行 (列) 乘以数c后加到另一行 (列) 上去.

上述3种变换统称为矩阵的初等变换.

Definition 3.4

一个矩阵 A 如果经过有限次初等变换变到 B，我们就称 A 与 B 等价或者相抵记作 $A\sim B$

Theorem 3.3.1

一个$m\times n$矩阵$A=(a_{ij})_{m\times n}$必相抵于下面形式的$m\times n$矩阵：$\begin{pmatrix} 1 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \cdots & 1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix}$.

上面的矩阵中前r行及前r列交点处有r个1, 其余元素皆为零.

换言之, 任一$m\times n$阵均与一个主对角线上元素等于1或0而其余元素均为0的$m\times n$矩阵相抵.

Proof 若$A=0$, 则结论显然成立.现设$A\neq 0$, 即A至少有个元素$a_{ij}\neq 0$.

如果a_{ij}不在第$(1,1)$位置, 那么可将它所在的行与第一行对换, 再将它所在的列与第一列对换就可将a_{ij}调至第$(1,1)$位置

所以我们不妨设$a_{11}\neq 0$.接下去将第一行依次乘以$-a_{11}^{-1}a_{i1}$加到第i行上去$(i=2,3,\cdots,m)$

于是第一列元素除a_{11}外都变成了零.再将第一列元素乘以一$-a_i'a_{1j}$后加到第j列上去$(j=2,3,\cdots,n)$

则第一行元素除了a_{11}外都变成零.再用a_{11}^{-1}乘以第一行, 就得到第$(1,1)$元素等于1而第一行及第一列其他元素都是零的矩阵

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & b_{22} & \cdots & b_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ 0 & b_{m2} & \cdots & b_{mn} \end{pmatrix}. \tag{1}$$

再对第二行第一列采用与上面相同的步骤使第$(2,2)$位置的元素不等于0, 并用同样办法消去除第$(2,2)$元素外第二行及第二列的所有元素显然, 在进行上述过程中第列及第一行的元素保持不变.这样不断做下去, 直到变成定理中式的形状为止

Theorem 3.3.2

设A是一个$m\times n$矩阵, 则经过若干次初等行变换, A可以化为阶梯形矩阵.

Proof 假定A的第一列元素全为零, 则初等变换从第二列开始.

现设A的第一列元素不全为零, 则用行对换将非零元素换到第$(1,1)$位置

再用第三类行初等变换即可将第一列其余汇素消为零.再看第二列, 如果这时从第$(2,2)$位置(包括$(2,2)$元素)

以下全为零, 则移至下一列.若否, 用行对换将非零元素换到$(2,2)$位算, 再用第三类初等变换消去这一列第$(2,2)$位置以下的元素.

就这样不断做下去即可得到

Definition 3.5

对单位阵I_n施以第一类、第二类、第三类初等变换后得到的矩阵分别称为第一类、第二类及第三类初等矩阵.

分别记为P_{ij}; $P_i(c)$; $T_{ij}(c)$ (第i行$\times c$加到第j行.或者第j列$\times c$加到第i列)

Theorem 3.3.3

设A是一个$m\times n$阵, 则对A作一次行初等变换后得到的矩阵等于用一个m阶相应的初等矩阵

(即第一类初等变换相应于第一类初等矩阵, 第二类初等变换相应于第二类初等矩阵, 等等) 左乘A后得到的积.

矩阵A作一次列初等变换后得到的矩阵等于用一个n阶相应的初等矩阵右乘A后所得到的积.

Proof 只需用定义证明即可

Proposition 3.2

1.初等矩阵都是非异阵且其逆阵仍是同类初等矩阵：$P_{ij}^{-1}=P_{ij},P_i(c)^{-1}=P_i\left(\frac{1}{c}\right),T_{ij}(c)^{-1}=T_{ij}(-c)$.

2.非异阵经初等变换后仍是非异阵, 奇异阵经初等变换后仍是奇异阵.

3.初等矩阵的行列式如下：$\left|P_{ij}\right|=-1,|P_i(c)|=c,\left|T_{ij}(c)\right|=1$.

Proof 证明由定理立即可得.证毕.

因为初等矩阵都是可逆阵, 可逆阵和可逆阵之积仍可逆, 所以第一个结论成立.
又设A是奇异阵, P是初等矩阵, 假定PA是非异阵, 注意到P^{-1}也是初等矩阵, 故$A=P^{-1}(PA)$将是非异阵, 矛盾
因此PA必是奇异阵.同理AP也是奇异阵.证毕.

证明因为单位阵的行列式等于1, 故交换单位阵两行而得到的矩阵P_{ij}的行列式等于−1.其余结论显然, 证毕.

Theorem 3.3.4

矩阵的相抵关系是一种等价关系

Proposition 3.3

矩阵的初等变换的对换操作可以通过另外两种来实现

Proof $$\begin{pmatrix}1&\cdots&0\\ \vdots&&\vdots\\ 0&\cdots&1\end{pmatrix}\xrightarrow{r_i=r_i+r_j}\begin{pmatrix}1&\cdots&1\\ \vdots&&\cdots\\ 0&\cdots&1\end{pmatrix}\xrightarrow{r_j=(-1)\cdot r_j}\begin{pmatrix}1&\cdots&1\\ \vdots&&\vdots\\ 0&\cdots&-1\end{pmatrix}\xrightarrow{r_j=r_j+r_i}\begin{pmatrix}1&\cdots&1\\ \vdots&&\vdots\\ 1&\cdots&0\end{pmatrix}\xrightarrow{r_i=r_i-r_j}\begin{pmatrix}0&\cdots&1\\ \vdots&&\vdots\\ 1&\cdots&0\end{pmatrix}$$

3.4 矩阵乘积的行列式和初等变换法

Lemma 3.1

设A是一个n阶可逆阵 (即非异阵), 则仅用行初等变换或仅用列初等变换就可以将它化为单位阵I_n. 且是当且仅当

Proof 我们先证明仅用行初等变换可以将A化为主对角元素全不等于零的上三角阵.因为A可逆,所以它没有整行或整列元素全为零
因此A的第一列至少有一个非零元素,通过行初等变换可以将它换到第 $(1,1)$位置.
用这个非零元素经过第三类行初等变换就可以将第一列元素 (除第 $(1,1)$元素) 全化为零.于是我们不妨假定

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \ldots \\ 0 & a_{n2} & \cdots & a_{nn} \end{pmatrix}. \tag{1}$$

我们现在要说明$a_{22},\cdots,a_{n2}$不全为零.若否,则通过第三类列初等变换用a_{11}可以消去a_{12},
于是非异阵A经过初等变换将化为一个第二列全为零的矩阵,即化为一个奇异阵,这与上一节的结论矛盾即命题.3.2
既然$a_{22},\cdots,a_{n2}$不全为零,我们又可以通过行初等变换将非零元素换到第 $(2,2)$位置上,再消去第二列除了第 $(2,2)$位置的元素.
如此下去就可以说明

Corollary 3.1

任一n阶非异阵均可表示成有限个初等矩阵的积, 且是当且仅当.

Proof 由上面的引理知道,存在有限个初等矩阵$P_1,\cdots,P_m$,使得$P_m\cdots P_1A=I_n$,因此$A=P_1^{-1}\cdots P_m^{-1}$.
而初等矩阵的逆阵仍是初等矩阵,故结论成立.证毕

Lemma 3.2

设A是一个n阶方阵, Q是一个n阶初等矩阵, 则$|QA|=|Q||A|=|AQ|$.

Proof 若$Q=P_{ij}$,则QA为A的第i行及第j行对换后所得之矩阵,其行列式的值等于$-|A|$.但$\left|P_{ij}\right|=-1$,因此$\left|P_{ij}A\right|=-|A|=\left|P_{ij}\right||A|$.
若$Q=P_i(c)(c\neq 0)$,同样不难验证$|P_i(c)A|=c|A|=|P_i(c)||A|$.最后,$\left|T_{ij}(c)A\right|=|A|$,$\left|T_{ij}(c)\right|=1$,故$\left|T_{ij}(c)A\right|=\left|T_{ij}(c)\right||A|$.
当Q作用在A右侧时也可类似证明.证毕.

Theorem 3.4.1

一个n阶方阵A为非异阵的充分必要条件是它的行列式的值不等于零.

Proof 根据定理3.2.1知道 若$|A|\neq 0$,则A是非异阵,故只需证明若A非异,则$|A|\neq 0$.
由上述推论,非异阵A等于有限个初等矩阵之积,设$A=P_1P_2\cdots P_t$,则从上面的引理知道,
$|A|=|P_1||P_2|\cdots|P_k|$.由于初等矩阵的行列式$\neq 0$.故$|A|\neq 0$,证毕

Theorem 3.4.2

设A,B都是n阶矩阵, 则$|AB|=|A||B|$.

Proof 第一种情形,设A是非异阵.则存在若干个初等矩阵$Q_1,\cdots,Q_m$,使$A=Q_1\cdots Q_m$,故
$$\begin{aligned}|AB|&=|Q_1\cdots Q_mB|=|Q_1|\cdots|Q_m||B|\\&=|Q_1\cdots Q_m||B|=|A||B|.\end{aligned}$$
第二种情形,设A为奇异阵,这时$|A|=0$,故只需证明$|AB|=0$.因为A是奇异阵,故存在初等矩阵$P_1,\cdots,P_s;Q_1,\cdots,Q_r$,

使得$P_s\cdots P_1AQ_1\cdots Q_r=D$, 其中$D$是$A$的相抵标准型, 它是一个对角阵, 主对角元素为1或0.
因为A奇异, 所以D也是奇异阵, 至少最后一行全为零.又$P_s\cdots P_1A=DQ_r^{-1}\cdots Q_1^{-1}$.
由矩阵乘法知道, 因为D的第n行元素全为零, 所以$DQ_r{}^{-1}\cdots Q_1^{-1}$从而$P_s\cdots P_1A$的第$n$行等于零.
于是$P_s\cdots P_1AB$的第n行也等于零.
故 $|P_s|\cdots|P_1|\,|AB|=|P_s\cdots P_1AB|=0$, 其中 $|P_i|\neq 0(i=1,\cdots,s)$, 于是$|AB|=0=|A||B|$.证毕.

Corollary 3.2

一个奇异阵和一个方阵的乘积仍为奇异阵. 两个非异阵乘积为非异阵
若A是非异阵, 则 $\left|A^{-1}\right|=|A|^{-1}$

Example 3.1 计算下列$n+1$阶矩阵A的行列式 : $A=\begin{pmatrix}(a_0+b_0)^n & (a_0+b_1)^n & \cdots & (a_0+b_n)^n\\ (a_1+b_0)^n & (a_1+b_1)^n & \cdots & (a_1+b_n)^n\\ \cdots & \cdots & \cdots & \cdots\\ (a_n+b_0)^n & (a_n+b_1)^n & \cdots & (a_n+b_n)^n\end{pmatrix}$.

Proof 将A分解为两个矩阵之积 : $A=\begin{pmatrix}1 & C_n^1a_0 & C_n^2a_0^2 & \cdots & a_0^n\\ 1 & C_n^1a_1 & C_n^2a_1^2 & \cdots & a_1^n\\ \cdots & \cdots & \cdots & \cdots & \cdots\\ 1 & C_n^1a_n & C_n^2a_n^2 & \cdots & a_n^n\end{pmatrix}\begin{pmatrix}b_0^n & b_1^n & b_2^n & \cdots & b_n^n\\ b_0^{n-1} & b_1^{n-1} & b_2^{n-1} & \cdots & b_n^{n-1}\\ \cdots & \cdots & \cdots & \cdots & \cdots\\ 1 & 1 & 1 & \cdots & 1\end{pmatrix}$.
上式左边矩阵的行列式每一列提出公因子后就是一个$VanderMonde$行列式.
右边矩阵的行列式也可以化为$VanderMonde$行列式并求出其值, 于是$|A|=C_n^1C_n^2\cdots C_n^n\prod\limits_{0\leqslant j<i\leqslant n}\left(a_i-a_j\right)\left(b_j-b_i\right)$.

Theorem 3.4.3

设A,C分别是m,n阶方阵, 则对分块上 (下) 三角行列式有 : $G=\begin{vmatrix}A & B\\ 0 & C\end{vmatrix}=|A||C|, H=\begin{vmatrix}A & 0\\ B & C\end{vmatrix}=|A||C|$.

Proof 本命题可以用$Laplace$定理立即得到, 但是我们采用另外一种方法来证明.只对第一式进行证明.
对A的阶m月归纳法, $m=1$时, 结论显然.假定对左上角是$m-1$阶矩阵的分块上三角行列式结论为真.
设$A=\begin{pmatrix}a_{11} & a_{12} & \cdots & a_{1m}\\ a_{21} & a_{22} & \cdots & a_{2m}\\ \cdots & \cdots & \cdots & \cdots\\ a_{m1} & a_{m2} & \cdots & a_{mm}\end{pmatrix}$
将G按第一列展开 (注意到第一列从第$m+1$行起的元素全为零) :
$|G|=a_{11}G_{11}+a_{21}G_{21}+\cdots+a_{m1}G_{m1}$, 其中$G_{i1}$是$a_{i1}$在$G$中的代数余子式.
与每个G_{i1}相应的余子式是一个左上角为$m-1$阶矩阵的分块上三角行列式, 由归纳假设可得$G_{i1}=A_{i1}|C|$
其中A_{i1}是元素a_{i1}在A中的代数余子式.因此$|G|=a_{11}A_{11}|C|+a_{21}A_{21}|C|+\cdots+a_{m1}A_{m1}|C|=|A||C|$.证毕.

Theorem 3.4.4 (打洞原理)

$$\begin{vmatrix}A & B\\ C & D\end{vmatrix}=\begin{cases}\text{若}A\text{可逆, 则 }=|A|\left|D-CA^{-1}B\right|\\ \text{若}D\text{可逆, 则 }=|D|\left|A-BD^{-1}C\right|\\ \text{若}A,D\text{可逆, 则 }|A|\left|D-CA^{-1}B\right|=|D|\left|A-BD^{-1}C\right|\end{cases}$$

Corollary 3.3

第一降阶公式

设A为n阶可逆矩阵,α,β为n维列向量,则$\left|A+\alpha\beta^T\right|=|A|\left(1+\beta^TA^{-1}\alpha\right)$.

第二降阶公式

设A为n阶可逆矩阵,B_1和B_2分别为$n\times m$和$m\times n$矩阵,则$|A+B_1B_2|=|A|\left|E_m+B_2A^{-1}B_1\right|$.

Proof 设A为n阶可逆矩阵,α,β为n维列向量,则$\left|A+\alpha\beta^T\right|=|A|\left(1+\beta^TA^{-1}\alpha\right)$.

$$\begin{pmatrix}E&0\\\beta^TA^{-1}&1\end{pmatrix}\begin{pmatrix}A&\alpha\\-\beta^T&1\end{pmatrix}=\begin{pmatrix}A&\alpha\\0&1+\beta^TA^{-1}\alpha\end{pmatrix}\qquad\begin{pmatrix}A&\alpha\\-\beta^T&1\end{pmatrix}\begin{pmatrix}E&0\\\beta^T&1\end{pmatrix}=\begin{pmatrix}A+\alpha\beta^T&\alpha\\0&1\end{pmatrix}$$

两边取行列式,则$\begin{vmatrix}A&\alpha\\0&1+\beta^TA^{-1}\alpha\end{vmatrix}=\begin{vmatrix}A&\alpha\\-\beta^T&1\end{vmatrix}=\begin{vmatrix}A+\alpha\beta^T&\alpha\\0&1\end{vmatrix}$.

即$\left|A+\alpha\beta^T\right|=|A|\left(1+\beta^TA^{-1}\alpha\right)$.

设A为n阶可逆矩阵,B_1和B_2分别为$n\times m$和$m\times n$矩阵,则$|A+B_1B_2|=|A|\left|E_m+B_2A^{-1}B_1\right|$.

$$\begin{pmatrix}E_n&0\\B_2A^{-1}&E_m\end{pmatrix}\begin{pmatrix}A&B_1\\-B_2&E_m\end{pmatrix}=\begin{pmatrix}A&B_1\\0&E_m+B_2A^{-1}B_1\end{pmatrix}\qquad\begin{pmatrix}A&B_1\\-B_2&E_m\end{pmatrix}\begin{pmatrix}E_n&0\\B_2&E_m\end{pmatrix}=\begin{pmatrix}A+B_1B_2&B_1\\0&E_m\end{pmatrix}$$

两边取行列式,则$\begin{vmatrix}A&B_1\\0&E_m+B_2A^{-1}B_1\end{vmatrix}=\begin{vmatrix}A&B_1\\-B_2&E_m\end{vmatrix}=\begin{vmatrix}A+B_1B_2&B_1\\0&E_m\end{vmatrix}$.

即$|A+B_1B_2|=|A|\left|E_m+B_2A^{-1}B_1\right|$.

Proposition 3.4

设$\boldsymbol{A},\boldsymbol{B}$分别是$s\times n,n\times s$矩阵,则

(1) $\begin{vmatrix}\boldsymbol{I}_n&\boldsymbol{B}\\\boldsymbol{A}&\boldsymbol{I}_s\end{vmatrix}=|\boldsymbol{I}_s-\boldsymbol{AB}|$;

(2) $\begin{vmatrix}\boldsymbol{I}_n&\boldsymbol{B}\\\boldsymbol{A}&\boldsymbol{I}_s\end{vmatrix}=|\boldsymbol{I}_n-\boldsymbol{BA}|$;

(3) $|\boldsymbol{I}_s-\boldsymbol{AB}|=|\boldsymbol{I}_n-\boldsymbol{BA}|$

Proof (1)设法把左端变成分块上三角矩阵的行列式

为此做分块矩阵的初等行变换:$\begin{pmatrix}\boldsymbol{I}_n&\boldsymbol{B}\\\boldsymbol{A}&\boldsymbol{I}_s\end{pmatrix}\xrightarrow{(2)+(-\boldsymbol{A})\cdot(1)}\begin{pmatrix}\boldsymbol{I}_n&\boldsymbol{B}\\0&\boldsymbol{I}_s-\boldsymbol{AB}\end{pmatrix}$,于是有$\begin{pmatrix}\boldsymbol{I}_n&0\\-\boldsymbol{A}&\boldsymbol{I}_s\end{pmatrix}\begin{pmatrix}\boldsymbol{I}_n&\boldsymbol{B}\\\boldsymbol{A}&\boldsymbol{I}_s\end{pmatrix}=\begin{pmatrix}\boldsymbol{I}_n&\boldsymbol{B}\\0&\boldsymbol{I}_s-\boldsymbol{AB}\end{pmatrix}$

在上式两边取行列式,得$\begin{vmatrix}\boldsymbol{I}_n&0\\-\boldsymbol{A}&\boldsymbol{I}_s\end{vmatrix}\begin{vmatrix}\boldsymbol{I}_n&\boldsymbol{B}\\\boldsymbol{A}&\boldsymbol{I}_s\end{vmatrix}=\begin{vmatrix}\boldsymbol{I}_n&\boldsymbol{B}\\0&\boldsymbol{I}_s-\boldsymbol{AB}\end{vmatrix}$

由此得出$|\boldsymbol{I}_n|\,|\boldsymbol{I}_s|\begin{pmatrix}\boldsymbol{I}_n&\boldsymbol{B}\\\boldsymbol{A}&\boldsymbol{I}_s\end{pmatrix}=|\boldsymbol{I}_n|\,|\boldsymbol{I}_s-\boldsymbol{AB}|$,从而$\begin{vmatrix}\boldsymbol{I}_n&\boldsymbol{B}\\\boldsymbol{A}&\boldsymbol{I}_s\end{vmatrix}=|\boldsymbol{I}_s-\boldsymbol{AB}|$

(2) 类似于第 (1)小题的证法,请同学们自己写出

(3) 由第 (1)(2) 小题即得

Corollary 3.4

设A、B分别是$s\times n$、$n\times s$矩阵,且$\lambda\neq0$

证明:$\lambda^n|\lambda I,-AB|=\lambda^s\,|\lambda I_n-BA|$.

同理运用上面的证明方法只需把I_n变为λI_n

3.5 Laplace 定理

Definition 3.6

设 $|A|$ 是一个n阶行列式, $k<n$; $i_1,i_2\cdots i_k$及$j_1\cdots j_k$是两组指标满足:

$1\leqslant i_1<i_2<\cdots<i_k\leqslant n; 1\leqslant j_1<j_2<\cdots<j_k\leqslant n;$取上述行列交点的元素按照原来的相对位置组成一个$k$阶行列式

我们称为 $|A|$ 的一个k阶子式记作$A\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{pmatrix}$; 划去这些行列, 剩下的元素按照原来相对位置组成一个$n-k$

阶行列式称为上述子式的余子式记$M\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{pmatrix}$.记$p=i_1+\cdots+i_k; q=j_1+\cdots+j_k$

记$\widehat{A}\begin{pmatrix} li_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{pmatrix}=(-1)^{p+q}M\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{pmatrix}$称为代数余子式

特别的当行列选取相等称为主子式

Theorem 3.5.1 (Laplace 定理)

设$|A|$是n阶行列式, 在$|A|$中任取k行 (列), 那么含于这k行 (列) 的全部k阶子式与它们对应的代数余子式的乘积之和等于$|A|$.

即若取定k个行: $1\leqslant i_1<i_2<\cdots<i_k\leqslant n$, 则$|A|=\displaystyle\sum_{1\leqslant j_1<j_2<\cdots<j_k\leqslant n} A\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{pmatrix}\widehat{A}\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{pmatrix}$.

同样若取定k个列: $1\leqslant j_1<j_2<\cdots<j_k\leqslant n$, 则$|A|=\displaystyle\sum_{1\leqslant i_1<i_2<\cdots<i_k\leqslant n} A\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{pmatrix}\widehat{A}\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{pmatrix}$.

Proof 分析:

1.$|A|$ 有组合定义知道有$n!$个单项

2.$Laplace$定理等式右边对于选定的k行 (列) 有C_n^k选法, A中$k!$个单项, $\widehat{A}$中有 $(n-k)!$个单项那么一共就$n!$个

关键证明: 1.等式右边单项都是 $|A|$ 中的一个且符号一致. 2.等式右边的单项各个不相同

$case1$.

选取的刚好就是前k行与前k列, 即$i_1=1, i_2=2\cdots i_k=k; j_1=1\cdots j_k=k$

$|A|=\begin{vmatrix} A_1 & * \\ * & A_2 \end{vmatrix}$. $|A_1|=A\begin{pmatrix} 1 & 2 & \cdots & k \\ 1 & 2 & \cdots & k \end{pmatrix}$; $|A_2|=M\begin{pmatrix} 1 & 2 & \cdots & k \\ 1 & 2 & \cdots & k \end{pmatrix}$

$|A_1|$ 中任一单项: $(-1)^{N(r_1\cdots r_k)}a_{r_1 1}a_{r_2}\cdots a_{r_k k}\quad (r_1\cdots r_k)\in(1,2,\cdots k)$

$|A_2|$ 中任一单项: $(-1)^{N(r_{k+1}\cdots r_n)}a_{r_{k+1},k+1}\cdots a_{r_n n}\quad (r_{k+1}\cdots r_n)\in(k+1,k+2\cdots n)$

那么 $|A_1||A_2|$ 中任一单项具有形式

$(-1)^{\sigma}a_{r_1 1}a_{r_2 2}\cdots a_{r_k k}a_{r_{k+1},k+1}\cdots a_{r_n n}$.那么 $(r_1\cdots r_k r_{k+1}\cdots r_n)$ 是 $(1,2\cdots n)$的全排列其中$\sigma=N(r_1\cdots r_k)+N(r_{k+1}\cdots r_n)$

那么等式右边$A\begin{pmatrix} 1 & 2 & \cdots & k \\ 1 & 2 & \cdots & k \end{pmatrix}\widehat{A}\begin{pmatrix} 1 & 2 & \cdots & k \\ 1 & 2 & \cdots & k \end{pmatrix}$某个单项就为 $(-1)^{2(1+\cdots+k)}(-1)^{\sigma}a_{r_1 1}a_{r_2 2}\cdots a_{r_k k}a_{r_{k+1},k+1}\cdots a_{r_n n}$

就为 $(-1)^{\sigma}a_{r_1 1}a_{r_2 2}\cdots a_{r_k k}a_{r_{k+1},k+1}\cdots a_{r_n n}$

而$|A|$中的单项显然与上式相同, 且符号由$N(r_1\cdots r_k r_{k+1}\cdots r_n)$ 决定 $=N(r_1\cdots r_k)+N(r_{k+1}\cdots r_n)=\sigma$

$r_1\sim r_k$中看后面有多少比它小, $r_{k+1}\sim r_n$看前面有多少个比它大, 贡献为0

这就证明了 (1)的$case1$较为简单的情形

$case2$:

一般的情形, 我们总可以通过相邻对换变到$case1$

第i_1行经过i_1-1次到第1行; 第i_2行经过i_2-2次到第2行...同理对列也这样操作即可得到

最终经过$\underbrace{i_1+\cdots+i_k}_{p}+\underbrace{j_1+\cdots+j_k}_{q}-(k+1)k$对换变化到

$|C|=\begin{vmatrix} D & * \\ * & B \end{vmatrix}$ 其中$|C|=(-1)^{p+q}|A|$其中$|D|=A\begin{pmatrix} i_1 & \cdots & i_k \\ j_1 & \cdots & j_k \end{pmatrix}, |B|=M\begin{pmatrix} i_1 & \cdots & i_k \\ j_1 & \cdots & j_k \end{pmatrix}$, $\left|\widehat{B}\right|=\widehat{A}\begin{pmatrix} i_1 & \cdots & i_k \\ j_1 & \cdots & j_k \end{pmatrix}=|B|(-1)^{p+q}$

由$case1$知道$|D||B|$的每一个单项都是$|C|$的单项且符号相同(在$case1$中因为$2(1+\cdots+k)$贡献无用)

$\Rightarrow |D||B|(-1)^{p+q}$ 与$|C|(-1)^{p+q}$ 单项相同且符号一致且$(-1)^{p+q}|C|=|A|$

$\Rightarrow A\begin{pmatrix} i_1 & \cdots & i_k \\ j_1 & \cdots & j_k \end{pmatrix}\widehat{A}\begin{pmatrix} i_1 & \cdots & i_k \\ j_1 & \cdots & j_k \end{pmatrix}$ 与$|A|$单项一致且符号保持一致。

预开始由题条件知道选取$1\leqslant i_1<\cdots<i_k\leqslant n$;剩余指标为$1\leqslant i_{k+1}<\cdots<i_n\leqslant n$

选取$1\leqslant j_1<\cdots<j_k\leqslant n$;剩余指标为$1\leqslant j_{k+1}<\cdots<j_n\leqslant n$

另选取$1\leqslant l_1<\cdots<l_k\leqslant n$;剩余指标为$1\leqslant l_{k+1}<\cdots<l_n\leqslant n$

现在我们取$(j_1\cdots j_k)$的一个全排列$(r_1\cdots r_k)$那么剩余指标的全排列$(r_{k+1}\cdots r_n)$

取$(l_1\cdots l_k)$的一个全排列$(s_1\cdots s_k)$那么剩余指标的全排列为$(s_{k+1}\cdots s_n)$

那么RHS的一个单项因为$a_{i_1r_1}\cdots a_{i_kr_k}a_{i_{k+1}r_{k+1}}\cdots a_{i_nr_n}$。另外也为$a_{i_1s_1}\cdots a_{i_ks_k}a_{i_{k+1}s_{k+1}}\cdots a_{i_ns_n}$

从逆否命题来看,若这两个相同,那么应当$r_1=s_1\cdots r_n=s_n$那么$(j_1\cdots j_k)=(l_1\cdots l_k)$

Proposition 3.5

设$A_{n\times n}$; $B_{n\times n}$则 $|AB|=|A||B|$

Proof 构建$C=\begin{bmatrix} A & O \\ -I_n & B \end{bmatrix}$则

$$C=\left[\begin{array}{ccc|ccc} a_{11} & \cdots & a_{1n} & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} & 0 & \cdots & 0 \\ \hline -1 & \cdots & 0 & b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & \cdots & -1 & b_{n1} & \cdots & b_{nn} \end{array}\right]$$

进行初等行变换得到

$$|C|=\left|\begin{array}{ccc|ccc} 0 & \cdots & 0 & \sum a_{1r}b_{r1} & \cdots & \sum a_{1r}b_{rn} \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & \sum a_{nr}b_{r1} & \cdots & \sum a_{nr}b_{rn} \\ \hline -1 & \cdots & 0 & b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & -1 & b_{n1} & \cdots & b_{nn} \end{array}\right|=\begin{vmatrix} O & AB \\ -I_n & B \end{vmatrix}$$

对两个式子分别使用$Laplace$定理对第一式前n行展开$=|A||B|$,第二式前n行展开$|AB|\Rightarrow |AB|=|A||B|$

3.6 Cauchy-Binet 公式

Theorem 3.6.1 (Cauchy-Binet 公式)

设$A=(a_{ij})$ 是$m\times n$矩阵, $B=(b_{ij})$ 是$n\times m$矩阵

(1)若$m>n$, 必有$|AB|=0$

(2)若$m\leqslant n$, 必有$|AB|=\sum\limits_{1\leqslant j_1<j_2<\cdots<j_m\leqslant n} A\begin{pmatrix}1 & 2 & \cdots & m\\ j_1 & j_2 & \cdots & j_m\end{pmatrix} B\begin{pmatrix}j_1 & j_2 & \cdots & j_m\\ 1 & 2 & \cdots & m\end{pmatrix}$

Proof 构建$m+n$阶方阵$\begin{bmatrix}A_{m\times n} & O_{m\times m}\\ -I_n & B_{n\times m}\end{bmatrix}\to\begin{bmatrix}O & (AB)_{m\times m}\\ -I_n & B_{n\times m}\end{bmatrix}$则有$|LHS|=|RHS|$

$(1):|RHS|$对前m行$Laplace$展开为$|AB|=|AB|(-1)^{1+2\cdots+m+(n+1)+\cdots+(n+m)}\cdot|-I_n|=|AB|(-1)^{n(m+1)}$

(其中1与$n+1$的1配对, 2与$n+2$的2配对$\cdots$还有$|-I_n|$里拆出来的负号

$(2):|LHS|$前m行展开.

$case1:m>n$, 则至少一列包含0所以$|LHS|=0\Rightarrow|RHS|=0\Rightarrow|AB|=0$

$case2:m\leqslant n$, 则$|LHS|=\sum\limits_{1\leqslant j_1<\cdots<j_m\leqslant n} A\begin{pmatrix}l1 & 2 & \cdots & m\\ j_1 & j_2 & \cdots & j_m\end{pmatrix} C\begin{pmatrix}l1 & 2 & \cdots & m\\ j_1 & j_2 & \cdots & j_m\end{pmatrix}$ (C为对应的代数余子式用新符号写罢)

而$C\begin{pmatrix}1 & 2 & \cdots & m\\ j_1 & j_2 & \cdots & j_m\end{pmatrix}=(-1)^{1+2+\cdots+m+j_1+j_2+\cdots j_m}\cdot\left|-e_{i_1},-e_{i_2},\cdots-e_{i_{n-m}},B\right|$ (其中$1\leqslant i_1<i_2<i_{n-m}\leqslant n$为$j_1\cdots j_m$剩下的余指标)

$\left|-e_{i_1},-e_{i_2},\cdots-e_{i_{n-m}},B\right|$计算该式子利用对前$n-m$列$Laplace$定理展开。只有一个不为0即刚好也选取对应拿到$-1$元素的行

即$|-I_{n-m}|\,(-1)^{1+2+\cdots+(n-m)+i_1+i_2+\cdots i_{n-m}}\times B\begin{pmatrix}j_1 & j_2 & \cdots & j_m\\ 1 & 2 & \cdots & m\end{pmatrix}$

(因为$1\leqslant i_1<i_2<i_{n-m}\leqslant n$为$j_1\cdots j_m$在$1\sim n$剩下的余指标, 现在在$n$行中又选取了$i_1i_2\cdots i_{n-m}$指标相当于又回去了)

那么$|LHS|=(-1)^{\sigma}\sum\limits_{1\leqslant j_1<\cdots<j_m\leqslant n} A\begin{pmatrix}l1 & 2 & \cdots & m\\ j_1 & j_2 & \cdots & j_m\end{pmatrix} B\begin{pmatrix}j_1 & j_2 & \cdots & j_m\\ 1 & 2 & \cdots & m\end{pmatrix}$ 其中$\sigma=\begin{cases}1+2+\cdots+m\\ j_1+\cdots+j_m+i_1+\cdots+i_{n-m}\\ n-m\\ 1+2+\cdots+(n-m)\end{cases}$

那么现在则需要证明$(-1)^{\sigma}=(-1)^{n(m+1)}$ 我们只需证明$\sigma+n\,(m+1)$为偶数

即$(1+\cdots+m)+(1+\cdots+n)+(1+\cdots+(n-m-1))+n\,(m+1)$ (约去两个$n-m$)

$=\underbrace{[(m+1)+\cdots+n]+[0+1+\cdots+(n-m-1)]}_{\text{首位配对}}+n\,(m+1)$

$=n\,(n-m)+n\,(m+1)=n\,(n+1)$ 偶数

证毕

Theorem 3.6.2

设$A=(a_{ij})$ 是$m\times n$矩阵, $B=(b_{ij})$ 是$n\times m$矩阵, r是一个自然数且$r\leqslant m$, 则

(1) 若$r>n$, 则AB的任意一个r阶子式等于零

(2) 若$r\leqslant n$, 则AB的r阶子式$AB\begin{pmatrix}i_1 & i_2 & \cdots & i_r\\ j_1 & j_2 & \cdots & j_r\end{pmatrix}=\sum\limits_{1\leqslant k_1<k_2<\cdots<k_r<n} A\begin{pmatrix}i_1 & i_2 & \cdots & i_r\\ k_1 & k_2 & \cdots & k_r\end{pmatrix} B\begin{pmatrix}k_1 & k_2 & \cdots & k_r\\ j_1 & j_2 & \cdots & j_r\end{pmatrix}$.

Proof 设$C=AB$, 则C为m阶矩阵, 且$c_{ij}=a_{i1}b_{1j}+\cdots+a_{in}b_{nj}$因此

$$C\begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ j_1 & j_2 & \cdots & j_r \end{pmatrix} = \begin{pmatrix} a_{i_1 1} & a_{i_1 2} & \cdots & a_{i_1 n} \\ a_{i_2 1} & a_{i_2 2} & \cdots & a_{i_2 n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{i_r 1} & a_{i_r 2} & \cdots & a_{i_r n} \end{pmatrix} \begin{pmatrix} b_{1 j_1} & b_{1 j_2} & \cdots & b_{1 j_r} \\ b_{2 j_1} & b_{2 j_2} & \cdots & b_{2 j_r} \\ \cdots & \cdots & \cdots & \cdots \\ b_{n j_1} & b_{n j_2} & \cdots & b_{n j_r} \end{pmatrix}$$

由上述定理可知：当$r > n$时, $C\begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ j_1 & j_2 & \cdots & j_r \end{pmatrix} = 0$;

当$r \leqslant n$时, $C\begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ j_1 & j_2 & \cdots & j_r \end{pmatrix} = \sum\limits_{1k_1<k_2<\cdots<k_r<n} A\begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ k_1 & k_2 & \cdots & k_r \end{pmatrix} B\begin{pmatrix} k_1 & k_2 & \cdots & k_r \\ j_1 & j_2 & \cdots & j_r \end{pmatrix}$.

证毕.

Corollary 3.5

1.设A为$m \times n$阶实矩阵, 则矩阵AA'的任一阶主子式非负

2.$A_{m\times n}$此时选定$i_1 \sim i_r$行, 求所有r阶子式的平方和, 实际上有

$$\boldsymbol{AA'}\begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ i_1 & i_2 & \cdots & i_r \end{pmatrix} = \sum_{1\le j_1<j_2<\cdots<j_r\le n} \boldsymbol{A}\begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ j_1 & j_2 & \cdots & j_r \end{pmatrix}^2$$

Proof 我们来证明1.

若$r \le n$, 则由$Cauchy-Binet$公式可得$\boldsymbol{AA'}\begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ i_1 & i_2 & \cdots & i_r \end{pmatrix} = \sum\limits_{1\le j_1<j_2<\cdots<j_r\le n} \boldsymbol{A}\begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ j_1 & j_2 & \cdots & j_r \end{pmatrix}^2 \ge 0$;

若$r > n$, 则$\boldsymbol{AA'}$的任一r阶主子式都等于零, 结论也成立.

Theorem 3.6.3 ($Lagrange$恒等式)

$$\left(\sum_{i=1}^n a_i^2\right)\left(\sum_{i=1}^n b_i^2\right) - \left(\sum_{i=1}^n a_i b_i\right)^2 = \sum_{1\leqslant i<j\leqslant n} (a_i b_j - a_j b_i)^2. \ (n \geqslant 2)$$

Proof 左边的式子等于$\begin{vmatrix} \sum\limits_{i=1}^n a_i^2 & \sum\limits_{i=1}^n a_i b_i \\ \sum\limits_{i=1}^n a_i b_i & \sum\limits_{i=1}^n b_i^2 \end{vmatrix}$这个行列式对应的矩阵可化为$\begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ b_1 & b_2 & \cdots & b_n \end{pmatrix}\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \\ \vdots & \vdots \\ a_n & b_n \end{pmatrix}$

用$Cauchy-Binet$公式得$\begin{vmatrix} \sum\limits_{i=1}^n a_i^2 & \sum\limits_{i=1}^n a_i b_i \\ \sum\limits_{i=1}^n a_i b_i & \sum\limits_{i=1}^n b_i^2 \end{vmatrix} = \sum\limits_{1\leqslant i<j\leqslant n} \begin{vmatrix} a_i & a_j \\ b_i & b_j \end{vmatrix}\begin{vmatrix} a_i & b_i \\ a_j & b_j \end{vmatrix} = \sum\limits_{1\leqslant i<j\leqslant n} (a_i b_j - a_j b_i)^2$.

证毕.

Theorem 3.6.4 (Cauchy-Schwarz 不等式)

设a_i, b_i都是实数$Cauchy-Schwarz$不等式$\left(\sum\limits_{i=1}^n a_i^2\right)\left(\sum\limits_{i=1}^n b_i^2\right) \geqslant \left(\sum\limits_{i=1}^n a_i b_i\right)^2$

Lemma 3.3

$$\sum a_i b_i \cdot \sum \overline{a_i b_i} = \left|\sum_{i=1}^{n} a_i b_i\right|^2$$

Proof 不防令$z_k = a_k b_k \quad z_k = x_k + y_k i$

那么$\sum a_i b_i \cdot \sum \overline{a_i b_i} = [(x_1 + \cdots x_n) + (y_1 + \cdots + y_n) i] [(x_1 + \cdots + x_n) - (y_1 + \cdots + y_n) i]$

$= (x_1 + \cdots + x_n)^2 + (y_1 + \cdots y_n)^2 = \left|\sum_{i=1}^{n} a_i b_i\right|^2$

Theorem 3.6.5 (复数形式拉格朗日恒等式)

a_i, b_i都是复数

$$\left|\sum_{i=1}^{n} a_i b_i\right|^2 = \left(\sum_{i=1}^{n} |a_i|^2\right)\left(\sum_{i=1}^{n} |b_i|^2\right) - \sum_{1 \leqslant i < j \leqslant n} \left|a_i \overline{b_j} - a_j \overline{b_i}\right|^2$$

Proof 构造$\begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ \overline{b_1} & \overline{b_2} & \cdots & \overline{b_n} \end{pmatrix} \begin{pmatrix} \overline{a_1} & b_1 \\ \overline{a_2} & b_2 \\ \vdots & \vdots \\ \overline{a_n} & b_n \end{pmatrix} = \begin{pmatrix} \sum |a_i|^2 & \sum a_i b_i \\ \sum \overline{a_i b_i} & \sum |b_i|^2 \end{pmatrix}$

$|LHS| = \sum\limits_{1 \leqslant i < j \leqslant n} \begin{vmatrix} la_i & a_j \\ \overline{b_i} & \overline{b_j} \end{vmatrix} \begin{vmatrix} \overline{a_i} & b_i \\ \overline{a_j} & b_j \end{vmatrix} = \sum\limits_{1 \leqslant i < j \leqslant n} \left(a_i \overline{b_j} - a_j \overline{b_i}\right)\left(\overline{a_i} b_j - b_i \overline{a_j}\right) = \sum\limits_{1 \leqslant i < j \leqslant n} \left|a_i \overline{b_j} - a_j \overline{b_i}\right|^2$

$|RHS| = \sum |a_i|^2 \cdot \sum |b_i|^2 - \sum a_i b_i \cdot \sum \overline{a_i b_i} = \sum |a_i|^2 \cdot \sum |b_i|^2 - \left|\sum\limits_{i=1}^{n} a_i b_i\right|^2$

证毕

Theorem 3.6.6 (复数形式 Cauchy-Schwarz 不等式)

a_i, b_i都是复数

$$\left(\sum_{i=1}^{n} |a_i|^2\right)\left(\sum_{i=1}^{n} |b_i|^2\right) \geqslant \left|\sum_{i=1}^{n} a_i b_i\right|^2 = \left|\sum_{i=1}^{n} a_i \overline{b_i}\right|^2$$

Proof 法一：由复数域上的$Largange$恒等式即可

法二：由实数形$Cauchy$不等式知道

$$\left(\sum_{i=1}^{n} |a_i|^2\right)\left(\sum_{i=1}^{n} |b_i|^2\right) \geqslant \left(\sum_{i=1}^{n} |a_i| \, |b_i|\right)^2 = \left(\sum_{i=1}^{n} \left|a_i \overline{b_i}\right|\right)^2 \geqslant \left|\sum_{i=1}^{n} a_i \overline{b_i}\right|^2 = \left|\sum_{i=1}^{n} a_i b_i\right|^2$$

法三：

$\sum\limits_{i=1}^{n} |a_i - \lambda b_i|^2 = \sum\limits_{i=1}^{n} \left(a_i - \lambda \overline{b_i}\right)\left(\overline{a_i} - \overline{\lambda} b_i\right) = \sum\limits_{i=1}^{n} \left[|a_i|^2 + |\lambda|^2 |b_i|^2 - 2\mathrm{Re}\left(\overline{\lambda} a_i b_i\right)\right]$

$= \sum\limits_{i=1}^{n} |a_i|^2 + |\lambda|^2 \sum\limits_{i=1}^{n} |b_i|^2 - 2\mathrm{Re}\left(\overline{\lambda} \sum\limits_{i=1}^{n} a_i b_i\right)$

令$\lambda = \dfrac{\sum\limits_{i=1}^{n} a_i b_i}{\sum\limits_{i=1}^{n} |b_i|^2}$ 则1. $|\lambda|^2 \sum\limits_{i=1}^{n} |b_i|^2 = \dfrac{\left|\sum\limits_{i=1}^{n} a_i b_i\right|^2}{\sum\limits_{i=1}^{n} |b_i|^2}$

$$2.\overline{\lambda}\sum_{i=1}^{n}a_ib_i=\frac{\sum\limits_{i=1}^{n}\overline{a_ib_i}}{\sum\limits_{i=1}^{n}|b_i|^2}\cdot\sum_{i=1}^{n}a_ib_i=\frac{\left|\sum\limits_{i=1}^{n}a_ib_i\right|^2}{\sum\limits_{i=1}^{n}|b_i|^2}\Rightarrow\mathrm{Re}\left(\overline{\lambda}\sum_{i=1}^{n}a_ib_i\right)=\frac{\left|\sum\limits_{i=1}^{n}a_ib_i\right|^2}{\sum\limits_{i=1}^{n}|b_i|^2}$$

$$\sum_{i=1}^{n}|a_i-\lambda b_i|^2=\sum_{i=1}^{n}|a_i|^2-\frac{\left|\sum\limits_{i=1}^{n}a_ib_i\right|^2}{\sum\limits_{i=1}^{n}|b_i|^2}\geqslant 0$$

证毕

3.7 矩阵常用公式结论

Proposition 3.6

1.逆阵唯一性

2.$\boldsymbol{A}$可逆则 $\left(\boldsymbol{A}^{-1}\right)^{-1}=\boldsymbol{A}$

3.$\boldsymbol{A},\boldsymbol{B}$均可逆则$\boldsymbol{AB}$可逆且 $(\boldsymbol{AB})^{-1}=\boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$

4.$\boldsymbol{A}$可逆, c为一非零数, 则$c\boldsymbol{A}$也可逆且 $(c\boldsymbol{A})^{-1}=c^{-1}\boldsymbol{A}^{-1}$

5.若$\boldsymbol{A}$是可逆阵, 则$\boldsymbol{A}'$也可逆且 $(\boldsymbol{A}')^{-1}=\left(\boldsymbol{A}^{-1}\right)'$

6.若A为对称阵那么A^{-1}也为对称阵

Theorem 3.7.1

设A,B为n阶矩阵那么 $(AB)^*=B^*A^*$

$(A')^*=(A^*)'$

$(cA)^*=c^{n-1}A^*$

若A可逆, 则A^*可逆, 且 $(A^*)^{-1}=\left(A^{-1}\right)^*$

若A为对称阵那么A^*也为对称阵

Proof 1.

$$(\boldsymbol{AB})^*(i;j)=(-1)^{i+j}\boldsymbol{AB}\begin{pmatrix}1,\cdots,j-1,j+1,\cdots,n\\1,\cdots,i-1,i+1,\cdots,n\end{pmatrix}$$
$$=(-1)^{i+j}\sum_{1\leqslant v_1<v_2<\cdots<v_{n-1}\leqslant n}\boldsymbol{A}\begin{pmatrix}1,\cdots,j-1,j+1,\cdots,n\\v_1,v_2,\cdots,v_{n-1}\end{pmatrix}\boldsymbol{B}\begin{pmatrix}v_1,v_2,\cdots,v_{n-1}\\1,\cdots,i-1,i+1,\cdots,n\end{pmatrix}$$
$$=(-1)^{i+j}\sum_{k=1}^{n}\boldsymbol{A}\begin{pmatrix}1,\cdots,j-1,j+1,\cdots,n\\1,\cdots,k-1,k+1,\cdots,n\end{pmatrix}\boldsymbol{B}\begin{pmatrix}1,\cdots,k-1,k+1,\cdots,n\\1,\cdots,i-1,i+1,\cdots,n\end{pmatrix}$$
$$=\sum_{k=1}^{n}(-1)^{j+k}\boldsymbol{A}\begin{pmatrix}1,\cdots,j-1,j+1,\cdots,n\\1,\cdots,k-1,k+1,\cdots,n\end{pmatrix}(-1)^{k+i}\boldsymbol{B}\begin{pmatrix}1,\cdots,k-1,k+1,\cdots,n\\1,\cdots,i-1,i+1,\cdots,n\end{pmatrix}$$
$$=\sum_{k=1}^{n}\boldsymbol{A}_{jk}\boldsymbol{B}_{ki}\ (\text{这里的}A_{jk},\ B_{ki}\text{表示代数余子式})$$
$$=\boldsymbol{B}^*\boldsymbol{A}^*(i;j),$$

因此 $(\boldsymbol{AB})^*=\boldsymbol{B}^*\boldsymbol{A}^*$

2.3.第二条第三条根据定义与行列式性质即可

4.$A^*\left(A^{-1}\right)^*=\left(A^{-1}A\right)^*=I_n^*=I_n$从而 $(A^*)^{-1}=\left(A^{-1}\right)^*$

5.同时取 (i,j) 元观察即可

Theorem 3.7.2

1. $\begin{bmatrix} I_r & O \\ O & O \end{bmatrix}^* = \begin{cases} O & r \leqslant n-2 \\ \begin{bmatrix} O & O \\ O & 1 \end{bmatrix} & r = n-1 \\ I_n & r = n \end{cases}$ $\quad \begin{bmatrix} I_r & O \\ O & O \end{bmatrix}^{**} = \begin{cases} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} & n = 2; r = 1 \\ I_n & r = n \\ O & \text{其他} \end{cases}$

2. $|A^*| = |A|^{n-1}$

3. $(A^*)^* = \begin{cases} |A|^{n-2}\, A \ (n \geqslant 3) \\ A \qquad\quad n = 2 \end{cases}$

4.设$\boldsymbol{A}$为m阶矩阵,$\boldsymbol{B}$为n阶矩阵,求分块对角阵$\boldsymbol{C}$的伴随矩阵 : $C = \begin{pmatrix} A & O \\ O & B \end{pmatrix}$.

Proof 1.利用定义即可

Proof 法一 : 若A可逆, 即$A^{-1} = \frac{1}{|A|} A^* \Rightarrow A^* = |A|\, A^{-1} \Rightarrow |A^*| = |A|^n \left|A^{-1}\right| = |A|^n\, |A|^{-1}$
$\Rightarrow |A^*| = |A|^{n-1}$
若A不可逆, 即 $|A| = 0$即证 $|A^*| = 0$
法一 : 反证法假设 $|A^*| \neq 0$; 则A^*可逆
$AA^* = |A|\, I_n = O \Rightarrow A = O$利用伴随阵定义零矩阵的伴随阵为零矩阵即$A^* = O$这与 $|A^*| \neq 0$矛盾
$\Rightarrow |A^*| = 0$

法二 : 若$\boldsymbol{A}$不可逆, 即$|\boldsymbol{A}| = 0$, 则存在可逆矩阵$\boldsymbol{P}, \boldsymbol{Q}$, 使$\boldsymbol{PAQ} = \Lambda = \begin{pmatrix} \boldsymbol{I}_r & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} \end{pmatrix}$, 其中$r < n$.

我们注意到 : 若$r \leq n-2$, 则$\Lambda^* = \boldsymbol{O}$; 若$r = n-1$, 则$\Lambda^* = \begin{pmatrix} \boldsymbol{O} & \boldsymbol{O} \\ \boldsymbol{O} & 1 \end{pmatrix}$.无论是哪种情况, 我们都有 $|\Lambda^*| = 0$

从而$0 = |(\boldsymbol{PAQ})^*| = |\boldsymbol{Q}^* \boldsymbol{A}^* \boldsymbol{P}^*| = |\boldsymbol{Q}^*|\, |\boldsymbol{A}^*|\, |\boldsymbol{P}^*|$.由题知$\boldsymbol{P}^*, \boldsymbol{Q}^*$都是可逆矩阵, 因此 $|\boldsymbol{A}^*| = 0 = |\boldsymbol{A}|^{n-1}$仍然成立.

法三 : 若$\boldsymbol{A} = 0$, 则结论显然成立
下设$\boldsymbol{A} \neq 0$, 我们知道, $\boldsymbol{AA}^* = |\boldsymbol{A}|\boldsymbol{I}$.
若$|\boldsymbol{A}| \neq 0$, 则$|\boldsymbol{A}|\, |\boldsymbol{A}^*| = |\boldsymbol{A}|^n$, 从而 $|\boldsymbol{A}^*| = |\boldsymbol{A}|^{n-1}$
若$|\boldsymbol{A}| = 0$, 则$\boldsymbol{AA}^* = 0$根据$slverster$不等式从而 $\operatorname{rank}(\boldsymbol{A}) + \operatorname{rank}(\boldsymbol{A}^*) \leqslant n$,
$\operatorname{rank}(\boldsymbol{A}^*) \leqslant n - \operatorname{rank}(\boldsymbol{A}) < n$,
因此 $|\boldsymbol{A}^*| = 0$, 从而结论成立。

Proof 法一: 当$n = 2$直接计算即可得到, 当$n = 3$时
若A可逆, 有$AA^* = |A|\, \boldsymbol{I}_n \Rightarrow A^*\,(A^*)^* = |A^*|\, I_n = |A|^{n-1}\, I_n$又$A^* = |A|\, A^{-1}$
$|A|\, A^{-1}\,(A^*)^* = |A|^{n-1}\, I_n \Rightarrow (A^*)^* = \left|A^{n-2}\right| A$

若$\boldsymbol{A}$不可逆, 即$|\boldsymbol{A}| = 0$, 则存在可逆矩阵$\boldsymbol{P}, \boldsymbol{Q}$, 使$\boldsymbol{PAQ} = \Lambda = \begin{pmatrix} \boldsymbol{I}_r & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} \end{pmatrix}$, 其中$r < n$.由1.可得$\Lambda^{**} = \boldsymbol{O}$, 从而
$\boldsymbol{O} = (\boldsymbol{PAQ})^{**} = P^{**} A^{**} Q^{**}$.知$\boldsymbol{P}^{**}, \boldsymbol{Q}^{**}$都是可逆矩阵, 因此$\boldsymbol{A}^{**} = \boldsymbol{O} = |\boldsymbol{A}|^{n-2}\boldsymbol{A}$仍然成立.

法二 :
当$n = 2$, 直接计算即可

设$n \geqslant 3$.

若$|\boldsymbol{A}| \neq 0$, 则 $|\boldsymbol{A}^*| = |\boldsymbol{A}|^{n-1}$. 由于$\boldsymbol{A}^*(\boldsymbol{A}^*)^* = |\boldsymbol{A}^*|\boldsymbol{I}$, 因此 $(\boldsymbol{A}^*)^* = |\boldsymbol{A}^*|(\boldsymbol{A}^*)^{-1} = |\boldsymbol{A}|^{n-1}\dfrac{1}{|\boldsymbol{A}|}\boldsymbol{A} = |\boldsymbol{A}|^{n-2}\boldsymbol{A}$

若$|\boldsymbol{A}| = 0$, 则据2.的结果得 $\operatorname{rank}(\boldsymbol{A}^*) \leqslant 1 < n-1$, 因此 $(\boldsymbol{A}^*)^* = 0$, 于是结论也成立。

Proof 4.设$\boldsymbol{A} = (a_{ij})_{m\times m}$, 元素$a_{ij}$的余子式和代数余子式分别记为$M_{ij}$和$A_{ij}$;

$\boldsymbol{B} = (b_{ij})_{n\times n}$, 元素$b_{ij}$的余子式和代数余子式分别记为$N_{ij}$和$B_{ij}$.

利用$Laplace$定理可以容易地计算出：当$1 \leq i,j \leq m$时, C的第(i,j)元素的代数余子式为$(-1)^{i+j}M_{ij}|\boldsymbol{B}| = |\boldsymbol{B}|A_{ij}$;

当$m+1 \leq i,j \leq m+n$时, $\boldsymbol{C}$的第(i,j)元素的代数余子式为$(-1)^{i+j}N_{i-m,j-m}|\boldsymbol{A}| = |\boldsymbol{A}|B_{i-m,j-m}$;

当i,j属于其他范围时, $\boldsymbol{C}$的第(i,j)元素的代数余子式等于零.因此我们有$C^* = \begin{pmatrix} |B|A^* & O \\ O & |\boldsymbol{A}|\boldsymbol{B}^* \end{pmatrix}$

Lemma 3.4 (Push-through 公式)

$A_{m\times n}; B_{n\times m}$使得$I_m + AB$可逆 $\Rightarrow I_n + BA$也可逆

Proof 注意到$A(I_n+BA) = (I_m+AB)A \Rightarrow \ (I_m+AB)^{-1}A(I_n+BA) = A$

$\Rightarrow B(I_m+AB)^{-1}A(I_n+BA) = BA$

那么$I_n = I_n + BA - BA = (I_n+BA) - B(I_m+AB)^{-1}A(I_n+BA)$

$\Rightarrow I_n = \left[I_n - B(I_m+AB)^{-1}A\right](I_n+BA)$

$(I_n+BA)^{-1} = I_n - B(I_m+AB)^{-1}A$

Proof 设$\boldsymbol{C}$为$\boldsymbol{E}+\boldsymbol{AB}$的逆矩阵, 则有$(\boldsymbol{E}+\boldsymbol{AB})\boldsymbol{C} = \boldsymbol{E}$, 即$\boldsymbol{C}+\boldsymbol{ABC} = \boldsymbol{E}$

$\Longrightarrow$ 左乘$\boldsymbol{B}$, 右乘$\boldsymbol{A}$可得$\boldsymbol{BCA}+\boldsymbol{BABCA} = \boldsymbol{BA}$, 即$(\boldsymbol{E}+\boldsymbol{BA})\boldsymbol{BCA} = \boldsymbol{BA}$

$\Longrightarrow$ 从而$(\boldsymbol{E}+\boldsymbol{BA})\boldsymbol{BCA} - (\boldsymbol{E}+\boldsymbol{BA}) = -\boldsymbol{E}$, 故$(\boldsymbol{E}+\boldsymbol{BA})(\boldsymbol{E}-\boldsymbol{BCA}) = \boldsymbol{E}$

于是$\boldsymbol{E}+\boldsymbol{BA}$可逆, 且$(\boldsymbol{E}+\boldsymbol{BA})^{-1} = \boldsymbol{E}-\boldsymbol{BCA} = \boldsymbol{E}-\boldsymbol{B}(\boldsymbol{E}+\boldsymbol{AB})^{-1}\boldsymbol{A}$

Proof 考虑分块矩阵$\begin{pmatrix} E & A \\ -B & E \end{pmatrix}$.由于

$$\begin{pmatrix} E & O \\ B & E \end{pmatrix}\begin{pmatrix} E & A \\ -B & E \end{pmatrix}\begin{pmatrix} E & -A \\ O & E \end{pmatrix} = \begin{pmatrix} E & O \\ O & E+BA \end{pmatrix}$$

$$\begin{pmatrix} E & -A \\ O & E \end{pmatrix}\begin{pmatrix} E & A \\ -B & E \end{pmatrix}\begin{pmatrix} E & O \\ B & E \end{pmatrix} = \begin{pmatrix} E+AB & O \\ O & E \end{pmatrix}$$

上两式两边取行列式可得$|\boldsymbol{E}+\boldsymbol{AB}| = |\boldsymbol{E}+\boldsymbol{BA}|$.故如果$\boldsymbol{E}+\boldsymbol{AB}$可逆, 则$\boldsymbol{E}+\boldsymbol{BA}$可逆.由于

$$\begin{pmatrix} \boldsymbol{E} & \boldsymbol{A} \\ -\boldsymbol{B} & \boldsymbol{E} \end{pmatrix}^{-1} = \begin{pmatrix} \boldsymbol{E} & -\boldsymbol{A} \\ \boldsymbol{O} & \boldsymbol{E} \end{pmatrix}\begin{pmatrix} \boldsymbol{E}^{-1} & \boldsymbol{O} \\ \boldsymbol{O} & (\boldsymbol{E}+\boldsymbol{BA})^{-1} \end{pmatrix}\begin{pmatrix} \boldsymbol{E} & \boldsymbol{O} \\ \boldsymbol{B} & \boldsymbol{E} \end{pmatrix} = \begin{pmatrix} \boldsymbol{E}-\boldsymbol{A}(\boldsymbol{E}+\boldsymbol{BA})^{-1}\boldsymbol{B} & -\boldsymbol{A}(\boldsymbol{E}+\boldsymbol{BA})^{-1} \\ (\boldsymbol{E}+\boldsymbol{BA})^{-1}\boldsymbol{B} & (\boldsymbol{E}+\boldsymbol{BA})^{-1} \end{pmatrix}$$

$$\begin{pmatrix} \boldsymbol{E} & \boldsymbol{A} \\ -\boldsymbol{B} & \boldsymbol{E} \end{pmatrix}^{-1} = \begin{pmatrix} \boldsymbol{E} & \boldsymbol{O} \\ \boldsymbol{B} & \boldsymbol{E} \end{pmatrix}\begin{pmatrix} (\boldsymbol{E}+\boldsymbol{AB})^{-1} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{E}^{-1} \end{pmatrix}\begin{pmatrix} \boldsymbol{E} & -\boldsymbol{A} \\ \boldsymbol{O} & \boldsymbol{E} \end{pmatrix} = \begin{pmatrix} (\boldsymbol{E}+\boldsymbol{AB})^{-1} & -(\boldsymbol{E}+\boldsymbol{AB})^{-1}\boldsymbol{A} \\ \boldsymbol{B}(\boldsymbol{E}+\boldsymbol{AB})^{-1} & \boldsymbol{E}-\boldsymbol{B}(\boldsymbol{E}+\boldsymbol{AB})^{-1}\boldsymbol{A} \end{pmatrix}$$

Theorem 3.7.3 (Sherman-Morrison-Woodbury 定理)

设A为n阶可逆阵, C为m阶可逆阵, $B_{n\times m}; D_{m\times n}$矩阵, 使得$C^{-1}+DA^{-1}B$可逆

则$A+BCD$可逆; $(A+BCD)^{-1} = A^{-1} - A^{-1}B\left(C^{-1}+DA^{-1}B\right)^{-1}DA^{-1}$

Proof $A+BCD = A\left(I_n + A^{-1}BCD\right)$ 那么我们想要知道$I_n + A^{-1}BCD$是否可逆

此时$I_m + (CD)\left(A^{-1}B\right) = C\left(C^{-1}+DA^{-1}B\right)$ 所以$I_m + (CD)\left(A^{-1}B\right)$ 可逆

且有$\left(I_m + CDA^{-1}B\right)^{-1} = \left(C^{-1}+DA^{-1}B\right)^{-1}C^{-1}$

那么由$Push-through$等式知道$I_n+\left(A^{-1}B\right)(CD)$也可逆

$\left(I_n+A^{-1}BCD\right)^{-1}=I_n-A^{-1}B\left(I_m+CDA^{-1}B\right)^{-1}CD=I_n-A^{-1}B\left(C^{-1}+DA^{-1}B\right)^{-1}D$

于是$A+BCD$可逆

且$(A+BCD)^{-1}=A^{-1}-A^{-1}B\left(C^{-1}+DA^{-1}B\right)^{-1}DA^{-1}$

Theorem 3.7.4 (Sherman-Morrison 公式)

设A为n阶可逆阵,α,β是n维列向量且$1+\beta'A^{-1}\alpha\neq 0$

$\Rightarrow (A+\alpha\beta')^{-1}=A^{-1}-\dfrac{1}{1+\beta'A^{-1}\alpha}A^{-1}\alpha\beta'A^{-1}$

Proof 这是上述定理的一个特殊形式

Theorem 3.7.5

设$\boldsymbol{A},\boldsymbol{B}$是$n$阶矩阵

(1)$\operatorname{tr}(\boldsymbol{A}+\boldsymbol{B})=\operatorname{tr}\boldsymbol{A}+\operatorname{tr}\boldsymbol{B}$; (2)$\operatorname{tr}(k\boldsymbol{A})=k(\operatorname{tr}\boldsymbol{A})$;

(3)$\operatorname{tr}\left(\boldsymbol{A}'\right)=\operatorname{tr}\boldsymbol{A}$; (4)$\operatorname{tr}(\boldsymbol{A}\boldsymbol{B})=\operatorname{tr}(\boldsymbol{B}\boldsymbol{A})$ (更进一步$A_{m\times n};B_{n\times m}$也一样)

Proof (1)、(2)以及(3)显然可得,只须证明(4)即可.

设$\boldsymbol{A}=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$, $\boldsymbol{B}=\begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{pmatrix}$,

设$\boldsymbol{C}=\boldsymbol{A}\boldsymbol{B},\boldsymbol{D}=\boldsymbol{B}\boldsymbol{A},\boldsymbol{C}=\left(c_{ij}\right),\boldsymbol{D}=\left(d_{ij}\right)$,则 $\operatorname{tr}(\boldsymbol{A}\boldsymbol{B})=\sum_{i=1}^n c_{ii}=\sum_{i=1}^n\sum_{k=1}^n a_{ik}b_{ki},\operatorname{tr}(\boldsymbol{B}\boldsymbol{A})=\sum_{i=1}^n d_{ii}=\sum_{i=1}^n\sum_{k=1}^n b_{ik}a_{ki}$.

由求和的交换性即得 : $\operatorname{tr}(\boldsymbol{A}\boldsymbol{B})=\operatorname{tr}(\boldsymbol{B}\boldsymbol{A})$.

3.8 特殊矩阵

Definition 3.7 (标准单位向量)

n维标准单位列向量是指下列n个n维列向量：$\boldsymbol{e}_1=\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix},\boldsymbol{e}_2=\begin{pmatrix}0\\1\\\vdots\\0\end{pmatrix},\cdots,\boldsymbol{e}_n=\begin{pmatrix}0\\0\\\vdots\\1\end{pmatrix}$.

向量组$\boldsymbol{e}_1',\boldsymbol{e}_2',\cdots,\boldsymbol{e}_n'$则被称为$n$维标准单位行向量.标准单位向量有下列基本性质

Proposition 3.7

(1)若$i\neq j$,则$\boldsymbol{e}_i'\boldsymbol{e}_j=0$,而$\boldsymbol{e}_i'\boldsymbol{e}_i=1$;

(2)若$\boldsymbol{A}=(a_{ij})$是$m\times n$矩阵,则$\boldsymbol{A}\boldsymbol{e}_i$是$\boldsymbol{A}$的第$i$个列向量;$\boldsymbol{e}_i'\boldsymbol{A}$是$\boldsymbol{A}$的第$i$个行向量;

(3)若$\boldsymbol{A}=(a_{ij})$是$m\times n$矩阵,则$\boldsymbol{e}_i'\boldsymbol{A}\boldsymbol{e}_j=a_{ij}$.

(4) 设$\boldsymbol{A},\boldsymbol{B}$都为$m\times n$矩阵则$\boldsymbol{A}=\boldsymbol{B}\Leftrightarrow Ae_i=Be_i\ (\text{或}e_i'A=e_i'B)$

Definition 3.8 (初级基础矩阵)

n阶基础矩阵(又称初级矩阵)是指n^2个n阶矩阵$\{\boldsymbol{E}_{ij}(i,j=1,2,\cdots,n)\}$.

这里$\boldsymbol{E}_{ij}$是一个n阶矩阵,它的第(i,j)元素等于1,其他元素全为0.基础矩阵也可以看成是标准单位向量的积：$\boldsymbol{E}_{ij}=\boldsymbol{e}_i\boldsymbol{e}_j'$

由此我们不难证明基础矩阵的下列性质：

Proposition 3.8

(1) $\boldsymbol{E}_{ij}\boldsymbol{E}_{kl}=\begin{cases}\boldsymbol{O} & j\neq k\\ \boldsymbol{E}_{il} & j=k\end{cases}$

(2)若$\boldsymbol{A}$是n阶矩阵且$\boldsymbol{A}=(a_{ij})$,则$\boldsymbol{A}=\sum\limits_{i,j=1}^{n}a_{ij}\boldsymbol{E}_{ij}$

(3)若$\boldsymbol{A}$是n阶矩阵且$\boldsymbol{A}=(a_{ij})$,则$\boldsymbol{E}_{ij}\boldsymbol{A}$将$\boldsymbol{A}$的第$j$行变为第$i$行,将其他元素全变为0;

(4) 若$\boldsymbol{A}$是n阶矩阵且$\boldsymbol{A}=(a_{ij})$,则$\boldsymbol{A}\boldsymbol{E}_{ij}$将$\boldsymbol{A}$的第$i$列变为第$j$列,将其他元素全变为0;

(5) $\boldsymbol{E}_{ij}\boldsymbol{A}\boldsymbol{E}_{kl}=a_{jk}\boldsymbol{E}_{il}$

Definition 3.9 (循环移位矩阵)

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

Proposition 3.9

(1)用A左乘一个矩阵,就相当于把这个矩阵的行向上移一行,第1行换到最后一行;
用A右乘一个矩阵,就相当于把这个矩阵的列向右移一列,最后一列换到第1列;

(2) $A^k = \begin{pmatrix} O & I_{n-k} \\ I_k & O \end{pmatrix} (k = 1, 2, \cdots, n)$.

(3) $\sum_{l=0}^{n-1} A^l = J$,其中J是元素全为1的n级矩阵

Proof 将A写为$A = (e_n, e_1, \cdots, e_{n-1})$,其中$e_i$是单位列向量.由分块矩阵乘法并注意$Ae_i$就是$A$的第$i$列,因此$A^2 = (Ae_n, Ae_1, Ae_2, \cdots, Ae_{n-1}) = (e_{n-1}, e_n, e_1, \cdots, e_{n-2})$不断这样做下去就可得到结论.

Definition 3.10 (循环矩阵)

设b为非零常数,下列形状的矩阵称为$b-$循环矩阵:$A = \begin{pmatrix} a_1 & a_2 & a_3 & \cdots & a_n \\ ba_n & a_1 & a_2 & \cdots & a_{n-1} \\ ba_{n-1} & ba_n & a_1 & \cdots & a_{n-2} \\ \vdots & \vdots & \vdots & & \vdots \\ ba_2 & ba_3 & ba_4 & \cdots & a_1 \end{pmatrix}$.

Proposition 3.10

(1)证明:同阶$b-$循环矩阵的乘积仍然是$b-$循环矩阵

(2)求上述b循环矩阵A的行列式的值.

Proof 设$b-$循环矩阵为$J_b = \begin{pmatrix} O & I_{n-1} \\ b & O \end{pmatrix} = (be_n, e_1, e_2 \cdots e_{n-1})$

那么$J_b^2 = J_b (be_n, e_1, e_2 \cdots e_{n-1}) = (be_{n-1}, be_n, e_1, e_2 \cdots e_{n-1})$以此类推得到$J_b^k = \begin{pmatrix} O & I_{n-k} \\ bI_k & O \end{pmatrix} (k = 1, \cdots n)$

$A = a_1 I_n + a_2 J_b + a_3 J_b^2 + \cdots + a_n J_b^{n-1}$循环矩阵具有该形式，具有该形式的矩阵就为$b-$循环矩阵

那么两个$b-$循环矩阵的乘积注意到$J_b^n = bI_n$就仍然是$b-$循环矩阵

(2)作多项式$f(x) = a_1 + a_2 x + a_3 x^2 + \cdots + a_n x^{n-1}$,令$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$是$b$的所有$n$次方根.

完全类似于第一章行列式的解法,最后可得$|A| = f(\varepsilon_1) f(\varepsilon_2) \cdots f(\varepsilon_n)$.

Proposition 3.11

求下列循环矩阵的特征值：$A=\begin{pmatrix} a_1 & a_2 & a_3 & \cdots & a_n \\ a_n & a_1 & a_2 & \cdots & a_{n-1} \\ a_{n-1} & a_n & a_1 & \cdots & a_{n-2} \\ \vdots & \vdots & \vdots & & \vdots \\ a_2 & a_3 & a_4 & \cdots & a_1 \end{pmatrix}$

设$\boldsymbol{J}=\begin{pmatrix} \boldsymbol{O} & \boldsymbol{I}_{n-1} \\ 1 & \boldsymbol{O} \end{pmatrix}$, $f(x)=a_1+a_2x+a_3x^2+\cdots+a_nx^{n-1}$，则可知$\boldsymbol{A}=f(\boldsymbol{J})$.

经简单计算可得$|\lambda \boldsymbol{I}_n-\boldsymbol{J}|=\lambda^n-1$，于是$\boldsymbol{J}$的特征值为$\omega_k=\cos\dfrac{2k\pi}{n}+\mathrm{i}\sin\dfrac{2k\pi}{n}$，$\quad 0\le k\le n-1$

因此$\boldsymbol{A}$的特征值为$f(1), f(\omega_1), \cdots, f(\omega_{n-1})$.

Proposition 3.12

求下列循环矩阵的逆矩阵 $(n\geqslant 2)$: $\boldsymbol{A}=\begin{pmatrix} 1 & 2 & 3 & \cdots & n \\ n & 1 & 2 & \cdots & n-1 \\ \vdots & \vdots & \vdots & & \vdots \\ 2 & 3 & 4 & \cdots & 1 \end{pmatrix}$

Proof 先解线性方程组$\boldsymbol{Ax}=\boldsymbol{\beta}$, 其中$\boldsymbol{\beta}=(b_1,b_2,\cdots,b_n)'$

将这n个方程相加, 得$\dfrac{1}{2}n(n+1)(x_1+x_2+\cdots+x_n)=\sum_{j=1}^{n}b_j$

令$y=x_1+x_2+\cdots+x_n$, 由上式得$y=\dfrac{2}{n(n+1)}\sum_{j=1}^{n}b_j$

从第1个方程减去第2个方程, 得

$(1-n)x_1+x_2+x_3+\cdots+x_n=b_1-b_2$, 由此得出$y-nx_1=b_1-b_2$, 从而

$$x_1=\frac{1}{n}(y-b_1+b_2)=\frac{1}{n}\left[\left(\frac{2}{n(n+1)}\sum_{j=1}^{n}b_j\right)-b_1+b_2\right]$$

类似地, 从第i个方程减去第$i+1$个方程 $(i=2,\cdots,n-1)$, 可求出

$$x_i=\frac{1}{n}\left[\left(\frac{2}{n(n+1)}\sum_{j=1}^{n}b_j\right)-b_i+b_{i+1}\right],\quad i=2,\cdots,n-1$$

从第n个方程减去第1个方程, 可求出$x_n=\dfrac{1}{n}\left[\left(\dfrac{2}{n(n+1)}\sum_{j=1}^{n}b_j\right)-b_n+b_1\right]$

记$s=\dfrac{2}{n(n+1)}$

分别令$\boldsymbol{\beta}$为$\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_n$, 得

$$\boldsymbol{A}^{-1}=\frac{1}{n}\begin{pmatrix} s-1 & s+1 & s & \cdots & s \\ s & s-1 & s+1 & \cdots & s \\ s & s & s-1 & \cdots & s \\ \vdots & \vdots & \vdots & & \vdots \\ s & s & s & \cdots & s+1 \\ s+1 & s & s & \cdots & s-1 \end{pmatrix}。$$

Definition 3.11 (幂零 Jodan 块)

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

Proposition 3.13

求证：$\boldsymbol{A}^k=\begin{pmatrix}\boldsymbol{O}&\boldsymbol{I}_{n-k}\\\boldsymbol{O}&\boldsymbol{O}\end{pmatrix}(k=1,2,\cdots,n)$.

Proof 同理上文即可

Definition 3.12

$f(x)=x^n+a_1x^{n-1}+\cdots+a_{n-1}x+a_n$

定义$f(x)$友阵$C(f(x))=\begin{pmatrix}0&0&\cdots&0&-a_n\\1&0&\cdots&0&-a_{n-1}\\0&1&\cdots&0&-a_{n-2}\\\vdots&\vdots&&\vdots&\vdots\\0&0&\cdots&1&-a_1\end{pmatrix}$ 定义$F(f(x))=C(f(x))'$称为$f(x)$的$Frobenius$块

Proposition 3.14

$|xI_n-C(f(x))|=f(x)$

$C(f(x))e_i=e_{i+1}\ (1\leqslant i\leqslant n-1)$

$C(f(x))e_n=-\sum_{i=1}^n a_{n-i+1}e_i$

Proof 我们只需要证明$D_n=\begin{vmatrix}x&0&0&\cdots&0&a_0\\-1&x&0&\cdots&0&a_1\\0&-1&x&\cdots&0&a_2\\\vdots&\vdots&\vdots&&\vdots&\vdots\\0&0&0&\cdots&x&a_{n-2}\\0&0&0&\cdots&-1&x+a_{n-1}\end{vmatrix}=x^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$

从第n行起各行的x倍依次加到上面一行

$$LHS=\begin{vmatrix}0&0&0&\cdots&0&x^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0\\-1&0&0&\cdots&0&x^{n-1}+a_{n-1}x^{n-2}+\cdots+a_2x+a_1\\0&-1&0&\cdots&0&x^{n-2}+a_{n-1}x^{n-3}+\cdots+a_3x+a_2\\\vdots&\vdots&\vdots&&\vdots&\vdots\\0&0&0&\cdots&0&x^2+a_{n-1}x+a_{n-2}\\0&0&0&\cdots&-1&x+a_{n-1}\end{vmatrix}$$

$=(\text{按第一行展开})=\left(x^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0\right)*(-1)^{1+n}*(-1)^{n-1}$

$=x^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$

Definition 3.13 (幂等矩阵)

幂等矩阵 : $A_{n\times n}$; $A^2=A$

Proposition 3.15

(1) : A为幂等矩阵 $\Leftrightarrow r(A)+r(I_n-A)=n$

(2) : $r(A)=r$; A为幂等矩阵 $\Leftrightarrow$ 存在$S_{n\times r}$列满秩矩阵和$T_{r\times n}$行满秩矩阵 $st.A=ST$; $TS=I_r$

Proof (1) : $\begin{pmatrix} A & O \\ O & I-A \end{pmatrix} \to \begin{pmatrix} A & A \\ O & I-A \end{pmatrix} \to \begin{pmatrix} A & A \\ A & I \end{pmatrix} \to \begin{pmatrix} A-A^2 & A \\ O & I \end{pmatrix} \to \begin{pmatrix} A-A^2 & O \\ O & I \end{pmatrix}$

因此$r\begin{pmatrix} A & O \\ O & I-A \end{pmatrix} = r\begin{pmatrix} A-A^2 & O \\ O & I \end{pmatrix} \Rightarrow r(A)+r(I-A)=r(A-A^2)+n$

(2) : 充分性显然, 下证必要性

设P,Q是n阶非异阵, 使得$A=P\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}Q$; 由$A^2=A$得到$\begin{pmatrix} I_r & O \\ O & O \end{pmatrix} = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}QP\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$

令$S=P\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}\begin{pmatrix} I_r \\ O \end{pmatrix}$ $T=\begin{pmatrix} I_r & O \end{pmatrix}\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}Q$即可

Definition 3.14

对合矩阵 : $A_{n\times n}$; $A^2=I_n$

Proposition 3.16

(1) : A为对合矩阵 $\Leftrightarrow r(I_n-A)+r(I_n+A)=n$

Proof (1) : $\begin{pmatrix} I_n-A & O \\ O & I_n+A \end{pmatrix} \to \begin{pmatrix} I_n-A & I_n-A \\ O & I_n+A \end{pmatrix} \to \begin{pmatrix} I_n-A & I_n-A \\ I_n-A & 2I_n \end{pmatrix} \to \begin{pmatrix} \frac{1}{2}(I_n-A^2) & I_n-A \\ O & 2I_n \end{pmatrix} \to \begin{pmatrix} \frac{1}{2}(I_n-A^2) & O \\ O & I_n \end{pmatrix}$

所以$r\begin{pmatrix} I_n-A & O \\ O & I_n+A \end{pmatrix} = r\begin{pmatrix} \frac{1}{2}(I_n-A)^2 & O \\ O & I_n \end{pmatrix} \Rightarrow r(I_n-A)+r(I_n+A)=n+r(I_n-A^2)$

Definition 3.15 (对称与反对称矩阵)

(1) 一个矩阵$\boldsymbol{A}$如果满足$\boldsymbol{A}'=\boldsymbol{A}$, 那么称$\boldsymbol{A}$是对称矩阵

1.对称矩阵一定是方阵

2.n级矩阵$\boldsymbol{A}$是对称矩阵当且仅当$\boldsymbol{A}(i;j)=\boldsymbol{A}(j;i)$, $i,j=1,2,\cdots,n$

3.设$\boldsymbol{A},\boldsymbol{B}$都是数域$K$上的$n$级对称矩阵, 则$\boldsymbol{A}+\boldsymbol{B}, k\boldsymbol{A}(k\in K)$都是对称矩阵

4.设$\boldsymbol{A},\boldsymbol{B}$都是$n$级对称矩阵, 则$\boldsymbol{AB}$为对称矩阵的充分必要条件是$\boldsymbol{A}$与$\boldsymbol{B}$可交换

(2) 一个矩阵$\boldsymbol{A}$如果满足$\boldsymbol{A}'=-\boldsymbol{A}$, 那么称$\boldsymbol{A}$是斜对称矩阵

1.斜对称矩阵一定是方阵

2.并且n级矩阵$\boldsymbol{A}$是斜对称矩阵当且仅当对于$i,j=1,2,\cdots,n$, 有$\boldsymbol{A}(i;j)=-\boldsymbol{A}(j;i), \boldsymbol{A}(i;i)=0$.

3.数域K上奇数级斜对称矩阵的行列式等于0

4.若$\boldsymbol{A}$与$\boldsymbol{B}$都是数域K上的n级斜对称矩阵, 则$\boldsymbol{A}+\boldsymbol{B}, k\boldsymbol{A}(k\in K)$也都是斜对称矩阵

Definition 3.16 (幂零矩阵)

方阵$\boldsymbol{A}$称为幂零矩阵, 如果存在正整数l, 使得$\boldsymbol{A}^l = 0$; 使$\boldsymbol{A}^l = 0$成立的最小正整数l称为$\boldsymbol{A}$的幂零指数

Proposition 3.17

(1)上 (下) 三角矩阵是幂零矩阵当且仅当它的主对角元全为0

(2)如果n级上 (下) 三角矩阵是幂零矩阵, 那么它的幂零指数$l \leqslant n$。

Proof 实际上我们可以证明如下命题：

设$\boldsymbol{A}$是n阶上三角矩阵且主对角线上元素全为零, 那么$\boldsymbol{A}^n = \boldsymbol{O}$.

$A = \left(0, k_{21}e_1, \cdots, k_{n1}e_1 + \cdots + k_{n,n-1}e_{n-1}\right)$

归纳假设证明$A^k e_k = 0\,(1 \leqslant k \leqslant n)$.那么就有$A^n e_i = 0\,(1 \leqslant i \leqslant n)$ 那么$A^n = O$

显然$Ae_1 = 0$成立, 假设$A^i e_i = 0$对当$i < k$时成立.那么当$i = k$时

$A^k e^k = A^{k-1}\left(Ae_k\right) = A^{k-1}\left(a_{1k}e_1 + a_{2k}e_2 + \cdots + a_{k-1,k}e_{k-1}\right) = 0$

所以 (2) 中幂零指数自然小于等于n

(1) 问中的充分性也证明完毕了.

(1) 问中的必要性：因为上三角矩阵的乘积仍然为上三角矩阵, 且主对角元是对应的操作,

那么$A^n(i,i) = (A(i,i))^n = 0 \Longrightarrow A(i,i) = 0$故主对角线全为0

Proposition 3.18

1.n阶矩阵$\boldsymbol{A}$是幂零矩阵的充要条件是A的所有特征值$= 0$.

2.求证：n阶矩阵$\boldsymbol{A}$是幂零矩阵的充要条件是 $\operatorname{tr}\left(\boldsymbol{A}^k\right) = 0(1 \leq k \leq n)$.

Proof 1.主要注意到ch定理

2.设$\boldsymbol{A}$的特征值为$\lambda_1, \lambda_2, \cdots, \lambda_n$.若$\boldsymbol{A}$是幂零矩阵, 则$\boldsymbol{A}$的特征值全为零, 从而 $\operatorname{tr}\left(\boldsymbol{A}^k\right) = s_k = 0(k \geq 1)$

若$s_k = \operatorname{tr}\left(\boldsymbol{A}^k\right) = 0(1 \leq k \leq n)$, 则由$Newton$公式可计算出$\sigma_r = 0(1 \leq r \leq n)$

由特征多项式与主子式的关系知道$\boldsymbol{A}$的特征多项式为λ^n, 从而$\boldsymbol{A}$的特征值全为零, 可知$\boldsymbol{A}$为幂零矩阵.

3.9 摄动法

Theorem 3.9.1

设$\boldsymbol{A}$是一个n阶方阵，求证：存在一个正数a，使得对任意的$0<t<a$，矩阵$t\boldsymbol{I}_n+\boldsymbol{A}$都是非异阵.

Proof 通过简单的行列式计算，可得$|t\boldsymbol{I}_n+\boldsymbol{A}|=t^n+a_1t^{n-1}+\cdots+a_{n-1}t+a_n$，这是一个关于未知数$t$的$n$次多项式.
可知上述多项式至多只有n个不同的根.若上述多项式的根都是零，则不妨取$a=1$;
若上述多项式有非零根，则令a为 $|t\boldsymbol{I}_n+\boldsymbol{A}|$ 所有非零根的模长的最小值.
因此对任意的$0<t_0<a$，t_0不是 $|t\boldsymbol{I}_n+\boldsymbol{A}|$ 的根，即 $|t_0\boldsymbol{I}_n+\boldsymbol{A}|\neq 0$，从而$t_0\boldsymbol{I}_n+\boldsymbol{A}$是非异阵.

摄动法原理

1. 证明矩阵问题对非异阵成立
2. 对任意的 n 阶矩阵 $\boldsymbol{A}$，存在一列有理数 $t_k\to 0$ 使得 t_kI_n+A 都是非异阵，验证上述同样满足题干中的条件从而该问题对 t_kI_n+A 也成立
3. 该问题关于 t_k 连续，则可取极限从而得到一般结论

Proposition 3.19

$(\boldsymbol{AB})^*=\boldsymbol{B}^*\boldsymbol{A}^*$

Proof 若$\boldsymbol{A},\boldsymbol{B}$均为非异阵，则$\boldsymbol{A}^*=|\boldsymbol{A}|\boldsymbol{A}^{-1}$，$\boldsymbol{B}^*=|\boldsymbol{B}|\boldsymbol{B}^{-1}$，从而 $(\boldsymbol{AB})^*=|\boldsymbol{AB}|(\boldsymbol{AB})^{-1}=|\boldsymbol{A}||\boldsymbol{B}|\left(\boldsymbol{B}^{-1}\boldsymbol{A}^{-1}\right)=\left(|\boldsymbol{B}|\boldsymbol{B}^{-1}\right)\left(|\boldsymbol{A}|\boldsymbol{A}^{-1}\right)=\boldsymbol{B}^*\boldsymbol{A}^*$.
对于一般的方阵$\boldsymbol{A},\boldsymbol{B}$，可取到一列有理数$t_k\to 0$，使得$t_k\boldsymbol{I}_n+\boldsymbol{A}$与$t_k\boldsymbol{I}_n+\boldsymbol{B}$均为非异阵.
由非异阵情形的证明可得 $\left(\left(t_k\boldsymbol{I}_n+\boldsymbol{A}\right)\left(t_k\boldsymbol{I}_n+\boldsymbol{B}\right)\right)^*=\left(t_k\boldsymbol{I}_n+\boldsymbol{B}\right)^*\left(t_k\boldsymbol{I}_n+\boldsymbol{A}\right)^*$.
上式两边均为n阶方阵，其元素都是t_k的多项式，从而关于t_k连续.上式两边同时取极限，令$t_k\to 0$，
即有 $(\boldsymbol{AB})^*=\boldsymbol{B}^*\boldsymbol{A}^*$成立.

Proposition 3.20

$|\boldsymbol{A}^*|=|\boldsymbol{A}|^{n-1}$

Proof 若$\boldsymbol{A}$是非异阵，用可逆阵来计算显然 $|\boldsymbol{A}^*|=|\boldsymbol{A}|^{n-1}$.
对于一般的方阵$\boldsymbol{A}$，可取到一列有理数$t_k\to 0$，使得$t_k\boldsymbol{I}_n+\boldsymbol{A}$为非异阵.由非异阵情形的证明可得 $\left|\left(t_k\boldsymbol{I}_n+\boldsymbol{A}\right)^*\right|=\left|t_k\boldsymbol{I}_n+\boldsymbol{A}\right|^{n-1}$.
注意到上式两边均为行列式的幂次，其值都是t_k的多项式，从而关于t_k连续.上式两边同时取极限
令$t_k\to 0$，即有 $|\boldsymbol{A}^*|=|\boldsymbol{A}|^{n-1}$成立.

Proposition 3.21

设A为m阶矩阵，B为n阶矩阵，求分块对角阵$C:\begin{pmatrix} A & O \\ O & B \end{pmatrix}$的伴随矩阵

Proof 若$\boldsymbol{A},\boldsymbol{B}$均为非异阵，则

$$\boldsymbol{C}\begin{pmatrix}|\boldsymbol{B}|\boldsymbol{A}^* & \boldsymbol{O}\\ \boldsymbol{O} & |\boldsymbol{A}|\boldsymbol{B}^*\end{pmatrix}=\begin{pmatrix}\boldsymbol{A} & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{B}\end{pmatrix}\begin{pmatrix}|\boldsymbol{B}|\boldsymbol{A}^* & \boldsymbol{O}\\ \boldsymbol{O} & |\boldsymbol{A}|\boldsymbol{B}^*\end{pmatrix}$$

$$=\begin{pmatrix}|\boldsymbol{B}|\boldsymbol{A}\boldsymbol{A}^* & \boldsymbol{O}\\ \boldsymbol{O} & |\boldsymbol{A}|\boldsymbol{B}\boldsymbol{B}^*\end{pmatrix}=\begin{pmatrix}|\boldsymbol{A}||\boldsymbol{B}|\boldsymbol{I}_m & \boldsymbol{O}\\ \boldsymbol{O} & |\boldsymbol{A}||\boldsymbol{B}|\boldsymbol{I}_n\end{pmatrix}$$

$$=|\boldsymbol{C}|\boldsymbol{I}_{m+n}=\boldsymbol{C}\boldsymbol{C}^*$$

注意到$\boldsymbol{C}$非异, 故由上式可得$\boldsymbol{C}^*=\begin{pmatrix}A & O\\ O & B\end{pmatrix}^*=\begin{pmatrix}|B|A^* & O\\ O & |A|B^*\end{pmatrix}$

对于一般的方阵$\boldsymbol{A},\boldsymbol{B}$, 可取到一列有理数$t_k\to 0$, 使得$t_k\boldsymbol{I}_m+\boldsymbol{A}$与$t_k\boldsymbol{I}_n+\boldsymbol{B}$均为非异阵.

由非异阵情形的证明可得$\begin{pmatrix}t_k\boldsymbol{I}_m+\boldsymbol{A} & \boldsymbol{O}\\ \boldsymbol{O} & t_k\boldsymbol{I}_n+\boldsymbol{B}\end{pmatrix}^*=\begin{pmatrix}|t_k\boldsymbol{I}_n+\boldsymbol{B}|\left(t_k\boldsymbol{I}_m+\boldsymbol{A}\right)^* & \boldsymbol{O}\\ \boldsymbol{O} & |t_k\boldsymbol{I}_m+\boldsymbol{A}|\left(t_k\boldsymbol{I}_n+\boldsymbol{B}\right)^*\end{pmatrix}$.

Proposition 3.22

设A,B,C,D是n阶矩阵且$AC=CA$, 求证 : $\begin{vmatrix}A & B\\ C & D\end{vmatrix}=|AD-CB|$.

Proof 若$\boldsymbol{A}$是非异阵, 则由降阶公式可得这是显然的

对于一般的方阵$\boldsymbol{A}$, 可取到一列有理数$t_k\to 0$, 使得$t_k\boldsymbol{I}_n+\boldsymbol{A}$为非异阵, 并且条件$\left(t_k\boldsymbol{I}_n+\boldsymbol{A}\right)\boldsymbol{C}=\boldsymbol{C}\left(t_k\boldsymbol{I}_n+\boldsymbol{A}\right)$仍然成立.

由非异阵情形的证明可得$\begin{vmatrix}t_k\boldsymbol{I}_n+\boldsymbol{A} & \boldsymbol{B}\\ \boldsymbol{C} & \boldsymbol{D}\end{vmatrix}=\left|\left(t_k\boldsymbol{I}_n+\boldsymbol{A}\right)\boldsymbol{D}-\boldsymbol{C}\boldsymbol{B}\right|$.

注意到上式两边均为行列式, 其值都是t_k的多项式, 从而关于t_k连续.上式两边同时取极限

令$t_k\to 0$, 即有$\begin{vmatrix}\boldsymbol{A} & \boldsymbol{B}\\ \boldsymbol{C} & \boldsymbol{D}\end{vmatrix}=|\boldsymbol{A}\boldsymbol{D}-\boldsymbol{C}\boldsymbol{B}|$成立.

Proposition 3.23

设$|A|=\left|a_{ij}\right|$是一个n阶行列式, A_{ij}是它的第(i,j)元素的代数余子式

求证 : $\begin{vmatrix}a_{11} & a_{12} & \cdots & a_{1n} & x_1\\ a_{21} & a_{22} & \cdots & a_{2n} & x_2\\ \vdots & \vdots & & \vdots & \vdots\\ a_{n1} & a_{n2} & \cdots & a_{nn} & x_n\\ y_1 & y_2 & \cdots & y_n & z\end{vmatrix}=z|A|-\sum_{i=1}^n\sum_{j=1}^n A_{ij}x_iy_j$

Proof 设$\boldsymbol{x}=\left(x_1,x_2,\cdots,x_n\right)',\boldsymbol{y}=\left(y_1,y_2,\cdots,y_n\right)'$.

若$\boldsymbol{A}$是非异阵, 则由降阶公式可得$\begin{vmatrix}\boldsymbol{A} & \boldsymbol{x}\\ \boldsymbol{y}' & 1\end{vmatrix}=|\boldsymbol{A}|\left(1-\boldsymbol{y}'\boldsymbol{A}^{-1}\boldsymbol{x}\right)=|\boldsymbol{A}|-\boldsymbol{y}'\boldsymbol{A}^*\boldsymbol{x}$.

对于一般的方阵$\boldsymbol{A}$, 可取到一列有理数$t_k\to 0$, 使得$t_k\boldsymbol{I}_n+\boldsymbol{A}$为非异阵.由非异阵情形的证明可得

$\begin{vmatrix}t_k\boldsymbol{I}_n+\boldsymbol{A} & \boldsymbol{x}\\ \boldsymbol{y}' & 1\end{vmatrix}=|t_k\boldsymbol{I}_n+\boldsymbol{A}|-\boldsymbol{y}'\left(t_k\boldsymbol{I}_n+\boldsymbol{A}\right)^*\boldsymbol{x}$.

注意到上式两边都是关于t_k的多项式, 从而关于t_k连续.上式两边同时取极限

令$t_k\to 0$, 即有$\begin{vmatrix}\boldsymbol{A} & \boldsymbol{x}\\ \boldsymbol{y}' & 1\end{vmatrix}=|\boldsymbol{A}|-\boldsymbol{y}'\boldsymbol{A}^*\boldsymbol{x}=|\boldsymbol{A}|-\sum_{i=1}^n\sum_{j=1}^n A_{ij}x_iy_j$

第 4 章 线性空间

4.1 行列向量

Definition 4.1

K是一个数域, V是一个集合, 在V上定义了加法十, 即对V中任意两个元素α,β, 存在V中唯一的元素γ与之对应, 记$\gamma = \alpha+\beta$.

在数域K与V之间定义了一种运算, 称为数乘, 即对K中任一数k及V中任一元α, 在V中总有唯一的元素δ与之对应, 记为$\delta = k\alpha$.

若上述加法及数乘满足下列运算规则

(1) 加法交换律 : $\alpha+\beta=\beta+\alpha$;

(2) 加法结合律 : $(\alpha+\beta)+\gamma=\alpha+(\beta+\gamma)$;

(3) 在V中存在一个元素0, 对于V中任一向量α, 都有$\alpha+0=\alpha$;

(4) 对于V中每个元素α, 存在元素β, 使$\alpha+\beta=0$;

(5) $1\cdot\alpha=\alpha$;

(6)$k(l\alpha)=(kl)\alpha$

(7)$k(\alpha+\beta)=k\alpha+k\beta, k\in \mathrm{K}$

(8)$(k+l)\alpha=k\alpha+l\alpha, k,l\in \mathrm{K}$;

在上述规则中, α,β,γ是V中任意的元素, k,l是K中任意的数.适合上述条件的集合V称为数域K上的线性空间或向量空间.

V中的元素称为向量, V中适合 (3)的元素0称为零向量.对V中的元素α适合$\alpha+\beta=0$的元素β称为α的负向量, 记为$-\alpha$.

Example 4.1 复数域C可看成是实数域R上的线性空间.这时C上向量的加法就是复数的加法.
R中元素对C中向量 (即复数) 的乘法就是通常的数的乘法.
一般来说, 若两个数域$\mathrm{K}_1\subseteq \mathrm{K}_2$, 则$\mathrm{K}_2$可以看成是$\mathrm{K}_1$上的线性空间.
向量就是K_2中的数, 向量的加法就是数的加法, 数乘就是K_1中的数乘认K_2中的数.
特别, 数域K也可以看成是K自身上的线性空间.

Proposition 4.1

零向量是唯一的.

负向量也是唯一的.

对任意的α,β,γ, 有

(1)从$\alpha+\beta=\alpha+\gamma$可推出$\beta=\gamma$, 即加法消去律成立

(2) $0\cdot\alpha=0$, 这里左边的0表示数零, 右边的0表示零向量;

(3)$k\cdot 0=0$

(4) $(-1)\alpha=-\alpha$;

(5) 若$k\alpha=0$, 则$\alpha=0$或$k=0$.

Proof 假设$0_1,0_2$是线性空间V中的两个零向量, 则$0_1=0_1+0_2=0_2$.这就证明了唯一性, 证毕.

设α是V中的向量, β_1,β_2也是V中的向量, 且$\alpha+\beta_1=0,\alpha+\beta_2=0$, 则

$\beta_1=\beta_1+0=\beta_1+(\alpha+\beta_2)=(\beta_1+\alpha)+\beta_2$

$=(\alpha+\beta_1)+\beta_2=0+\beta_2=\beta_2.$

这说明负向量是唯一的,证毕.

(1)$(-\alpha)+(\alpha+\beta)=(-\alpha)+(\alpha+\gamma)$,再由结合律得$((-\alpha)+\alpha)+\beta=((-\alpha)+\alpha)+\gamma$,即$0+\beta=0+\gamma$,于是$\beta=\gamma$.

(2)$0\cdot\alpha=(0+0)\alpha=0\cdot\alpha+0\cdot\alpha\Rightarrow\ 0+0\cdot\alpha=0\cdot\alpha+0\cdot\alpha$再由(1)两边消去$0\cdot\alpha$即得$0=0\cdot\alpha$.

(3)$k\alpha+k\cdot 0=k(\alpha+0)=k\alpha=k\alpha+0$,两边消去$k\alpha$即得$k\cdot 0=0$.

(4)$\alpha+(-1)\alpha=1\cdot\alpha+(-1)\alpha=[1+(-1)]\alpha=0\cdot\alpha=0$,因此$(-1)\alpha=-\alpha$.

(5)假定$k\neq 0$且$k\alpha=0$,则k^{-1}存在,故$\alpha=\left(k^{-1}\cdot k\right)\alpha=k^{-1}(k\alpha)=k^{-1}\cdot 0=0$.证毕. □

Definition 4.2 (线性组合)

设V是K上的线性空间,$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$和$\boldsymbol{\beta}$均是$V$中的向量,若存在K中$n$个数$k_1,k_2,\cdots,k_n$,使$\boldsymbol{\beta}=k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_n\boldsymbol{\alpha}_n$则称$\boldsymbol{\beta}$是$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$的线性组合或$\boldsymbol{\beta}$可由$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$线性表示.

设V是数域K上的线性空间,$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$是$V$中$n$个向量

若存在K中n个不全为零的数$k_1,k_2,\cdots,k_n$,使$k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_n\boldsymbol{\alpha}_n=0$,则称$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$线性相关

反之若K中不存在不全为零的数$k_1,k_2,\cdots,k_n$使上式成立,则称$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$线性无关或线性独立.

Definition 4.3

设有n个未知数m个方程式的线性方程组:$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1,\\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2,\\ \cdots\cdots\cdots\cdots\\ a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=b_m.\end{cases}$

记$\boldsymbol{A}=\begin{pmatrix} a_{11}&a_{12}&\cdots&a_{1n}\\ a_{21}&a_{22}&\cdots&a_{2n}\\ \vdots&\vdots&&\vdots\\ a_{m1}&a_{m2}&\cdots&a_{mn}\end{pmatrix}$ $\widetilde{\boldsymbol{A}}=\begin{pmatrix} a_{11}&a_{12}&\cdots&a_{1n}&b_1\\ a_{21}&a_{22}&\cdots&a_{2n}&b_2\\ \vdots&\vdots&&\vdots&\vdots\\ a_{m1}&a_{m2}&\cdots&a_{mn}&b_m\end{pmatrix}$ $\boldsymbol{x}=\begin{pmatrix} x_1\\ x_2\\ \vdots\\ x_n\end{pmatrix}$

分别用$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n;\boldsymbol{\beta}$表示上述矩阵的列向量,即$\boldsymbol{\alpha}_1=\begin{pmatrix} a_{11}\\ a_{21}\\ \vdots\\ a_{m1}\end{pmatrix},\boldsymbol{\alpha}_2=\begin{pmatrix} a_{12}\\ a_{22}\\ \vdots\\ a_{m2}\end{pmatrix},\cdots,\boldsymbol{\alpha}_n=\begin{pmatrix} a_{1n}\\ a_{2n}\\ \vdots\\ a_{mn}\end{pmatrix};\boldsymbol{\beta}=\begin{pmatrix} b_1\\ b_2\\ \vdots\\ b_m\end{pmatrix}$

则方程组等价于下列向量形式的方程式:$x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+\cdots+x_n\boldsymbol{\alpha}_n=\boldsymbol{\beta}$.或者$\boldsymbol{Ax}=0$

注 方程组有解$\Leftrightarrow x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+\cdots+x_n\boldsymbol{\alpha}_n=\boldsymbol{\beta}$中$\boldsymbol{\beta}$可由$\boldsymbol{\alpha}_1\sim\boldsymbol{\alpha}_n$线性表示

齐次方程组有非零解$\Leftrightarrow\boldsymbol{\alpha}_1\sim\boldsymbol{\alpha}_n$线性相关.

Example 4.2 标准单位行列向量组线性无关

只含有一个向量的向量组线性无关$\Leftrightarrow$该向量等于0

笔记 接下来我们讨论一种题型已知一些线性无关向量如何判断这些向量的组合的线性无关性

Proposition 4.2

设$\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_s$线性无关,并且$\begin{cases}\boldsymbol{\beta}_1=a_{11}\boldsymbol{\alpha}_1+\cdots+a_{1s}\boldsymbol{\alpha}_s\\ \boldsymbol{\beta}_2=a_{21}\boldsymbol{\alpha}_1+\cdots+a_{2s}\boldsymbol{\alpha}_s\\ \cdots\cdots\cdots\\ \boldsymbol{\beta}_s=a_{s1}\boldsymbol{\alpha}_1+\cdots+a_{ss}\boldsymbol{\alpha}_s\end{cases}$

证明：$\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_s$线性无关的充分必要条件是：$\begin{vmatrix}a_{11}&\cdots&a_{s1}\\ \vdots&&\vdots\\ a_{1s}&\cdots&a_{ss}\end{vmatrix}\neq 0.$

Proof 设$k_1\vec{\boldsymbol{\beta}}_1+\cdots+k_s\vec{\boldsymbol{\beta}}_s=0$,即$k_1\left(a_{11}\boldsymbol{\alpha}_1+\cdots+a_{1s}\boldsymbol{\alpha}_s\right)+\cdots+k_s\left(a_{s,1}\boldsymbol{\alpha}_1+\cdots+a_s\boldsymbol{\alpha}_s\right)=0.$

$\Longrightarrow\left(k_1a_{11}+\cdots+k_sa_{s1}\right)\boldsymbol{\alpha}_1+\cdots+\left(k_1a_{1s}+\cdots+k_sa_{ss}\right)\boldsymbol{\alpha}_s=0.$

则由于$\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_s$线性无关,因此从上式得$\begin{cases}k_1a_{11}+\cdots+k_sa_{s1}=0,\\ \qquad\cdots\cdots\cdots\cdots\\ k_1a_{1s}+\cdots+k_sa_{ss}=0.\end{cases}$

这个齐次线性方程组的系数行列式$|A|$为$|A|=\begin{vmatrix}a_{11}&\cdots&a_{s1}\\ \vdots&&\vdots\\ a_{1s}&\cdots&a_{ss}\end{vmatrix}.$

于是向量组$\vec{\boldsymbol{\beta}}_1,\cdots,\vec{\boldsymbol{\beta}}_s$线性无关

$\Longleftrightarrow k_1=0,\cdots,k_s=0$

$\Longleftrightarrow$ 上述齐次线性方程组只有零解

$\Longleftrightarrow |A|\neq 0$

Theorem 4.1.1 (线性相关性之替换定理)

设向量组$\vec{\boldsymbol{\alpha}}_1,\cdots,\vec{\boldsymbol{\alpha}}_s$线性无关,$\vec{\boldsymbol{\beta}}=b_1\vec{\boldsymbol{\alpha}}_1+\cdots+b_s\vec{\boldsymbol{\alpha}}_s$

证明：若$b_i\neq 0$,那么用$\vec{\boldsymbol{\beta}}$替换$\vec{\boldsymbol{\alpha}}_i$以后得到的向量组$\vec{\boldsymbol{\alpha}}_1,\cdots,\vec{\boldsymbol{\alpha}}_{i-1},\vec{\boldsymbol{\beta}},\vec{\boldsymbol{\alpha}}_{i+1},\cdots,\vec{\boldsymbol{\alpha}}_s$也线性无关

Proof 由于$\vec{\boldsymbol{\alpha}}_1,\cdots,\vec{\boldsymbol{\alpha}}_{i-1},\vec{\boldsymbol{\beta}},\vec{\boldsymbol{\alpha}}_{i+1},\cdots,\vec{\boldsymbol{\alpha}}_s$用$\vec{\boldsymbol{\alpha}}_1,\cdots,\vec{\boldsymbol{\alpha}}_s$表示出的行列式为下述

而$\vec{\boldsymbol{\alpha}}_1,\cdots,\vec{\boldsymbol{\alpha}}_{i-1},\vec{\boldsymbol{\beta}},\vec{\boldsymbol{\alpha}}_{i+1},\cdots,\vec{\boldsymbol{\alpha}}_s$线性无关的充要条件是行列式不等于0

把下述行列式按第i行展开,得$\begin{vmatrix}1&\cdots&0&b_1&0&\cdots&0\\ 0&\cdots&0&b_2&0&\cdots&0\\ \vdots&&\vdots&\vdots&\vdots&&\vdots\\ 0&\cdots&1&b_{i-1}&0&\cdots&0\\ 0&\cdots&0&b_i&0&\cdots&0\\ 0&\cdots&0&b_{i+1}&1&\cdots&0\\ \vdots&&\vdots&\vdots&\vdots&&\vdots\\ 0&\cdots&0&b_s&0&\cdots&1\end{vmatrix}=b_i\neq 0$

因此得,$\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_{i-1},\boldsymbol{\beta},\boldsymbol{\alpha}_{i+1},\cdots,\boldsymbol{\alpha}_s$线性无关

4.2 向量组的秩

Definition 4.4 (极大无关组)
设在线性空间V中有一族向量S(其中可能只有有限个向量,也可能有无限多个向量)
如果在S中存在一组向量$\{\alpha_1,\alpha_2,\cdots,\alpha_r\}$适合如下条件:
(1)$\alpha_1,\alpha_2,\cdots,\alpha_r$线性无关
(2) 这族向量中的任意一个向量都可以用$\alpha_1,\alpha_2,\cdots,\alpha_r$线性表示.
那么称$\{\alpha_1,\alpha_2,\cdots,\alpha_r\}$是向量族$S$的极大线性无关组,简称极大无关组.

Definition 4.5
向量族 S 的极大无关组所含的向量个数称为 S 的秩, 记作 $\mathrm{rank}(S)$ 或者 $\mathrm{r}(S)$

注 上述定义 (2)表明若将S中任一向量α加入$\{\alpha_1,\alpha_2,\cdots,\alpha_r\}$,则向量组$\{\alpha_1,\alpha_2,\cdots,\alpha_r,\alpha\}$一定线性相关.
正是在这个意义上,我们称$\{\alpha_1,\alpha_2,\cdots,\alpha_r\}$为极大无关组.

Theorem 4.2.1 (极大无关组的存在性定理)
设S为一个向量组且至少包含一个非零向量,则S的极大无关组一定存在

Proof 设S的向量个数为k,即$\sharp S=k$;对k用$mathematical\ induction$
显然当$\sharp S=1$时,成立. 假设当$\sharp S<k$时成立,下面当$\sharp S=k$时
$If\ S\ is\ linear\ uncorrelated, so\ the\ Maximal\ linear\ uncorrelated\ group\ is\ S$
$If\ S\ is\ linear\ correlated. According\ to\ former\ conclusion\ we\ know\ must$
$\exists\alpha_m\in S\ st.\alpha_m=\alpha_1+\alpha_2+\cdots+\alpha_k$其中$\alpha_1\cdots\alpha_k\in S\backslash\{\alpha\}$
此时$\sharp S\backslash\{\alpha\}=k-1$.那么$S\backslash\{\alpha\}$中必定存在一个非零向量.否则的话若全部为0由线性组合知道α_m也为0
那么S全部为0不符合题意.
那么针对$S\backslash\{\alpha\}$就可以使用归纳假设,存在最大无关组不妨设为$\alpha_1\sim\alpha_r$
那么$\alpha_1+\alpha_2+\cdots+\alpha_k$都可以被$\alpha_1\sim\alpha_r$表示那么$\alpha_m$也可以被$\alpha_1\sim\alpha_r$表示
所以$\alpha_1\sim\alpha_r$就为S的一个极大无关组证毕 □

Definition 4.6 (向量组与向量组的关系, 向量组的等价)
1.如果向量组$\alpha_1,\cdots,\alpha_s$的每一个向量都可以由向量组$\beta_1,\cdots,\beta_r$线性表出
那么称向量组$\alpha_1,\cdots,\alpha_s$可以由向量组$\beta_1,\cdots,\beta_r$线性表出
2.如果向量组$\alpha_1,\cdots,\alpha_s$与向量组$\beta_1,\cdots,\beta_r$可以互相线性表出
那么称向量组$\alpha_1,\cdots,\alpha_s$与向量组$\beta_1,\cdots,\beta_r$等价,记作$\{\alpha_1,\cdots,\alpha_s\}\cong\{\beta_1,\cdots,\beta_r\}$

Proposition 4.3
向量组的等价是一等价关系

Proof (1)反身性,即任何一个向量组都与自身等价
(2)对称性,即如果$\alpha_1,\cdots,\alpha_s$与$\beta_1,\cdots,\beta_r$等价,那么$\beta_1,\cdots,\beta_r$与$\alpha_1,\cdots,\alpha_s$等价
(3)传递性,即如果$\{\alpha_1,\cdots,\alpha_s\}\cong\{\beta_1,\cdots,\beta_r\},\{\beta_1,\cdots,\beta_r\}\cong\{\gamma_1,\cdots,\gamma_t\}$.那么$\{\alpha_1,\cdots,\alpha_s\}\cong\{\gamma_1,\cdots,\gamma_t\}$。
关于反身性和对称性,容易从定义立即得出
关于传递性,只需证明线性表出有传递性:

设$\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_s$可以由$\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_r$线性表出：$\boldsymbol{\alpha}_i=\sum_{j=1}^{r}b_{ij}\boldsymbol{\beta}_j,\quad i=1,\cdots,s$

且$\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_r$可以由$\boldsymbol{\gamma}_1,\cdots,\boldsymbol{\gamma}_t$线性表出：$\boldsymbol{\beta}_j=\sum_{l=1}^{t}c_{jl}\boldsymbol{\gamma}_l,\quad j=1,\cdots,r$

则$\boldsymbol{\alpha}_i=\sum_{j=1}^{r}b_{ij}\left(\sum_{l=1}^{t}c_{jl}\boldsymbol{\gamma}_l\right)=\sum_{l=1}^{t}\left(\sum_{j=1}^{r}b_{ij}c_{jl}\right)\boldsymbol{\gamma}_l,\quad i=1,\cdots,s$

即$\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_s$可以由$\boldsymbol{\gamma}_1,\cdots,\boldsymbol{\gamma}_t$线性表出。

Example 4.3 向量族的极大无关组并不唯一例子

设有向量组$S=\{(1,0),(0,1),(1,1)\}$.这3个向量线性相关

但是S的3个子集$\{(1,0),(0,1)\};\{(1,0),(1,1)\}$以及$\{(0,1),(1,1)\}$不难验证都是$S$的极大无关组.

因此一般来说,向量族的极大无关组并不唯一.

Corollary 4.1

1. 向量组与其极大无关组等价 2. 向量组的任意两个极大无关组等价

Proof 由定义很显然能得到

Lemma 4.1 (向量组与向量组关系与个数关系)

设A,B是V的两组向量,A含有r个向量,B含有s个向量,且A中每个向量均可由B中线性表示

若A中向量线性无关$\Rightarrow r\leqslant s$

逆否形式：若严格多的向量组可由严格少的向量组线性表示,则多的向量组线性相关

Proof 我们用反证法.设$A=\{\alpha_1,\alpha_2,\cdots,\alpha_r\},B=\left\{\beta_1,\beta_2,\cdots,\beta_s\right\}$.假设$r>s$.假定$A$组向量线性无关.

由已知,A中向量α_1可由B组向量的线性组合来表示,即存在数$\lambda_1,\lambda_2,\cdots,\lambda_s$,使$\alpha_1=\lambda_1\beta_1+\lambda_2\beta_2+\cdots+\lambda_s\beta_s$.$(*)$

因为A中向量线性无关,因$\alpha_1\neq 0$,故λ_i中至少有一个不为零,不妨假定$\lambda_1\neq 0$.由上$(*)$式解出

$\beta_1:\beta_1=\frac{1}{\lambda_1}\alpha_1-\frac{\lambda_2}{\lambda_1}\beta_2-\cdots-\frac{\lambda_s}{\lambda_1}\beta_s$.$(**)$ 但对任意的$\alpha_i(i=2,3,\cdots,r)$,已知α_i可由$\left\{\beta_1,\beta_2,\cdots,\beta_s\right\}$的线性组合表示

将$(**)$式代入α_i的表示式,则α_i可由$\left\{\alpha_1,\beta_2,\cdots,\beta_s\right\}$的线性组合表示.

这样,我们可将B组向量中的β_1换成α_1,这时A组中任一向量仍可用B组中向量的线性组合来表示.

现在我们归纳法,设B组问量已经换成$\left\{\alpha_1,\cdots,\alpha_k;\beta_{k+1},\cdots,\beta_s\right\}$,且$A$中任一向量都可以用$\left\{\alpha_1,\cdots,\alpha_k;\beta_{k+1},\cdots,\beta_s\right\}$线性表示

假设$k<r$,则α_{k+1}可表示为$\alpha_{k+1}=\mu_1\alpha_1+\cdots+\mu_k\alpha_k+\mu_{k+1}\beta_{k+1}+\cdots+\mu_s\beta_s$,其中至少有一个$\mu_i(i=k+1,\cdots,s)$不为零.

这是因为若$\mu_{k+1}=\cdots=\mu_s=0$,则α_{k+1}将可用$\{\alpha_1,\cdots,\alpha_k\}$线性表示,这与$A$组向量线性无关矛盾.

不失一般性,可设$\mu_{k+1}\neq 0$.用与上述相同的论证,又可将β_{k+1}换成α_{k+1},得到问量组$\left\{\alpha_1,\cdots,\alpha_{k+1};\beta_{k+2},\cdots,\beta_s\right\}$

且A中任一向量均可表示为这组向量的线性组合.这一事实表明,我们可将A中向量依次换入B.

但$r>s$,因此可将A中s个向量换入B组.不妨设B经调换以后的向量组为$\{\alpha_1,\alpha_2,\cdots,\alpha_s\}$

则A组向量中α_r也可用$\{\alpha_1,\alpha_2,\cdots,\alpha_s\}$的线性组合来表示,从而向量组$\{\alpha_1,\alpha_2,\cdots,\alpha_s,\alpha_r\}$线性相关引出矛盾.证毕. □

Proof 法二：我们证明引理的逆否形式

为了证明$\beta_1,\beta_2,\cdots,\beta_r$线性相关,需要找一组不全为0的数$k_1,k_2,\cdots,k_r$使得$k_1\beta_1+k_2\beta_2+\cdots+k_r\beta_r=0$
为此考虑$\beta_1,\beta_2,\cdots,\beta_r$的线性组合$x_1\beta_1+x_2\beta_2+\cdots+x_r\beta_r$

由已知条件,可设$\begin{cases}\beta_1=a_{11}\alpha_1+a_{21}\alpha_2+\cdots+a_{s1}\alpha_s\\ \beta_2=a_{12}\alpha_1+a_{22}\alpha_2+\cdots+a_{s2}\alpha_s\\ \cdots\cdots\cdots\\ \beta_r=a_{1r}\alpha_1+a_{2r}\alpha_2+\cdots+a_{sr}\alpha_s\end{cases}$

于是

$$\begin{aligned}x_1\beta_1+x_2\beta_2+\cdots+x_r\beta_r=&x_1\left(a_{11}\alpha_1+a_{21}\alpha_2+\cdots+a_{s1}\alpha_s\right)\\ &+x_2\left(a_{12}\alpha_1+a_{22}\alpha_2+\cdots+a_{s2}\alpha_s\right)\\ &+\cdots\cdots\cdots\\ &+x_r\left(a_{1r}\alpha_1+a_{2r}\alpha_2+\cdots+a_{sr}\alpha_s\right)\\ =&\left(a_{11}x_1+a_{12}x_2+\cdots+a_{1r}x_r\right)\alpha_1\\ &+\left(a_{21}x_1+a_{22}x_2+\cdots+a_{2r}x_r\right)\alpha_2\\ &+\cdots+\left(a_{s1}x_1+a_{s2}x_2+\cdots+a_{sr}x_r\right)\alpha_s\end{aligned}$$

考虑下述齐次线性方程组$\begin{cases}a_{11}x_1+a_{12}x_2+\cdots+a_{1r}x_r=0\\ a_{21}x_1+a_{22}x_2+\cdots+a_{2r}x_r=0\\ \cdots\cdots\cdots\\ a_{s1}x_1+a_{s2}x_2+\cdots+a_{s}x_r=0\end{cases}$

由已知条件$s<r$,因此方程组上述齐次方程组必有非零解.取它的一个非零解$(k_1,k_2,\cdots,k_r)$
则从式得$k_1\beta_1+k_2\beta_2+\cdots+k_r\beta_r=0\alpha_1+0\alpha_2+\cdots+0\alpha_s=0$因此$\beta_1,\beta_2,\cdots,\beta_r$线性相关。

Theorem 4.2.2

1.A,B都是向量族S的极大无关组,则A,B所含的向量个数相同,即秩相同
2.等价向量组具有相同的秩

Proof 由上述引理容易知道 □

Corollary 4.2 (秩与向量组表示的关系)

1.若$\{\alpha_1\cdots\alpha_m\}$能够被$\{\beta_1\cdots\beta_n\}$表示那么$r\left(\alpha_1\cdots\alpha_m\right)\leqslant r\left(\beta_1\cdots\beta_n\right)$
2.设向量组$\{\alpha_1\sim\alpha_s\};\{\beta_1\sim\beta_t\};\{\alpha_1\sim\alpha_s,\beta_1\sim\beta_t\}$的秩分别为$r_1,r_2,r_3\Longrightarrow\max\left(r_1,r_2\right)\leqslant r_3\leqslant r_1+r_2$
且证明r_3即为α与β的极大无关组放在一起再取秩的大小
3.向量组$\alpha_1\cdots\alpha_s$线性无关$\Longleftrightarrow r\left(\alpha_1\cdots\alpha_s\right)=s$

Proof

1.取二者的极大无关组利用引理即可,也称为表出向量组的秩不超过原向量组的秩,从几何上子空间也能看出这一点

2.由题干知道$\alpha_1\sim\alpha_s$可以由$\alpha_1\sim\alpha_s,\beta_1\sim\beta_t$表示; $\beta_1\sim\beta_t$可以由$\alpha_1\sim\alpha_s,\beta_1\sim\beta_t$表示;$\Rightarrow r_1\leqslant r_3;r_2\leqslant r_3\Rightarrow\max\left(r_1,r_2\right)\leqslant r_3$
不妨分别设极大线性无关组为$\alpha_{i_1}\cdots\alpha_{i_{r_1}};\beta_{j_1}\cdots\beta_{j_{r_2}};\gamma_{k_1}\cdots\gamma_{k_{r_3}}$所以$\gamma_{k_1}\cdots\gamma_{k_{r_3}}\overset{\text{等价}}{\Longleftrightarrow}\alpha_1\sim\alpha_s,\beta_1\sim\beta_t\overset{\text{等价}}{\Longleftrightarrow}\alpha_{i_1}\cdots\alpha_{i_{r_1}},\beta_{j_1}\cdots\beta_{j_{r_2}}$;
所以$r_3=R\left(\alpha_{i_1}\cdots\alpha_{i_{r_1}},\beta_1\cdots\beta_{j_{r_2}}\right)\leqslant r_1+r_2$(秩小于等于向量个数)

3.$\alpha_1\cdots\alpha_s$线性无关$\Longleftrightarrow$其极大无关组为自身$\Longleftrightarrow r\left(\alpha_1\cdots\alpha_s\right)=s$

Definition 4.7

设V是数域$\mathbb{K}$上线性空间，若在V中存在线性无关的向量$\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$，使$V$中任一向量均可表示为这组向量的线性组合

则称$\{\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n\}$是$V$的一组基，线性空间$V$称为$n$维线性空间(具有维数$n$).

此时由基的表述方法是唯一的，可由本节4.8命题得到

如果不存在有限个向量组成V的一组基，则称V是无限维线性空间.

对任一无限维空间，也有基的概念.无限维线性空间基的存在性证明超出了本课程的范围.

注 显然V中的一极大无关组是V的一组基，n维线性空间任意一组基都只含有n个向量.

若V是数域$\mathbb{K}$上的n维线性空间那么记作 $\dim_{\mathbb{K}} V = n$

Corollary 4.3

1.n维线性空间中任一组超过n个向量的向量组必线性相关.

2.逆否形式：若一个线性空间中，有含有超过n个向量的向量组是线性无关，那么秩$> n$

Proof

若m个问量$\alpha_1, \alpha_2, \cdots, \alpha_m$线性无关且$m > n$，$n$维线性空间的基为$\{e_1, e_2, \cdots, e_n\}$，则$\alpha_i$均可由$e_1, e_2, \cdots, e_n$的线性组合来表示

由引理逆否形式知$\alpha_1, \alpha_2, \cdots, \alpha_m$线性相关，证毕. □

判断一组基的条件比较严苛但如果我们事先知道了空间的维数则可以适当降低条件

Theorem 4.2.3 (有限维线性空间基判别法)

已知V是n维向量空间，$\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$是$V$中$n$个向量.若它们适合下列条件之一，则$\{\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n\}$是$V$的一组基.

(1) $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$线性无关.

(2) V中的任一向量均可由$\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$线性表示.

Proof (1)因为V中任意$n+1$个向量一定线性相关，故对任意V中向量$\boldsymbol{v}$，向量组$\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n, \boldsymbol{v}$线性相关.

于是存在不全为零的数$a_1, a_2, \cdots, a_n, c$，使$a_1\boldsymbol{e}_1 + a_2\boldsymbol{e}_2 + \cdots + a_n\boldsymbol{e}_n + c\boldsymbol{v} = 0$，其中$c \neq 0$.

事实上若$c = 0$，因为$\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$线性无关，将导致$a_1 = a_2 = \cdots = a_n = c = 0$，与假设矛盾.

由$c \neq 0$可得$\boldsymbol{v} = -\frac{a_1}{c}\boldsymbol{e}_1 - \frac{a_2}{c}\boldsymbol{e}_2 - \cdots - \frac{a_n}{c}\boldsymbol{e}_n$.因此$\boldsymbol{v}$可用向量组$\{e_1, e_2, \cdots, e_n\}$线性表示，即$\{e_1, e_2, \cdots, e_n\}$是一组基.

(2)不妨设$\{\boldsymbol{e}_1, \boldsymbol{e}_2 \cdots \boldsymbol{e}_r\}$是$\{\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n\}$的极大无关组，其中$r \leqslant n$.

那么我们就知道V中任一向量都可以由$\{\boldsymbol{e}_1, \boldsymbol{e}_2 \cdots \boldsymbol{e}_r\}$表示.因此$\{\boldsymbol{e}_1, \boldsymbol{e}_2 \cdots \boldsymbol{e}_r\}$是$V$的基

$\Rightarrow r = \dim V = n \quad \Rightarrow \{\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n\}$是$V$的一组基 □

Theorem 4.2.4 (基扩张定理)

设V是n维向量空间，$\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_m$是$V$中$m(m < n)$个线性无关的向量，又假定$\{\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n\}$是$V$的基

则必可在$\{\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n\}$中选出$n - m$个向量，使之和$\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_m$一起组成$V$的一组基.

Proof 将$\boldsymbol{e}_i (i = 1, \cdots, n)$依次放入$\{\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_m\}$，则必有一个$\boldsymbol{e}_i$，使$\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_m, \boldsymbol{e}_i$线性无关.

这是因为若任一$\boldsymbol{e}_i$加入$\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_m$后线性相关，则每个$\boldsymbol{e}_i$可用$\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_m$线性表示，将和引理的结论矛盾.

现不妨设$i = m + 1$.若$m + 1 < n$，又可从$\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$中找到一个向量，加入进去以后仍线性无关.

不断这样做下去，便可将$\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_m$扩张成为$V$的一组基.证毕. □

注 通常称为基扩张定理,常用的形式是:
n维线性空间V中任意$m(m<n)$个线性无关的向量均可扩张为V的一组基,或V的任意一个子空间的基均可扩张为V的一组基.

两组向量称为等价,若它们可以互相线性表示.注意向量组的等价与矩阵的等价不是一回事.
两个矩阵$\boldsymbol{A}$和$\boldsymbol{B}$等价是指通过初等变换可将$\boldsymbol{A}$变成$\boldsymbol{B}$.两个矩阵等价的充要条件是它们具有相同的秩.
但是两个向量组如果只具有相同的秩还不能保证它们等价.比如$A=\{(1,0)\}, B=\{(0,1)\}$.它们的秩都等于1,但它们不等价.
下面给出两个向量组等价的充要条件.

4.3 矩阵的秩

Theorem 4.3.1 (矩阵的行秩与列秩与初等变换的关系)
矩阵的行秩与列秩在初等变换下不变.

Proof 我们分两步走.第一步证明矩阵的行秩在行初等变换下不变,列秩在列初等变换下不变.
第二步证明列秩在行初等变换下不变,行秩在列初等变换下不变.

第一步,设$A=\left(a_{ij}\right)_{m\times n}$,为简单起见将它写成分块的形状:$A=\begin{pmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_m\end{pmatrix}$
其中$\alpha_i=(a_{i1},a_{i2},\cdots,a_{in})\ (i=1,2,\cdots,m)$是$A$的第$i$个行向量.
对换A的任意两行并不改变A的行向量组,因此也不改变A的行秩.这表明A在第一种行初等变换下行秩不变.

又若以一个非零常数k乘以A的第i行,则A变成:$A_1=\begin{pmatrix}\alpha_1\\ \vdots\\ k\alpha_i\\ \vdots\\ \alpha_m\end{pmatrix}$显然,$A_1$的$m$个行向量可用$A$的$m$个行向量的线性组合来表示.
反之A的m个行向量也可用A_1的行向量的线性组合来表示,因此A的行秩与A_1的行向量组等价因此秩相等.

接下来再看第三种行初等变换.将矩阵A的第i行乘以k后加到第j行上去,矩阵A变成了下列矩阵:$A_2=\begin{pmatrix}\alpha_1\\ \vdots\\ \alpha_i\\ \vdots\\ k\alpha_i+\alpha_j\\ \vdots\\ \alpha_m\end{pmatrix}$.
显然,A_2的向量是A中向量的线性组合.反之,$\alpha_j=(-k)\alpha_i+\left(k\alpha_i+\alpha_j\right)$.因此$A$与$A_2$行向量组等价因此秩相等
这就证明了A的行秩在行初等变换下不变.同理,A的列秩在列初等变换下也不变.

第二步,我们证明A的列秩在行初等变换下不变.
由于A的行初等变换等价于用一个初等矩阵左乘以A,我们只需证明对任一初等矩阵Q,QA与A的列秩相等就可以了.
现把A写成列向量形状:$A=\left(\beta_1,\beta_2,\cdots,\beta_n\right)$,其中$\beta_j$是$A$的第$j$个列向量.由分块矩阵的乘法得$QA=\left(Q\beta_1,Q\beta_2,\cdots,Q\beta_n\right)$.
设A的列向量的极大线性无关组为$\beta_1,\cdots,\beta_t$　现在我们要证明$\left\{Q\beta_1,\cdots,Q\beta_t\right\}$是$QA$的极大线性无关组.
这样的假定不失一般性,因为交换任意两列均不改变矩阵的列秩,我们经过若干次的交换后总可将极大线性无关组换到前t个.
先证明$Q\beta_1,\cdots,Q\beta_t$线性无关.设有$\lambda_1,\lambda_2,\cdots,\lambda_t\in\mathrm{K}$,使$\lambda_1Q\beta_1+\lambda_2Q\beta_2+\cdots+\lambda_tQ\beta_t=0$
则$Q\left(\lambda_1\beta_1+\lambda_2\beta_2+\cdots+\lambda_1\beta_t\right)=0$.但$Q$是非异阵,在上式两边左乘$Q^{-1}$即得$\lambda_1\beta_1+\lambda_2\beta_2+\cdots+\lambda_1\beta_t=0$.
再由$\beta_1,\cdots,\beta_t$线性无关即得$\lambda_1=\cdots=\lambda_t=0$.这就证明了$\{Q\beta_1,\cdots,Q\beta\}$是一组线性无关的向量.
再证明任一$Q\beta_j$均可表示为$Q\beta_1,\cdots,Q\beta_t$,的线性组合.由于$\beta_1,\cdots,\beta_t$是A的列向量组的极大线性无关组
故$\beta_j=\mu_1\beta_1+\mu_2\beta_2+\cdots+\mu_t\beta_t$.上式两边作用$Q$即得:$Q\beta_j=\mu_1Q\beta_1+\mu_2Q\beta_2+\cdots+\mu_tQ\beta_t$.
由上面的论证知道,A与QA的极大线性无关组都有相同个数的向量,因此A与QA的列秩相等.
同理可证明A的行秩在列初等变换下不变.证毕. □

Corollary 4.4

任一矩阵的行秩等于列秩.

有了这个推论, 我们今后不再讲行秩与列秩, 统称为秩.矩阵A的秩用$r(A)$或 rank$A$来表示.

Proof 任一矩阵A经初等变换后均可变成下列分块对角阵 : $B=\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$, 其中$B$是分块矩阵, I_r为r阶单位阵.
显然, B的行秩与列秩都等于r, 因此A的行秩与列秩都等于r.证毕. □

Corollary 4.5

1.任一矩阵A的转置A'与A有相同的秩.

2.任一矩阵与一非异阵相乘, 其秩不变.

3.n阶方阵为非异阵 $\Leftrightarrow$ 满秩阵

4.两个$m \times n$矩阵等价的充分必要条件是它们具有相同的秩.

Proof 显然

任一非异阵均可化为有限个初等矩阵的积, 由此即有结论.证毕.

若A为非异阵则$r(A)=r(AI_n)=r(I_n)=n \quad \Rightarrow A$满秩
若A为满秩阵则A经过初等变换可化为单位阵I_n从而非异

设矩阵A, B秩都等于r, 则它们都等价于r阶相抵标准型的矩阵, 因此A和B等价.
反之, 由于矩阵秩在初等变换下不变, 从A和B等价可知它们的秩相同.证毕. □

Corollary 4.6

设A为$m \times n$阶矩阵, 且A的第$j_1 \cdots j_r$列向量是A列向量的极大无关组
则对m阶非异阵Q, 矩阵QA的第$j_1 \cdots j_r$列向量也是QA列向量的极大无关组且具有相同线性关系

Proof 根据定理的证明过程即可知道 □

Theorem 4.3.2

设A为阶梯型矩阵.
则A的秩等于非零行个数且阶梯点所在列向量就为A列向量组的极大无关组

Proof 设A有r个非零行, 阶梯点为$a_{1k_1}, a_{2k_2} \cdots a_{rk_r}$

$$A=\begin{pmatrix} 0 & \cdots & a_{1k_1} & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & 0 & \cdots & a_{2k_2} & \cdots & \cdots & \cdots \\ \vdots & & \vdots & & \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & a_{rk_r} & \cdots \\ & & & & \boldsymbol{O} & & & \end{pmatrix}$$

先用第三类初等列变换和阶梯点上的元素依次消去同行的其他非零元素, 再用第二类初等列变换将阶梯点上的元素全部变成

1, 最后用列对换将 r 个阶梯点换到前面 r 个对角线点上得到

$$\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$$

进而得到 $r(A)=r$

对于第二个结论将 r 个阶梯点所在的列向量去除拼接成

$$\begin{pmatrix} a_{1k_1} & \cdots & \cdots & \cdots \\ 0 & a_{2k_2} & \cdots & \cdots \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{r_{kr}} \\ & \boldsymbol{O} & & \end{pmatrix}$$

用同样方法化为相抵标准型 $\begin{pmatrix} I_r \\ \boldsymbol{O} \end{pmatrix}$ 则r个阶梯点所在的列向量组的秩为r即A的秩为A的列秩, 则阶梯点的列向量是极大无关组

Theorem 4.3.3 (矩阵秩与子式行列式的关系)

设$m\times n$矩阵$A=(a_{ij})$有一个r阶子式不等于零, 且A中任意$r+1$阶子式 (如存在) 都等于零, 则$r(A)=r$.

反之, 若$r(A)=r$, 则A中必有一个r阶子式不等于零, 而所有$r+1$阶子式都等于零

Proof 设$r(A)=r$, 则A中任意$r+1$行都线性相关, 因此A的任意$r+1$阶子式的行向量也线性相关

由满秩阵和非异阵的等价性可知这些$r+1$阶子式的值均为零.

再证明A至少有一个r阶子式不等于零.因为A的秩为r, A中有r行线性无关.

不失一般性, 设为前r行.把这r行取出得到一个矩阵 : $B=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{r1} & a_{r2} & \cdots & a_{rn} \end{pmatrix}$.

显然$r(B)=r$, 因此B有r列线性无关, 同样不妨设为前r列, 则由B的前r列组成的行列式不等于零, 即A有一个r阶子式不等于零.

反之, 设A有一个r阶子式不为零而A的所有$r+1$阶子式全等于零.这时由行列式的定义可知, A的所有高于r阶的子式均等于零.

设$r(A)=t$, 则由前面的论述可知$t\geqslant r$, 否则A的r阶子式无一不为零.

但t也不能大于r, 否则A就要有一个大于r阶的子式不等于零而与假定矛盾, 因此$t=r$.证毕. □

Corollary 4.7

子矩阵的秩不超过大矩阵的秩

Proof 利用矩阵秩与子式行列式的关系得到

Proposition 4.4

设$\boldsymbol{A}$是数域K上的一个$s\times n$矩阵

证明 : 如果$\boldsymbol{A}$的秩为r, 那么$\boldsymbol{A}$的行向量组的一个极大线性无关组与$\boldsymbol{A}$的列向量组的一个极大线性无关组

交叉位置的元素按原来的排法组成的r阶子式不等于0

Proof 设$\boldsymbol{\gamma}_{i_1},\boldsymbol{\gamma}_{i_2},\cdots,\boldsymbol{\gamma}_{i_r}$是$\boldsymbol{A}$的行向量组$\boldsymbol{\gamma}_1,\boldsymbol{\gamma}_2,\cdots,\boldsymbol{\gamma}_s$的一个极大线性无关组

$\boldsymbol{\alpha}_{j_1},\boldsymbol{\alpha}_{j_2},\cdots,\boldsymbol{\alpha}_{j_r}$是$\boldsymbol{A}$的列向量组$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$的一个极大线性无关组

令$A_1 = \begin{pmatrix} \gamma_{i_1} \\ \gamma_{i_2} \\ \vdots \\ \gamma_{i_r} \end{pmatrix}$, 则 $\operatorname{rank}(A_1) = r$

A_1的列向量记作$\tilde{\alpha}_1, \tilde{\alpha}_2, \cdots, \tilde{\alpha}_n$, 它们是$\alpha_1, \alpha_2, \cdots, \alpha_n$的缩短组

由于A的每一列α_l可以由$\alpha_{j_1}, \alpha_{j_2}, \cdots, \alpha_{j_r}$线性表出, 因此$A_1$的每一列$\tilde{\alpha}_l$可以由$\tilde{\alpha}_{j_1}, \tilde{\alpha}_{j_2}, \cdots, \tilde{\alpha}_{j_r}$线性表出

由于 $\operatorname{rank}(A_1) = r$, 因此$\tilde{\alpha}_{1j}, \tilde{\alpha}_{j_2}, \cdots, \tilde{\alpha}_{j_r}$是$A_1$的列向量组的一个极大线性无关组

从而由$\tilde{\alpha}_{j_1}, \tilde{\alpha}_{j_2}, \cdots, \tilde{\alpha}_{j_r}$组成的子矩阵$A_2$的行列式不等于0, 即$A\begin{pmatrix} i_1, & i_2, & \cdots, & i_r \\ j_1, & j_2, & \cdots, & j_r \end{pmatrix} \neq 0$

Proposition 4.5

设A, B分别是数域K上的$s \times n$、$s \times m$矩阵, 用(A, B)表示在A的右边添写上B得到的$s \times (n + m)$矩阵

证明 : $\operatorname{rank}(A) = \operatorname{rank}((A, B))$当且仅当$B$的列向量组可以由$A$的列向量组线性表出

Proof 证法一

设A的列向量组为$\alpha_1, \alpha_2, \cdots, \alpha_n$; B的列向量组为$\beta_1, \beta_2, \cdots, \beta_m$, 则$(A, B)$的列向量组为$\{\alpha_1, \alpha_2, \cdots, \alpha_n, \beta_1, \beta_2, \cdots, \beta_m\}$

显然 $\langle \alpha_1, \alpha_2, \cdots, \alpha_n \rangle \subseteq \langle \alpha_1, \alpha_2, \cdots, \alpha_n, \beta_1, \beta_2, \cdots, \beta_m \rangle$

那么 $\operatorname{rank}(A) = \operatorname{rank}((A, B))$

$\Longleftrightarrow \dim \langle \alpha_1, \alpha_2, \cdots, \alpha_n \rangle = \dim \langle \alpha_1, \cdots, \alpha_n, \beta_1, \cdots, \beta_m \rangle$

$\Longleftrightarrow \langle \alpha_1, \alpha_2, \cdots, \alpha_n \rangle = \langle \alpha_1, \cdots, \alpha_n, \beta_1, \cdots, \beta_m \rangle$ (运用若$U \subseteq W$, 且$dimU = dimW$则$U = W$的子空间的结论)

$\Longleftrightarrow \beta_1, \beta_2, \cdots, \beta_m \in \langle \alpha_1, \alpha_2, \cdots, \alpha_n \rangle$

$\Longleftrightarrow B$ 的列向量组可以由A 的列向量组线性表出

证法二

设A, B的列向量组分别为$\alpha_1, \alpha_2, \cdots, \alpha_n; \beta_1, \beta_2, \cdots, \beta_m$

显然向量组$\alpha_1, \alpha_2, \cdots, \alpha_n$可以由向量组$\alpha_1, \alpha_2, \cdots, \alpha_n, \beta_1, \beta_2, \cdots, \beta_m$线性表出

那么$\operatorname{rank}(A) = \operatorname{rank}((A, B))$

$\Longleftrightarrow \operatorname{rank}\{\alpha_1, \alpha_2, \cdots, \alpha_n\} = \operatorname{rank}\{\alpha_1, \cdots, \alpha_n, \beta_1, \cdots, \beta_m\}$

$\Longleftrightarrow \{\alpha_1, \alpha_2, \cdots, \alpha_n\} \cong \{\alpha_1, \cdots, \alpha_n, \beta_1, \cdots, \beta_m\}$ (利用两个向量组等价的充要条件是秩相等且一组可被另一组表示)

$\Longleftrightarrow \beta_1, \cdots, \beta_m$可以由$\alpha_1, \alpha_2, \cdots, \alpha_n$线性表出

$\Longleftrightarrow B$ 的列向量组可以由A的列向量组线性表出

Proposition 4.6

1.设A是实数域上的$s \times n$矩阵, 则 $\operatorname{rank}(A'A) = \operatorname{rank}(AA') = \operatorname{rank}(A)$

2.证明 : 对于任意复矩阵A, 有 $\operatorname{rank}\left(A\overline{A}'\right) = \operatorname{rank}\left(\overline{A}'A\right) = \operatorname{rank}(A)$

Proof 如果能够证明n元齐次线性方程组$(A'A)x = 0$与$Ax = 0$同解

那么它们的解空间一致, 从而由解空间的维数公式, 得$n - \operatorname{rank}(A'A) = n - \operatorname{rank}(A)$由此得出, $\operatorname{rank}(A'A) = \operatorname{rank}(A)$

现在来证明$(A'A)x = 0$与$Ax = 0$同解

设η是$Ax = 0$的任意一个解, 则$A\eta = 0$, 从而$(A'A)\eta = 0$, 因此η是$(A'A)x = 0$的一个解;

反之, 设δ是$(A'A)x = 0$的任意一个解, 则$(A'A)\delta = 0$上式两边左乘δ', 得即

$\delta' A' A\delta = 0, (A\delta)' A\delta = 0$

设 $(A\delta)' = (c_1, c_2, \cdots, c_s)$, 由于$A$是实数域上的矩阵, 因此$c_1, c_2, \cdots, c_s$都是实数

得$c_1^2 + c_2^2 + \cdots + c_s^2 = 0$

由此推出, $c_1 = c_2 = \cdots = c_s = 0$, 从而 $\boldsymbol{A}\boldsymbol{\delta} = 0$, 即 $\boldsymbol{\delta}$ 是 $\boldsymbol{A}\boldsymbol{x} = 0$ 的一个解, 因此 $(\boldsymbol{A}'\boldsymbol{A})\boldsymbol{x} = 0$ 与 $\boldsymbol{A}\boldsymbol{x} = 0$ 同解

于是 $\operatorname{rank}(\boldsymbol{A}'\boldsymbol{A}) = \operatorname{rank}(\boldsymbol{A})$。由这个结论立即得出 $\operatorname{rank}(\boldsymbol{A}\boldsymbol{A}') = \operatorname{rank}\left[(\boldsymbol{A}')'(\boldsymbol{A}')\right] = \operatorname{rank}(\boldsymbol{A}') = \operatorname{rank}(\boldsymbol{A})$。

Proof 同样利用方程组同解可以证明 $\operatorname{rank}\left(\overline{\boldsymbol{A}}'\boldsymbol{A}\right) = \operatorname{rank}(\boldsymbol{A})$

接下来只要说明 $\boldsymbol{A}$ 在取转置和共轭操作下保持秩不变即可

转置显然不变

若是取共轭, 我们知道 $\boldsymbol{A}$ 的秩可以看为 $\boldsymbol{A}$ 的行向量组的秩, 不妨设 $\boldsymbol{\alpha}_1 \cdots \boldsymbol{\alpha}_r$ 是极大无关组

那么任意 $\boldsymbol{\alpha}_l$ 可以由 $\boldsymbol{\alpha}_1 \cdots \boldsymbol{\alpha}_r$ 表示 $\boldsymbol{\alpha}_l = k_1\boldsymbol{\alpha}_1 + \cdots + k_r\boldsymbol{\alpha}_r \Longrightarrow \overline{\boldsymbol{\alpha}_l} = \overline{k_1}\overline{\boldsymbol{\alpha}_1} + \cdots + \overline{k_r}\overline{\boldsymbol{\alpha}_r}$

Theorem 4.3.4 (矩阵幂次的秩控制定理)

1设 $\boldsymbol{A}$ 是数域 $\boldsymbol{K}$ 上的 n 级矩阵

证明：如果存在正整数 m, 使得 $\operatorname{rank}(\boldsymbol{A}^m) = \operatorname{rank}(\boldsymbol{A}^{m+1})$, 那么对一切正整数 k, 有 $\operatorname{rank}(\boldsymbol{A}^m) = \operatorname{rank}\left(\boldsymbol{A}^{m+k}\right)$

2.设 $\boldsymbol{A}$ 是数域 $\boldsymbol{K}$ 上的 n 级矩阵, 证明：对任意正整数 k, 有 $\operatorname{rank}\left(\boldsymbol{A}^{n+k}\right) = \operatorname{rank}(\boldsymbol{A}^n)$

证明如果 $\boldsymbol{A}$ 可逆, 那么 $\boldsymbol{A}^{n+k}, \boldsymbol{A}^n$ 都可逆, 从而 $\operatorname{rank}\left(\boldsymbol{A}^{n+k}\right) = n = \operatorname{rank}(\boldsymbol{A}^n)$

Proof 下面设 $\boldsymbol{A}$ 不可逆, 则 $\operatorname{rank}(\boldsymbol{A}) < n$.

由于 $\operatorname{rank}(\boldsymbol{A}) \geqslant \operatorname{rank}(\boldsymbol{A}^2) \geqslant \cdots \geqslant \operatorname{rank}(\boldsymbol{A}^n) \geqslant \operatorname{rank}(\boldsymbol{A}^{n+1})$

并且小于 n 的自然数只有 n 个, 因此上述 n 个 $\geqslant$ 中至少有一个取 $=$

即存在正整数 $m \leqslant n$, 使得 $\operatorname{rank}(\boldsymbol{A}^m) = \operatorname{rank}(\boldsymbol{A}^{m+1})$

据上个命题知道对一切正整数 k, 有 $\operatorname{rank}(\boldsymbol{A}^m) = \operatorname{rank}\left(\boldsymbol{A}^{m+k}\right)$

由于 $m \leqslant n$, 因此有 $\operatorname{rank}(\boldsymbol{A}^n) = \operatorname{rank}\left(\boldsymbol{A}^{n+k}\right)$。

我们来证此时会有 $\operatorname{rank}(\boldsymbol{A}^m) = \operatorname{rank}(\boldsymbol{A}^{m+1}) = \operatorname{rank}(\boldsymbol{A}^{m+2})$ 类似的用数学归纳即可

还是通过方程组同解来看, 此时 $\boldsymbol{A}^{m+1}\boldsymbol{x} = 0$ 的解一定是 $\boldsymbol{A}^{m+2}\boldsymbol{x} = 0$ 的解

取 $\boldsymbol{A}^{m+2}\boldsymbol{x} = 0$ 的解此时为 $\boldsymbol{A}^{m+1}(\boldsymbol{A}\boldsymbol{x}) = 0$ 把 $\boldsymbol{A}\boldsymbol{x}$ 看成 $\boldsymbol{A}^{m+1}\boldsymbol{y} = 0$ 的解

那么 $\boldsymbol{A}^{m}(\boldsymbol{A}\boldsymbol{x}) = 0 \Longrightarrow \boldsymbol{A}^{m+1}\boldsymbol{x} = 0$ 证毕

4.4 坐标向量

Lemma 4.2 (向量在基下的坐标唯一)

设 $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n\}$ 是n维线性空间V的一组基, 且$\boldsymbol{\alpha} = a_1\boldsymbol{e}_1 + a_2\boldsymbol{e}_2 + \cdots + a_n\boldsymbol{e}_n = b_1\boldsymbol{e}_1 + b_2\boldsymbol{e}_2 + \cdots + b_n\boldsymbol{e}_n$
则$a_1 = b_1, a_2 = b_2, \cdots, a_n = b_n$.

Proof 由假设得 : $(a_1 - b_1)\boldsymbol{e}_1 + (a_2 - b_2)\boldsymbol{e}_2 + \cdots + (a_n - b_n)\boldsymbol{e}_n = 0$.但$\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$线性无关
因此$a_i - b_i = 0 (i = 1, 2, \cdots, n)$, 即$a_i = b_i$. □

在建立了上述的引理之后我们就知道, 在线性空间 V 和 K^n 之间建立了一个双射 $\varphi : \begin{array}{l} V \to K^n \\ a_1e_1 + \cdots a_ne_n \longrightarrow (a_1, \cdots a_n) \end{array}$ 而且很显然该双射保持了加法和数乘这一点很重要

Definition 4.8 (线性空间的同构)

设V, U是数域K上的两个线性空间, 若存在V到U上一个 − −对应的映射φ, 使得对任意V中向量$\boldsymbol{\alpha}, \boldsymbol{\beta}$以及K中的数$k$均有$\varphi(\boldsymbol{\alpha} + \boldsymbol{\beta}) = \varphi(\boldsymbol{\alpha}) + \varphi(\boldsymbol{\beta}); \varphi(k\boldsymbol{\alpha}) = k\varphi(\boldsymbol{\alpha})$, 则称$V$与$U$这两个线性空间同构, 记为$V \cong U$.

Theorem 4.4.1

数域K上的任一n维线性空间均与K上的n维行向量空间同构.

Theorem 4.4.2

(1)设V, U同构, 同构映射为φ, 则$\varphi(0) = 0$;
(2) φ将线性相关的向量组映成线性相关的向量组, 将线性无关的向量组映成线性无关的向量组;
(3)同构关系是一个等价关系
即 $(i) V \cong V; (ii)$若$V \cong U$, 则$U \cong V; (iii)$若$V \cong U, U \cong W$, 则$V \cong W$;
(4) 数域K上的两个有限维线性空间同构的充分必要条件是它们具有相同的维数.

Proof (1)显然有$\varphi(0) = \varphi(0 + 0) = \varphi(0) + \varphi(0)$.消去一个$\varphi(0)$就有$\varphi(0) = 0$.

(2)若$\alpha_1, \alpha_2, \cdots, \alpha_m$是$V$中线性相关向量, 则存在不全为零的数$k_1, k_2, \cdots, k_m$, 使$k_1\alpha_1 + k_2\alpha_2 + \cdots + k_m\alpha_m = 0$
于是由 (1), 有$\varphi(k_1\alpha_1 + k_2\alpha_2 + \cdots + k_m\alpha_m) = 0$.上式左端为
$\varphi(k_1\alpha_1) + \varphi(k_2\alpha_2) + \cdots + \varphi(k_m\alpha_m) = k_1\varphi(\alpha_1) + k_2\varphi(\alpha_2) + \cdots + k_m\varphi(\alpha_m)$.
因此$\varphi(\alpha_1), \varphi(\alpha_2), \cdots, \varphi(\alpha_m)$是一组线性相关的向量.
另一方面, 若$\alpha_1, \alpha_2, \cdots, \alpha_m$线性无关且如果$\varphi(\alpha_1), \varphi(\alpha_2), \cdots, \varphi(\alpha_m)$在$U$中线性相关, 则存在一组不全为零的数$c_1, c_2, \cdots, c_m$
使$c_1\varphi(\alpha_1) + c_2\varphi(\alpha_2) + \cdots + c_m\varphi(\alpha_m) = 0$.但上式左端等于$\varphi(c_1\alpha_1 + c_2\alpha_2 + \cdots + c_m\alpha_m)$.
由于φ是一一对应的映射且已证明V中的零向量与U中的零向量对应, 因此$c_1\alpha_1 + c_2\alpha_2 + \cdots + c_m\alpha_m = 0$.
但$\alpha_1, \alpha_2, \cdots, \alpha_m$线性无关, 这就引出了矛盾, 故$\varphi(\alpha_1), \varphi(\alpha_2), \cdots, \varphi(\alpha_m)$必线性无关.

(3) (i) 显然V与自身同构, 这时取φ为恒同映射, 即$\varphi(\boldsymbol{\alpha}) = \boldsymbol{\alpha}, \boldsymbol{\alpha} \in V$.
(ii)设φ是$V \to U$上的一一对应.φ^{-1}是其逆对应 : $U \to V$.φ^{-1}也是一一对应.
设$\boldsymbol{x}, \boldsymbol{y}$是$U$中的向量, 由于$\varphi$是一一对应, 故存在$\boldsymbol{\alpha}, \boldsymbol{\beta} \in V$, 使$\varphi(\boldsymbol{\alpha}) = \boldsymbol{x}, \varphi(\boldsymbol{\beta}) = \boldsymbol{y}$, 也就是$\boldsymbol{\alpha} = \varphi^{-1}(\boldsymbol{x}), \boldsymbol{\beta} = \varphi^{-1}(\boldsymbol{y})$.
由φ是同构可知$\varphi(\boldsymbol{\alpha} + \boldsymbol{\beta}) = \boldsymbol{x} + \boldsymbol{y}, \varphi(k\boldsymbol{\alpha}) = k\boldsymbol{x}$, 因此$\varphi^{-1}(\boldsymbol{x} + \boldsymbol{y}) = \boldsymbol{\alpha} + \boldsymbol{\beta} = \varphi^{-1}(\boldsymbol{x}) + \varphi^{-1}(\boldsymbol{y}), \varphi^{-1}(k\boldsymbol{x}) = k\boldsymbol{\alpha} = k\varphi^{-1}(\boldsymbol{x})$.

这表明φ^{-1}是$U \to V$上的同构, 故$U \cong V$.
(iii)若φ是$V \to U$上的同构,ψ是$U \to W$上的同构.令$\xi = \psi\varphi$, 即对任意的$\alpha \in V$, 有$\xi(\alpha) = \psi(\varphi(\alpha))$, 则$\xi$是$V \to W$的一一对应, 且

$$\begin{aligned}\xi(\alpha+\beta) &= \psi(\varphi(\alpha+\beta)) = \psi(\varphi(\alpha)+\varphi(\beta)) \\ &= \psi(\varphi(\alpha)) + \psi(\varphi(\beta)) \\ &= \xi(\alpha)+\xi(\beta),\end{aligned}$$

$\xi(k\alpha) = \psi(\varphi(k\alpha)) = \psi(k\varphi(\alpha)) = k\psi(\varphi(\alpha)) = k\xi(\alpha)$.
这就是说ξ是$V \to W$上的同构.

(4)设V与U是K上的两个线性空间, $V \cong U$.
设 $\dim V = n$, $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n\}$ 是V的一组基, 则 $\{\varphi(\boldsymbol{e}_1), \varphi(\boldsymbol{e}_2), \cdots, \varphi(\boldsymbol{e}_n)\}$ 是U的一组线性无关的向量.
又若$\boldsymbol{x} \in U$, 则由于φ是一一对应, 因此存在$\alpha \in V$, 使$\boldsymbol{x} = \varphi(\alpha)$.设$\alpha = k_1\boldsymbol{e}_1 + k_2\boldsymbol{e}_2 + \cdots + k_n\boldsymbol{e}_n$,
则$\boldsymbol{x} = \varphi(\alpha) = k_1\varphi(\boldsymbol{e}_1) + k_2\varphi(\boldsymbol{e}_2) + \cdots + k_n\varphi(\boldsymbol{e}_n)$.即$U$中任一向量可表示为 $\{\varphi(\boldsymbol{e}_1), \varphi(\boldsymbol{e}_2), \cdots, \varphi(\boldsymbol{e}_n)\}$ 的线性组合,
故 $\{\varphi(\boldsymbol{e}_1), \varphi(\boldsymbol{e}_2), \cdots, \varphi(\boldsymbol{e}_n)\}$ 是U的一组基.因此 $\dim U = n = \dim V$.
反之, 若 $\dim U = \dim V = n$, 则U与V皆同构于n维行向量空间K^n, 由 (3) 可知V和U同构.证毕. □

Corollary 4.8

设 $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n\}$ 是线性空间V的基, $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$是$V$中向量.它们在这组基下的坐标向量依次为$\widetilde{\alpha}_1, \widetilde{\alpha}_2, \cdots, \widetilde{\alpha}_m$,
则向量组$\widetilde{\alpha}_1, \widetilde{\alpha}_2, \cdots, \widetilde{\alpha}_m$, 和向量组$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$有相同的秩.
且若$\boldsymbol{\beta}$能够被$\boldsymbol{\alpha}_1 \sim \boldsymbol{\alpha}_m$线性表示 $\Leftrightarrow$ $\widetilde{\boldsymbol{\beta}}$能够被$\widetilde{\boldsymbol{\alpha}_1} \sim \widetilde{\boldsymbol{\alpha}_m}$表示

Proof 同构映射将 $\{\alpha_1, \alpha_2, \cdots, \alpha_m\}$ 的极大无关组映为 $\{\tilde{\alpha}_1, \tilde{\alpha}_2, \cdots, \tilde{\alpha}_m\}$ 的极大无关组, 所以这两组向量的秩相同. □

这个推论可以使我们将一般的线性空间中向量组的求秩问题归结为行向量或列向量的求秩问题, 我们已经知道它可以用矩阵来处理. 特别, 判断向量组是否线性相关也可以这样做.

Example 4.4 设 $\{e_1, e_2, e_3\}$ 是线性空间V的基, 又 $\begin{cases}\boldsymbol{\alpha}_1 = \boldsymbol{e}_1 + 2\boldsymbol{e}_2 + 3\boldsymbol{e}_3, \\ \boldsymbol{\alpha}_2 = 2\boldsymbol{e}_1 - \boldsymbol{e}_2 - 3\boldsymbol{e}_3, \\ \boldsymbol{\alpha}_1 = \boldsymbol{e}_1 - 3\boldsymbol{e}_2 - 6\boldsymbol{e}_3,\end{cases}$ 求向量组 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3\}$ 的秩并判断它们是否线性相关.

Proof 因为矩阵$\begin{pmatrix}1 & 2 & 3 \\ 2 & -1 & -3 \\ 1 & -3 & -6\end{pmatrix}$的秩等于2, 所以向量组 $\{\alpha_1, \alpha_2, \alpha_3\}$ 的秩等于2.这3个向量线性相关. □

4.5 基变换与过渡矩阵

Definition 4.9 (过渡矩阵)

设 $\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\}$ 是数域K上线性空间V的一组基, $\{\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n\}$ 是另一组基

若$\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n$可用$\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n$的下列线性组合表示

$$\begin{cases}\boldsymbol{f}_1=a_{11}\boldsymbol{e}_1+a_{12}\boldsymbol{e}_2+\cdots+a_{1n}\boldsymbol{e}_n,\\ \boldsymbol{f}_2=a_{21}\boldsymbol{e}_1+a_{22}\boldsymbol{e}_2+\cdots+a_{2n}\boldsymbol{e}_n,\\ \qquad\ldots\cdots\cdots\cdots\\ \boldsymbol{f}_n=a_{n1}\boldsymbol{e}_1+a_{n2}\boldsymbol{e}_2+\cdots+a_{nn}\boldsymbol{e}_n,\end{cases}$$ 上述表示式中$\boldsymbol{e}_i$的系数组成了一个元素在K上的n阶矩阵.

这个矩阵的转置$A=\begin{pmatrix}a_{11}&a_{21}&\cdots&a_{n1}\\a_{12}&a_{22}&\cdots&a_{n2}\\\cdots&\cdots&\cdots&\cdots\\a_{1n}&a_{2n}&\cdots&a_{nn}\end{pmatrix}$称为从基 $\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\}$ 到基 $\{\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n\}$ 的过渡矩阵.

总结有 $\{\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n\}=\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\}\boldsymbol{A}$

Theorem 4.5.1 (坐标变换公式)

设 $\{e_1\cdots e_n\}\longrightarrow\{f_1\cdots f_n\}$ 的过渡矩阵为A, 设α在两组基e_i与f_i下的坐标为 $(\lambda_1\cdots\lambda_n)'$ 与 $(\mu_1\cdots\mu_n)'$

那么 $(\lambda_1\cdots\lambda_n)'=A\,(\mu_1\cdots\mu_n)'$

反之, 若V中向量在 $\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\}$ 下的坐标向量 $(\lambda_1,\lambda_2,\cdots,\lambda_n)'$ 和α在基 $\{\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n\}$ 下的坐标向量 $(\mu_1,\mu_2,\cdots,\mu_n)'$ 适合如变换公式的关系.注意到$\boldsymbol{f}_1$在 $\{\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n\}$ 下的坐标向量为 $(1,0,\cdots,0)'$

可知它在 $\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\}$ 下的坐标向量为 $(a_{11},a_{12},\cdots,a_{1n})'$, 这即是说$\boldsymbol{f}_1=a_{11}\boldsymbol{e}_1+a_{12}\boldsymbol{e}_2+\cdots+a_{1n}\boldsymbol{e}_n$.

同理, 对$\boldsymbol{f}_i$有$\boldsymbol{f}_i=a_{i1}\boldsymbol{e}_1+a_{i2}\boldsymbol{e}_2+\cdots+a_{in}\dot{\boldsymbol{e}}_n$.因此, 矩阵$A$就是从基 $\{e_1,e_2,\cdots,e_n\}$ 到基 $\{\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n\}$ 的过渡矩阵.

Theorem 4.5.2 (过渡矩阵的可逆性质)

两组基之间互相的过渡矩阵n阶方阵A,B互为逆阵, 即$B=A^{-1}$.

Proof 设$\alpha=\lambda_1\boldsymbol{e}_1+\lambda_2\boldsymbol{e}_2+\cdots+\lambda_n\boldsymbol{e}_n=\mu_1\boldsymbol{f}_1+\mu_2\boldsymbol{f}_2+\cdots+\mu_n\boldsymbol{f}_n$, 则

$$\begin{pmatrix}\lambda_1\\\lambda_2\\\vdots\\\lambda_n\end{pmatrix}=\begin{pmatrix}a_{11}&a_{21}&\cdots&a_{n1}\\a_{12}&a_{22}&\cdots&a_{n2}\\\cdots&\cdots&\cdots&\cdots\\a_{1n}&a_{2n}&\cdots&a_{nn}\end{pmatrix}\begin{pmatrix}\mu_1\\\mu_2\\\vdots\\\mu_n\end{pmatrix},\qquad\begin{pmatrix}\mu_1\\\mu_2\\\vdots\\\mu_n\end{pmatrix}=\begin{pmatrix}b_{11}&b_{21}&\cdots&b_{n1}\\b_{12}&b_{22}&\cdots&b_{n2}\\\cdots&\cdots&\cdots&\cdots\\b_{1n}&b_{2n}&\cdots&b_{nn}\end{pmatrix}\begin{pmatrix}\lambda_1\\\lambda_2\\\vdots\\\lambda_n\end{pmatrix},$$

因此 $\begin{pmatrix}\lambda_1\\\lambda_2\\\vdots\\\lambda_n\end{pmatrix}=\begin{pmatrix}a_{11}&a_{21}&\cdots&a_{n1}\\a_{12}&a_{22}&\cdots&a_{n2}\\\cdots&\cdots&\cdots&\cdots\\a_{1n}&a_{2n}&\cdots&a_{nn}\end{pmatrix}\begin{pmatrix}b_{11}&b_{21}&\cdots&b_{n1}\\b_{12}&b_{22}&\cdots&b_{n2}\\\cdots&\cdots&\cdots&\cdots\\b_{1n}&b_{2n}&\cdots&b_{nn}\end{pmatrix}\begin{pmatrix}\lambda_1\\\lambda_2\\\vdots\\\lambda_n\end{pmatrix}$.但$\lambda_i$可取K中任意数, 因此$AB=I_n$. □

问题是：我们假定从基 $\{e_1,e_2,\cdots,e_n\}$ 到基 $\{f_1,f_2,\cdots,f_n\}$ 的过渡矩阵为A.从 $\{\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n\}$ 到 $\{\boldsymbol{g}_1,\boldsymbol{g}_2,\cdots,\boldsymbol{g}_n\}$ 的过渡矩阵为B

那么从 $\{e_1,e_2,\cdots,e_n\}$ 到 $\{g_1,g_2,\cdots,g_n\}$ 的过渡矩阵是什么?

假定从基 $\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\}$ 到基 $\{\boldsymbol{g}_1,\boldsymbol{g}_2,\cdots,\boldsymbol{g}_n\}$ 的过渡矩阵是C.又假定

$$\begin{aligned}\boldsymbol{\alpha} &= \lambda_1 e_1 + \lambda_2 e_2 + \cdots + \lambda_n e_n \\ &= \mu_1 f_1 + \mu_2 f_2 + \cdots + \mu_n f_n \text{则} \\ &= \xi_1 g_1 + \xi_2 g_2 + \cdots + \xi_n g_n,\end{aligned}$$

$$\begin{pmatrix}\lambda_1\\ \lambda_2\\ \vdots\\ \lambda_n\end{pmatrix} = A\begin{pmatrix}\mu_1\\ \mu_2\\ \vdots\\ \mu_n\end{pmatrix}, \begin{pmatrix}\mu_1\\ \mu_2\\ \vdots\\ \mu_n\end{pmatrix} = B\begin{pmatrix}\xi_1\\ \xi_2\\ \vdots\\ \xi_n\end{pmatrix}, \text{于是}\begin{pmatrix}\lambda_1\\ \lambda_2\\ \vdots\\ \lambda_n\end{pmatrix} = AB\begin{pmatrix}\xi_1\\ \xi_2\\ \vdots\\ \xi_n\end{pmatrix}$$

这就是说AB是基$\{\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n\}$到基$\{\boldsymbol{g}_1, \boldsymbol{g}_2, \cdots, \boldsymbol{g}_n\}$的过渡矩阵.

Example 4.5 设在K^3中有两组基$\boldsymbol{f}_1 = (1, 0, -1), \boldsymbol{f}_2 = (2, 1, 1), \boldsymbol{f}_3 = (1, 1, 1)$和$g_1 = (0, 1, 1), g_2 = (-1, 1, 0), g_3 = (1, 2, 1)$. 求从$\{f_1, f_2, f_3\}$到$\{g_1, g_2, g_3\}$的过渡矩阵.

Proof 这道题如果我们直接做, 将面临解一个九元一次方程组, 非常麻烦. 我们利用上面的结论可以很快得到所要求的结果.

设$\boldsymbol{e}_1 = (1, 0, 0), \boldsymbol{e}_2 = (0, 1, 0), \boldsymbol{e}_3 = (0, 0, 1)$, 则从$\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$到$\{\boldsymbol{f}_1, f_2, f_3\}$的过渡矩阵为$A = \begin{pmatrix}1 & 2 & 1\\ 0 & 1 & 1\\ -1 & 1 & 1\end{pmatrix}$

从$\{e_1, e_2, e_3\}$到$\{g_1, g_2, g_3\}$的过渡矩阵为$B = \begin{pmatrix}0 & -1 & 1\\ 1 & 1 & 2\\ 1 & 0 & 1\end{pmatrix}$

则从$\{f_1, f_2, f_3\}$到$\{g_1, g_2, g_3\}$的过渡矩阵为$A^{-1}B$. 求矩阵$A^{-1}B$:

$$(A|B) = \left(\begin{array}{ccc|ccc}1 & 2 & 1 & 0 & -1 & 1\\ 0 & 1 & 1 & 1 & 1 & 2\\ -1 & 1 & 1 & 1 & 0 & 1\end{array}\right) \to \left(\begin{array}{ccc|ccc}1 & 2 & 1 & 0 & -1 & 1\\ 0 & 1 & 1 & 1 & 1 & 2\\ 0 & 3 & 2 & 1 & -1 & 2\end{array}\right) \to \left(\begin{array}{ccc|ccc}1 & 2 & 1 & 0 & -1 & 1\\ 0 & 1 & 1 & 1 & 1 & 2\\ 0 & 0 & -1 & -2 & -4 & -4\end{array}\right) \to \left(\begin{array}{ccc|ccc}1 & 2 & 1 & 0 & -1 & 1\\ 0 & 1 & 1 & 1 & 1 & 2\\ 0 & 0 & 1 & 2 & 4 & 4\end{array}\right) \to$$

$$\left(\begin{array}{ccc|ccc}1 & 2 & 0 & -2 & -5 & -3\\ 0 & 1 & 0 & -1 & -3 & -2\\ 0 & 0 & 1 & 2 & 4 & 4\end{array}\right) \to \left(\begin{array}{ccc|ccc}1 & 0 & 0 & 0 & 1 & 1\\ 0 & 1 & 0 & -1 & -3 & -2\\ 0 & 0 & 1 & 2 & 4 & 4\end{array}\right)$$ 因此从基$\{f_1, f_2, f_3\}$到基$\{g_1, g_2, g_3\}$的过渡矩阵为$P = \begin{pmatrix}0 & 1 & 1\\ -1 & -3 & -2\\ 2 & 4 & 4\end{pmatrix}$

4.6 子空间与直和与商空间

Definition 4.10 (线性子空间)

设V是数域$\mathbb{K}$上的线性空间,V_0是V的非空子集,且对V_0中的任意两个向量$\boldsymbol{\alpha},\boldsymbol{\beta}$及$\mathbb{K}$中任一数$k$,总有$\boldsymbol{\alpha}+\boldsymbol{\beta}\in V_0$及$k\boldsymbol{\alpha}\in V_0$
则称V_0是V的线性子空间,简称子空间
显然V_0在V的加法及数乘下是数域$\mathbb{K}$上的线性空间.

Definition 4.11 (平凡子空间)

任一线性空间V至少有两个子空间,一是零向量$\{0\}$组成的子空间,称为零子空间(维数规定为0);
另一个是V自身.这两个子空间通常称为平凡子空间.

如果V是n维线性空间,则由知道,V的任一子空间的维数不超过n.若V_0是V的非平凡子空间,则$0<\dim V_0<\dim V=n$.
事实上,若$\dim V_0=n$,则V_0有一组由n个线性无关向量组成的基.但任意n个线性无关的向量均可组成V的一组基
因此这时$V_0=V$,与V_0是非平凡子空间矛盾.

Theorem 4.6.1 (子空间之间的维数联系)

1.设U和W都是K^n的非零子空间,如果$U\subseteq W$,那么$\dim U\leqslant\dim W$
2.设U和W是K^n的两个非零子空间,且$U\subseteq W$,如果$\dim U=\dim W$,那么$U=W$

Proof 1.在U和W中分别取一个基$\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_r;\boldsymbol{\eta}_1,\cdots,\boldsymbol{\eta}_t$
因为$U\subseteq W$,所以$\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_r$可以由$\boldsymbol{\eta}_1,\cdots,\boldsymbol{\eta}_t$线性表出,从而$r\leqslant t$

2.U中取一个基$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$.由于$U\subseteq W$,因此$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r\in W$
由于$\dim W=\dim U=r$,因此$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$是$W$的一个基
从而W中任一向量$\boldsymbol{\beta}$可以由$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$线性表出,于是$\boldsymbol{\beta}\in U$,因此$W\subseteq U$,从而$U=W$。

Definition 4.12 (子空间的交与和)

若V_1,V_2是V的子空间,定义它们的交为既在V_1又在V_2中的全体向量所成的集合$V_1\cap V_2$.
定义它们的和为$V_1+V_2=\{\alpha+\boldsymbol{\beta}\mid\alpha\in V_1,\boldsymbol{\beta}\in V_2\}$,即所有形如$\boldsymbol{\alpha}+\boldsymbol{\beta}$的向量的集合,其中要求$\boldsymbol{\alpha}\in V_1,\boldsymbol{\beta}\in V_2$.

Proposition 4.7

$V_1\cap V_2,V_1+V_2$都是V的子空间.

Definition 4.13 (生成子空间)

设S是线性空间V的子集,记$L(S)$为S中向量所有可能的线性组合构成的子集
则不难看出,$L(S)$是V的一个子空间,称之为由集合S生成的子空间,或称之为由S张成的子空间.

Theorem 4.6.2 (生成子空间与极大无关组的结构定理)

设S是线性空间V的子集,$L(S)$为由S张成的子空间,则
(1)$S\subseteq L(S)$且若V_0是包含集合S的子空间,则$L(S)\subseteq V_0$,也即$L(S)$是包含S的V的最小子空间;
(2)$L(S)$的维数等于S中极大无关组所含向量的个数,且若$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$是$S$的极大无关组,则$L(S)=L\left(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m\right)$.

Proof (1)显然$S \subseteq L(S)$.设$\beta \in L(S)$,则β是S中若干个向輿$\alpha_1, \alpha_2, \cdots, \alpha_r$的线性组合:$\beta = \lambda_1\alpha_1 + \lambda_2\alpha_2 + \cdots + \lambda_r\alpha_r$.
因为$S \subseteq V_0$,由子空间的定义可知$\beta \in V_0$,因此$L(S) \subseteq V_0$.

(2)设$\alpha_1, \alpha_2, \cdots, \alpha_m$是$S$的极大无关组,则$S$中任一向量都是$\alpha_1, \alpha_2, \cdots, \alpha_r$的线性组合,即$S \subseteq L(\alpha_1, \alpha_2, \cdots, \alpha_m)$
因此$L(S) \subseteq L(\alpha_1, \alpha_2, \cdots, \alpha_m)$另一方面,显然有$L(\alpha_1, \alpha_2, \cdots, \alpha_m) \subseteq L(S)$,因此$L(S) = L(\alpha_1, \alpha_2, \cdots, \alpha_m)$.$L(S)$的维数等于$m$.□

Proposition 4.8

设V_1, V_2是线性空间V的子空间,则$L(V_1 \cup V_2) = V_1 + V_2$.
一般地,我们不难证明:若$V_1, V_2, \cdots, V_m$是V的子空间,则$L(V_1 \cup V_2 \cup \cdots \cup V_m) = V_1 + V_2 + \cdots + V_m$.

Proof 由生成的定义,对任意的$\alpha_1 \in V_1, \alpha_2 \in V_2, \alpha_1 + \alpha_2 \in L(V_1 \cup V_2)$,故$V_1 + V_2 \subseteq L(V_1 \cup V_2)$.
另一方面,因为$V_1 \subseteq V_1 + V_2, V_2 \subseteq V_1 + V_2$,由定理3.9.1(1),$L(V_1 \cup V_2) \subseteq V_1 + V_2$.于是$L(V_1 \cup V_2) = V_1 + V_2$.证毕. □

Theorem 4.6.3 (维数公式)

设V_1, V_2是V的子空间,则$\dim(V_1 + V_2) = \dim V_1 + \dim V_2 - \dim(V_1 \cap V_2)$.
设 $\dim V_1 = n_1, \dim V_2 = n_2, \dim V_1 \cap V_2 = m$.

Proof 取$V_1 \cap V_2$的一组基$\{\alpha_1, \cdots, \alpha_m\}$,由于$V_1 \cap V_2$是$V_1$的子空间
故可添上V_1中的向量$\alpha_{m+1}, \cdots, \alpha_{n_1}$,使$\{\alpha_1, \cdots, \alpha_m, \alpha_{m+1}, \cdots, \alpha_{n_1}\}$是$V_1$的一组基.
同样道理,可添上V_2中的向量$\beta_{m+1}, \cdots, \beta_{n_2}$,使$\{\alpha_1, \cdots, \alpha_m, \beta_{m+1}, \cdots, \beta_{n_2}\}$成为$V_2$的一组基.
由于$V_1 + V_2$中的任意向量均可由下列向量组:$\alpha_1, \cdots, \alpha_m, \alpha_{m+1}, \cdots, \alpha_{n_1}, \beta_{m+1}, \cdots, \beta_{n_2}$的线性组合给出.
如能证明上式中的向量线性无关,则它们构成V的一组基,由此即可推出所要的结论.
现假定$\lambda_1\alpha_1 + \cdots + \lambda_m\alpha_m + \lambda_{m+1}\alpha_{m+1} + \cdots + \lambda_{n_1}\alpha_{n_1} + \mu_{m+1}\beta_{m+1} + \cdots + \mu_{n_2}\beta_{n_2} = 0$,则
$\lambda_1\alpha_1 + \cdots + \lambda_m\alpha_m + \lambda_{m+1}\alpha_{m+1} + \cdots + \lambda_{n_1}\alpha_{n_1} = -(\mu_{m+1}\beta_{m+1} + \cdots + \mu_{n_2}\beta_{n_2})$.
上式左端属于V_1,右端属于V_2,故$\mu_{m+1}\beta_{m+1} + \cdots + \mu_{n_2}\beta_{n_2} \in V_1 \cap V_2$
即存在$\xi_1, \cdots, \xi_m \in \mathrm{K}$,使$\mu_{m+1}\beta_{m+1} + \cdots + \mu_{n_2}\beta_{n_2} = \xi_1\alpha_1 + \cdots + \xi_m\alpha_m$.
但$\alpha_1, \cdots, \alpha_m, \beta_{m+1}, \cdots, \beta_{n_2}$是$V_2$的基,因此$\mu_{m+1} = \cdots = \mu_{n_2} = \xi_1 = \cdots = \xi_m = 0$.
再由$\alpha_1, \cdots, \alpha_m, \alpha_{m+1}, \cdots, \alpha_{n_1}$线性无关得$\lambda_1 = \cdots = \lambda_m = \lambda_{m+1} = \cdots = \lambda_{n_1} = 0$.证毕. □

Definition 4.14 (直和定义)

设$V_1, V_2, \cdots, V_m$是线性空间V的子空间且对一切$i(= 1, 2, \cdots, m)$有 $V_i \cap (V_1 + \cdots + V_{i-1} + V_{i+1} + \cdots + V_m) = 0$
则称和$V_1 + V_2 + \cdots + V_m$为直接和,简称直和.记为$V_1 \oplus V_2 \oplus \cdots \oplus V_m$.

Theorem 4.6.4 (直和的等价判据)

设$V_1, V_2, \cdots, V_m$是线性空间V的子空间,$V_0 = V_1 + V_2 + \cdots + V_m$则下列命题等价:
(1)和$V_1 \oplus V_2 \oplus \cdots \oplus V_m$是直和;
(2) 对任意的$i\ (2 \leqslant i \leqslant m)$, $V_i \cap (V_1 + V_2 + \cdots + V_{i-1}) = 0$;
(3) $\dim(V_1 + V_2 + \cdots + V_m) = \dim V_1 + \dim V_2 + \cdots + \dim V_m$
(4) $V_1, V_2, \cdots V_m$的基可以拼成V_0的一组基
(5)V_0中的元素表示为$V_1, V_2, \cdots, V_m$中元素之和时其表示唯一,
即若$\alpha \in V_0$且$\alpha = v_1 + v_2 + \cdots + v_m = u_1 + u_2 + \cdots + u_m$,其中$\boldsymbol{v}_i, \boldsymbol{u}_i \in V_i$,则$\boldsymbol{u}_i = \boldsymbol{v}_i (i = 1, 2, \cdots, m)$; ($\Leftrightarrow$ 零向量表示唯一)

Proof (1) $\Rightarrow$ (2):显然.

$(2) \Rightarrow (3)$ 由维数公式知道对任意$2 \leqslant i \leqslant m$有
$\dim\left(V_1+V_2+\cdots+V_i\right)=\dim\left(V_1+V_2+\cdots+V_{i-1}\right)+\dim V_i$不断迭代即可得到

$(3) \Rightarrow (4)$ 依次取$V_1 \sim V_m$的一组基为$\boldsymbol{e}_{11}, \cdots \boldsymbol{e}_{1n_1}; \boldsymbol{e}_{21} \cdots \boldsymbol{e}_{2n_2}; \cdots\cdots; \boldsymbol{e}_{m1} \cdots \boldsymbol{e}_{mn_m}$
则 $\dim V_0=n_1+n_2+\cdots+n_m$且由$V_0=V_1+V_2+\cdots+V_m$知道$V_0$中向量就可以表示成上述的线性组合
那么由有限维空间的基向量判别法知道$\boldsymbol{e}_{11}, \cdots \boldsymbol{e}_{1n_1}; \boldsymbol{e}_{21} \cdots \boldsymbol{e}_{2n_2}; \cdots\cdots; \boldsymbol{e}_{m1} \cdots \boldsymbol{e}_{mn_m}$即为$V_0$的一组基

下面先证明 (5) 的等价命题 $(5')$ 零向量表示唯一：即若零向量表示唯一, 即$0=v_1+\cdots+v_m$其中$v_i \in V_i$则$v_i=0$
$(5) \Rightarrow (5')$ 显然
$(5') \Rightarrow (5)$ 若$\alpha=v_1+v_2+\cdots+v_m=u_1+u_2+\cdots+u_m$, 其中$\boldsymbol{v}_i, \boldsymbol{u}_i \in V_i$, 则$\boldsymbol{u}_i=\boldsymbol{v}_i (i=1,2,\cdots,m)$;
那么$0=\left(u_1-v_1\right)+\left(u_2-v_2\right)+\cdots+\left(u_m-v_m\right)$ 利用 $(5')$ 的条件易知

$(4) \Rightarrow (5')$ 依次取$V_1 \sim V_m$的一组基为$\boldsymbol{e}_{11}, \cdots \boldsymbol{e}_{1n_1}; \boldsymbol{e}_{21} \cdots \boldsymbol{e}_{2n_2}; \cdots\cdots; \boldsymbol{e}_{m1} \cdots \boldsymbol{e}_{mn_m}$于是他们拼成了$V_0$的一组基, 自然线性无关.
其次设$v_i=\lambda_{i1} e_{i1}+\cdots+\lambda_{in_i} e_{in_i}$由$V_0=V_1+V_2+\cdots+V_m$知道$0=v_1+v_2+\cdots+v_m$其中$v_i \in V_i$
那么$0=\lambda_{11} e_{11}+\cdots+\lambda_{1n_1} e_{1n_1}+\cdots+\lambda_{m1} e_{m1}+\cdots+\lambda_{mn_m} e_{mn_m} \Rightarrow \lambda_{ij}=0 \Rightarrow v_i=0$这就证明了 $(5')$

$(5') \Rightarrow (1)$ 任取$v \in V_i \cap\left(V_1+\cdots+V_{i-1}+V_{i+1}+\cdots+V_m\right)$则
$v=v_1+\cdots+v_{i-1}+v_{i+1}+\cdots+v_m$其中$v_j \in V_j\ (j=1 \cdots i-1, i+1, \cdots m)$
$\Rightarrow 0=v_1+\cdots+v_{i-1}+(-v)+v_{i+1}+\cdots+v_m$注意到$(-v) \in V_i$那么由 $(5')$ 知道$v=0$
于是由定义可得 □

下面来叙述如何求解两个子空间的交空间, 和空间的维数和基

Example 4.6 设$\boldsymbol{\alpha}_1=(1,0,-1,0),\boldsymbol{\alpha}_2=(0,1,2,1),\boldsymbol{\alpha}_3=(2,1,0,1)$是四维实行向量空间$V$中的向量, 它们生成的子空间为$V_1$
又向量$\boldsymbol{\beta}_1=(-1,1,1,1),\boldsymbol{\beta}_2=(1,-1,-3,-1),\boldsymbol{\beta}_3=(-1,1,-1,1)$生成的子空间为$V_2$
求子空间V_1+V_2和$V_1\cap V_2$的基.

解 1.利用矩阵进行简化不难得到 $\dim(V_1)=2\left(V_1\text{基为}\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2\right)\quad\dim(V_2)=2\left(V_2\text{基为}\boldsymbol{\beta}_1,\boldsymbol{\beta}_2\right)$

2.V_1+V_2由$\boldsymbol{\alpha}_i$和$\boldsymbol{\beta}_i$生成, 只要求出这6个向量的极大无关组即可.将6个向量拼成矩阵, 用初等行变换将其化为阶梯形矩阵:

$$\begin{pmatrix}1&0&2&-1&1&-1\\0&1&1&1&-1&1\\-1&2&0&1&-3&-1\\0&1&1&1&-1&1\end{pmatrix}\to\begin{pmatrix}1&0&2&-1&1&-1\\0&1&1&1&-1&1\\0&2&2&0&-2&-2\\0&0&0&0&0&0\end{pmatrix}\to\begin{pmatrix}1&0&2&-1&1&-1\\0&1&1&1&-1&1\\0&0&0&-2&0&-4\\0&0&0&0&0&0\end{pmatrix}$$

故可取$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1$为$V_1+V_2$的基 (不唯一).至此求出 $\dim(V_1+V_2)=3\quad\dim(V_1)=2\left(V_1\text{基为}\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2\right)$
注意V_1的基和维数在2.步骤中可以顺带求出来诚然也可以在1.问中求出; 但是在2.中V_2中就不是那么简单了

3.利用维数公式得到 $\dim(V_1\cap V_2)=1$; 接下来求基;
对任一向量$\boldsymbol{\gamma}\in V_1\cap V_2$, $\boldsymbol{\gamma}=x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2=(-x_3)\boldsymbol{\beta}_1+(-x_4)\boldsymbol{\beta}_2$.
因此, 求向量$\boldsymbol{\gamma}$等价于求解线性方程组$x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+x_3\boldsymbol{\beta}_1+x_4\boldsymbol{\beta}_2=0$

通过初等行变换将其系数矩阵$(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1,\boldsymbol{\beta}_2)$进行化简: $\begin{pmatrix}1&0&-1&1\\0&1&1&-1\\0&0&-2&0\\0&0&0&0\end{pmatrix}\to\begin{pmatrix}1&0&0&1\\0&1&0&-1\\0&0&1&0\\0&0&0&0\end{pmatrix}$,

故上述线性方程组的通解为$(x_1,x_2,x_3,x_4)=k(-1,1,0,1)$, 从而$\boldsymbol{\gamma}=-k(\boldsymbol{\alpha}_1-\boldsymbol{\alpha}_2)=-k\boldsymbol{\beta}_2(k\in\mathbb{R})$, 于是$\boldsymbol{\beta}_2$是$V_1\cap V_2$的基 □

Example 4.7 设V_1,V_2分别是数域$\mathbb{F}$上的齐次线性方程组$x_1=x_2=\cdots=x_n$与$x_1+x_2+\cdots+x_n=0$的解空间, 求证: $\mathbb{F}^n=V_1\oplus V_2$.

Proof 由线性方程组解的定理知, V_1的维数是1, V_2的维数是$n-1$.
若列向量$\boldsymbol{\alpha}\in V_1\cap V_2$, 则$\boldsymbol{\alpha}$既是第一个线性方程组的解, 也是第二个线性方程组的解, 不难看出$\boldsymbol{\alpha}$只能等于零向量, 因此$V_1\cap V_2=0$.
又因为 $\dim(V_1\oplus V_2)=\dim V_1+\dim V_2=1+(n-1)=n=\dim\mathbb{F}^n$, 故$\mathbb{F}^n=V_1\oplus V_2$.
一般来说证明两个空间的和是直和有以下思路:
1.先证明全空间$V=V_1+V_2$ 2.再利用直和证明条件来证明
1.先证明V_1+V_2是直和利用直和证明条件 2.利用 $\dim(V_1+V_2)=\dim(V_1\oplus V_2)=\dim(V_1)+dim(V_2)=\dim(V)$ □

4.7 商空间与补空间

Definition 4.15 (外直和定义)

设U,V是数域$\mathbb{K}$上的两个线性空间, $W=U\times V$是U和V的积集合, 即$W=\{(\boldsymbol{u},\boldsymbol{v})\mid \boldsymbol{u}\in U,\boldsymbol{v}\in V\}$.

现在W上定义加法和数乘 : $(\boldsymbol{u}_1,\boldsymbol{v}_1)+(\boldsymbol{u}_2,\boldsymbol{v}_2)=(\boldsymbol{u}_1+\boldsymbol{u}_2,\boldsymbol{v}_1+\boldsymbol{v}_2)$, $k(\boldsymbol{u},\boldsymbol{v})=(k\boldsymbol{u},k\boldsymbol{v})$.

设$U'=\{(\boldsymbol{u},0)\mid \boldsymbol{u}\in U\}, V'=\{(0,\boldsymbol{v})\mid \boldsymbol{v}\in V\}$,

1.W是$\mathbb{K}$上的线性空间 (这个线性空间称为U和V的外直和).

2.U',V'是W的子空间

3.U'和U同构, V'和V同构

4.$W=U'\oplus V'$.

Definition 4.16 (补空间定义以及存在性)

设U是V的子空间, 存在V的子空间W, 使得$V=U\oplus W$.这样的子空间W称为子空间U在V中的补空间.

Proof 取子空间U的一组基$\{\boldsymbol{e}_1,\cdots,\boldsymbol{e}_m\}$, 由基扩张定理可将其扩张为$V$的一组基$\{\boldsymbol{e}_1,\cdots,\boldsymbol{e}_m,\boldsymbol{e}_{m+1},\cdots,\boldsymbol{e}_n\}$.
令$W=L(\boldsymbol{e}_{m+1},\cdots,\boldsymbol{e}_n)$, 则$V=U+W$.事实上, $\{\boldsymbol{e}_{m+1},\cdots,\boldsymbol{e}_n\}$是$W$的一组基, 故 $\dim V=\dim U+\dim W$, 从而$V=U\oplus W$. □

注 在此$U\cap W=\{0\}$绝不是$\varnothing$; 同时$V=U+W$而不是$V=U\cup W$; 因此补空间绝不是补集;
一般来说补空间并不唯一; 若 $\dim V-\dim U\geqslant 1$且 $\dim U\geqslant 1$则U有无限个补空间, 可由证明过程感知取不同基

Definition 4.17 (商空间定义)

设V是数域$\mathbb{K}$上的线性空间, U是V的子空间.

对任意的$\boldsymbol{v}\in V$, 集合$\boldsymbol{v}+U:=\{\boldsymbol{v}+\boldsymbol{u}\ \ |\boldsymbol{u}\in U\}$称为$\boldsymbol{v}$的$U-$陪集　　所有$U-$陪集构成的集合$S=\{\boldsymbol{v}+U\ \ |\boldsymbol{v}\in V\}$

定义S中的加法和数乘如下 $(\boldsymbol{v}_1,\boldsymbol{v}_2\in V;k\in\mathbb{K})$:

$(\boldsymbol{v}_1+U)+(\boldsymbol{v}_2+U)=(\boldsymbol{v}_1+\boldsymbol{v}_2)+U \qquad k\cdot(\boldsymbol{v}_1+U)=k\cdot\boldsymbol{v}_1+U$

(1) : $U-$陪集之间的关系是 : 作为集合或者相等或者不相交

(2) : $\boldsymbol{v}_1+U=\boldsymbol{v}_2+U$相等 $\Leftrightarrow(\boldsymbol{v}_1-\boldsymbol{v}_2)\in U$; 特别的$\boldsymbol{v}+U$是$V$的子空间 $\Leftrightarrow \boldsymbol{v}\in U$

(3) : S中的加法和数乘不依赖于代表元的选取, 即 :

若$\boldsymbol{v}_1+U=\boldsymbol{v}_1'+U$且$\boldsymbol{v}_2+U=\boldsymbol{v}_2'+U$则 $(\boldsymbol{v}_1+U)+(\boldsymbol{v}_1'+U)=(\boldsymbol{v}_2+U)+(\boldsymbol{v}_2'+U)$ 以及$k\cdot(\boldsymbol{v}_1+U)=k\cdot(\boldsymbol{v}_1'+U)$

(4) : S在上述加法和数乘下称为数域$\mathbb{K}$上的线性空间, 称为V关于子空间的商空间记作V/U

(5) : 若W是U的补空间, 那么 $\dim(V/U)=\dim(V)-\dim(U)$ 且存在线性同构$\varphi:W\to V/U$

Proof (1) 设 $(\boldsymbol{v}_1+U)\cap(\boldsymbol{v}_2+U)=\varnothing$那么存在$\boldsymbol{u}_1,\boldsymbol{u}_2$ $st.$ $\boldsymbol{v}_1+\boldsymbol{u}_1=\boldsymbol{v}_2+\boldsymbol{u}_2\Rightarrow\boldsymbol{v}_1-\boldsymbol{v}_2=\boldsymbol{u}_2-\boldsymbol{u}_1\in U$
从而$\boldsymbol{v}_1+U=\boldsymbol{v}_2+(\boldsymbol{v}_1-\boldsymbol{v}_2)+U\subseteq\boldsymbol{v}_2+U$同理可证另一边证毕

(2) : 若$\boldsymbol{v}_1+U=\boldsymbol{v}_2+U$相等那么对于$\boldsymbol{v}_1+\boldsymbol{u}_1$必定存在一个$\boldsymbol{u}_2$ $st.$ $\boldsymbol{v}_2+\boldsymbol{u}_2=\boldsymbol{v}_1+\boldsymbol{u}_1$那么同样可以得到 $(\boldsymbol{v}_1-\boldsymbol{v}_2)\in U$
反之若 $(\boldsymbol{v}_1-\boldsymbol{v}_2)\in U$那么存在$\boldsymbol{u}_3$ $st.$ $\boldsymbol{v}_1-\boldsymbol{v}_2=\boldsymbol{u}_3\Rightarrow\boldsymbol{v}_1+\boldsymbol{u}_i=\boldsymbol{v}_2+\boldsymbol{u}_3+\boldsymbol{u}_i$所以 $(\boldsymbol{v}_1+U)\cap(\boldsymbol{v}_2+U)=\varnothing$推出相等

若$\boldsymbol{v}\in U$要证$\boldsymbol{v}+U$是V的子空间; 只要证对V中的加法和数乘封闭
任取$\boldsymbol{\gamma}_1$, $\boldsymbol{\gamma}_2\in\boldsymbol{v}+U$即$\boldsymbol{\gamma}_1=\boldsymbol{v}+\boldsymbol{u}_1$ $\boldsymbol{\gamma}_2=\boldsymbol{v}+\boldsymbol{u}_2$那么由$\boldsymbol{\gamma}_1+\boldsymbol{\gamma}_2=\boldsymbol{v}+(\boldsymbol{v}+\boldsymbol{u}_1+\boldsymbol{u}_2)\in\boldsymbol{v}+U$　数乘同理可证
反之若$\boldsymbol{v}+U$是V的子空间要证$\boldsymbol{v}\in U$
那么对于$0\in\boldsymbol{v}+U$那么存在$\boldsymbol{u}_1$ $st.$ $\boldsymbol{v}+\boldsymbol{u}_1=0\Rightarrow\boldsymbol{v}=-\boldsymbol{u}_1$那么$\boldsymbol{v}\in U$

(3) : 若$\boldsymbol{v}_1+U=\boldsymbol{v}_1'+U$且$\boldsymbol{v}_2+U=\boldsymbol{v}_2'+U$那么根据 (2) 知道存在$\boldsymbol{u}_1,\boldsymbol{u}_2$有$\boldsymbol{v}_1-\boldsymbol{v}_1'=\boldsymbol{u}_1\quad \boldsymbol{v}_2-\boldsymbol{v}_2'=\boldsymbol{u}_2$
要证 $(\boldsymbol{v}_1+U)+(\boldsymbol{v}_2+U)=(\boldsymbol{v}_2+U)+(\boldsymbol{v}_2'+U)$ 即证 $(\boldsymbol{v}_1+\boldsymbol{v}_2)+U=(\boldsymbol{v}_1'+\boldsymbol{v}_2')+U$
即证 $(\boldsymbol{v}_1+\boldsymbol{v}_2)-(\boldsymbol{v}_1'+\boldsymbol{v}_2')\in U$　即$\boldsymbol{u}_1+\boldsymbol{u}_2\in U$显然证毕　　另一个式子同理可以得到

(4) : 简单验证即可 (注意到 : 数乘单位元是1　零元是$0+U$)

(5) : 取子空间U的一组基 $\{\boldsymbol{e}_1\cdots\boldsymbol{e}_m\}$; 补空间的一组基 $\{\boldsymbol{e}_{m+1}\cdots\boldsymbol{e}_n\}$ 则 $\{e_1\cdots e_m,e_{m+1}\cdots e_n\}$ 是V的一组基
那么我们就证 $\{\boldsymbol{e}_{m+1}+U\cdots e_n+U\}$ 是商空间V/U的一组基.
一方面, 对任意的$\boldsymbol{v}\in V$; 设$\boldsymbol{v}=\sum_{i=1}^{n}a_i\boldsymbol{e}_i$那么$\boldsymbol{v}+U=\sum_{i=1}^{n}a_i\boldsymbol{e}_i+U=\sum_{i=m+1}^{n}a_i\boldsymbol{e}_i+U=\sum_{i=m+1}^{n}a_i(\boldsymbol{e}_i+U)$.
另一方面设有$k_{m+1}\cdots k_n\ \ st.\ \sum_{i=m+1}^{n}k_i(\boldsymbol{e}_i+U)=0+U\Rightarrow\left(\sum_{i=m+1}^{n}k_i\boldsymbol{e}_i\right)+U=U\Rightarrow\sum_{i=m+1}^{n}k_i\boldsymbol{e}_i\in U$
于是存在系数$l_1\cdots l_m\ \ st.\ \sum_{i=m+1}^{n}k_i\boldsymbol{e}_i=-\sum_{i=1}^{m}l_i\boldsymbol{e}_i\Rightarrow k_i=l_i=0$
因此证毕所以 $\dim(V/U)=\dim(V)-\dim(U)$
对任意的$\boldsymbol{w}\in W$; 设$\boldsymbol{w}=\sum_{i=m+1}^{n}a_i\boldsymbol{e}_i$　定义映射$\varphi:W\to V/U\quad \varphi(\boldsymbol{w})=\boldsymbol{w}+U=\sum_{i=m+1}^{n}a_i\boldsymbol{e}_i+U$
容易验证φ满足加法和数乘; 且是一一对应从而是线性同构 $\square$

4.8 线性方程组的解的结构

Theorem 4.8.1 (线性方程组解的存在定理)

设有n个未知数m个方程式组成的线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ \qquad \cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m, \end{cases}$$

它的系数矩阵记为A,增广矩阵记为$\widetilde{A}$,即$\widetilde{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{pmatrix}$则

(1) $\widetilde{A}$与A的秩相等则方程组有解.若$\widetilde{A}$与A的秩不相等,则该方程组无解.

(2)若$\widetilde{A}$与A的秩都等于n,则该方程组有且只有唯−−组解;

(3)若$\widetilde{A}$与A的秩相等但小于n,即$r(\widetilde{A}) = r(A) < n$,则该方程组有无穷多组解;

Proof 首先我们证明方程组有解的充分必要条件是$r(\tilde{A}) = r(A)$.把方程组写成向量形式就是$x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n = \beta$,
其中α_i为矩阵A的第i个列向量,β为常数项向量.方程组有解等同于β是$\alpha_1, \alpha_2, \cdots, \alpha_n$的线性组合.因此$r(\widetilde{A}) = r(A)$.
反过来,若$r(\widetilde{A}) = r(A)$,则A的列向量的极大线性无关组就是$\widetilde{A}$的列向量的极大线性无关组.
因此β可表示为A的列向量的线性组合,即方程组有解.

若再有$r(\widetilde{A}) = r(A) = n$,此时$A$的$n$个列向量线性无关,所以$\beta$只有唯一种方法表示为$\alpha_1, \alpha_2, \cdots, \alpha_n$的线性组合,即方程组只有唯一组解.

若$r(\widetilde{A}) = r(A) = r < n$.则$\alpha_1, \alpha_2, \cdots, \alpha_n$线性相关,即存在不全为零的数$k_1, k_2, \cdots, k_n$,使得$k_1\alpha_1 + k_2\alpha_2 + \cdots + k_n\alpha_n = 0$.
这时对任意的数k, $kk_1\alpha_1 + kk_2\alpha_2 + \cdots + kk_n\alpha_n = 0$.
因此若$x_1 = c_1, x_2 = c_2, \cdots, x_n = c_n$是方程组的解,则$x_1 = kk_1 + c_1, x_2 = kk_2 + c_2, \cdots, x_n = kk_n + c_n$都是解,显然这样的解有无穷多组.

Theorem 4.8.2 (齐次与非齐次方程组解的结构)

设有齐次线性方程组$A\boldsymbol{x} = 0$,其中$A = (a_{ij})$是$m \times n$矩阵.若$r(A) = r < n$,则上述方程组有非零解.
它的解构成n维列向量空间的一个$n - r$维子空间.也就是说,存在$n - r$个向量构成的基础解系$\{\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \cdots, \boldsymbol{\eta}_{n\cdots r}\}$
使齐次方程组的任一组解均可表示为$\{\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \cdots, \boldsymbol{\eta}_{n-r}\}$的线性组合.

设有非齐次线性方程组,它的系数矩阵A及其增广矩阵$\widetilde{A}$的秩都等于$r, r < n$.
又假定方程组的相伴齐次方程组有基础解系$\{\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \cdots, \boldsymbol{\eta}_{n-r}\}$,又$\boldsymbol{\gamma}$是方程组的任一特解,则其所有解均可表示为如下形状
$k_1\boldsymbol{\eta}_1 + k_2\boldsymbol{\eta}_2 + \cdots + k_{n-r}\boldsymbol{\eta}_{n\cdots r} + \boldsymbol{\gamma}$,其中$k_1, k_2, \cdots, k_{n-r}$可取任何数.

Proof 先来证明齐次线性方程组的解的结构
若 $\mathrm{rank}(\boldsymbol{A}) = n$,则方程组只有唯一解即零解,则$W = \{0\}$,从而解的结构定理显然成立
下面设 $\mathrm{rank}(\boldsymbol{A}) = r < n$.
第一步:
把$\boldsymbol{A}$经过初等行变换化成简化行阶梯形矩阵$\boldsymbol{J}$, $\boldsymbol{J}$有r个主元,不妨设它们分别在第$1, 2, \cdots, r$列

于是齐次线性方程组的一般解为 $\begin{cases} x_1 = -b_{1,r+1}x_{r+1} - \cdots - b_{1n}x_n \\ x_2 = -b_{2,r+1}x_{r+1} - \cdots - b_{2n}x_n \\ \cdots\cdots\cdots \\ x_r = -b_{r,r+1}x_{r+1} - \cdots - b_{rn}x_n \end{cases}$ 其中$x_{r+1},\cdots,x_n$是自由未知量

第二步

让自由未知量$x_{r+1},\cdots,x_n$分别取下述$n-r$组数：$\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix},\begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix},\cdots,\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}(*)$

则由一般解公式得到方程组的$n-r$个解：$\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\cdots,\boldsymbol{\eta}_{n-r}$

其中$\boldsymbol{\eta}_1 = \begin{pmatrix} -b_{1,r+1} \\ -b_{2,r+1} \\ \vdots \\ -b_{r,r+1} \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix},\cdots,\boldsymbol{\eta}_{n-r} = \begin{pmatrix} -b_{1n} \\ -b_{2n} \\ \vdots \\ -b_{rn} \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$由于式 $(*)$ 中的向量组线性无关, 因此它们的延伸组$\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\cdots,\boldsymbol{\eta}_{n-r}$也线性无关

第三步

任取齐次线性方程组的一个解$\boldsymbol{\eta}$：$\boldsymbol{\eta} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$, 于是$\boldsymbol{\eta}$满足齐次线性方程组的一般解公式 $\begin{cases} c_1 = -b_{1,r+1}c_{r+1} - \cdots - b_{1n}c_n \\ c_2 = -b_{2,r+1}c_{r+1} - \cdots - b_{2n}c_n \\ \cdots\cdots\cdots \\ c_r = -b_{r,r+1}c_{r+1} - \cdots - b_{rn}c_n \end{cases}$

即从而解向量$\boldsymbol{\eta}$可以写成下述形式：

$$\boldsymbol{\eta} = \begin{pmatrix} c_1 \\ \vdots \\ c_r \\ c_{r+1} \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} -b_{1,r+1}c_{r+1} & - & \cdots & - & b_{1n}c_n \\ \vdots & & & & \vdots \\ -b_{r,r+1}c_{r+1} & - & \cdots & - & b_{rn}c_n \\ 1c_{r+1} & + & \cdots & + & 0c_n \\ \vdots & & & & \vdots \\ 0c_{r+1} & + & \cdots & + & 1c_n \end{pmatrix} = \begin{pmatrix} -b_{1,r+1} \\ \vdots \\ -b_{r,r+1} \\ 1 \\ \vdots \\ 0 \end{pmatrix} c_{r+1} + \cdots + \begin{pmatrix} -b_{1n} \\ \vdots \\ -b_{rn} \\ 0 \\ \vdots \\ 1 \end{pmatrix} c_n = c_{r+1}\boldsymbol{\eta}_1 + \cdots + c_n\boldsymbol{\eta}_{n-r},$$

因此齐次线性方程组的每一个解$\boldsymbol{\eta}$可以由$\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\cdots,\boldsymbol{\eta}_{n-r}$线性表出, 从而$\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\cdots,\boldsymbol{\eta}_{n-r}$是方程组的一个基础解系

它包含的解向量的个数为$n-\mathrm{rank}(\boldsymbol{A})$, 于是 $\dim W = n-\mathrm{rank}(\boldsymbol{A})$

Proof

如果数域K上n元非齐次线性方程组$\boldsymbol{A}\boldsymbol{x}=\boldsymbol{\beta}$有解, 那么它的解集$U$为$U=\{\boldsymbol{\gamma}_0+\boldsymbol{\eta} \mid \boldsymbol{\eta}\in W\}$

其中$\boldsymbol{\gamma}_0$是非齐次线性方程组$\boldsymbol{A}\boldsymbol{x}=\boldsymbol{\beta}$的一个解 (称$\boldsymbol{\gamma}_0$是特解), W是方程组$\boldsymbol{A}\boldsymbol{x}=\boldsymbol{\beta}$的导出组$\boldsymbol{A}\boldsymbol{x}=\vec{0}$的解空间

任取$\boldsymbol{\eta}\in W$, 那么得, $\boldsymbol{\gamma}_0+\boldsymbol{\eta}\in U$, 因此 $\{\boldsymbol{\gamma}_0+\boldsymbol{\eta} \mid \boldsymbol{\eta}\in W\}\supseteq U$

反之, 任取$\boldsymbol{\gamma}\in U$得, $\boldsymbol{\gamma}-\boldsymbol{\gamma}_0\in W$.记$\boldsymbol{\gamma}-\boldsymbol{\gamma}_0=\boldsymbol{\eta}$, 则$\boldsymbol{\gamma}=\boldsymbol{\gamma}_0+\boldsymbol{\eta}$.因此$U$包含 $\{\boldsymbol{\gamma}_0+\boldsymbol{\eta} \mid \boldsymbol{\eta}\in W\}$

我们把 $\{\boldsymbol{\gamma}_0+\boldsymbol{\eta} \mid \boldsymbol{\eta}\in W\}$ 记作$\boldsymbol{\gamma}_0+W$, 称它是一个W型的线性流形 (或子空间W的一个陪集), 把 $\dim W$称为线性流形γ_0+W的维数

Proposition 4.9

设V_0是数域$\mathbb{K}$上n维列向量空间的真子空间

求证：必存在矩阵$\boldsymbol{A}$，使V_0是n元齐次线性方程组$\boldsymbol{A}\boldsymbol{x}=0$的解空间.

Proof 设$\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_r$是子空间$V_0$的一组基.令$\boldsymbol{B}=(\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_r)$，这是一个$n\times r$矩阵.
考虑齐次线性方程组$\boldsymbol{B}'\boldsymbol{x}=0$.因为$\boldsymbol{B}$的秩等于$r$，故其基础解系含$n-r$个向量，记为$\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_{n-r}$.
令$\boldsymbol{A}=(\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_{n-r})'$，这是个$(n-r)\times n$矩阵且秩为$n-r$.
显然齐次线性方程组$\boldsymbol{A}\boldsymbol{x}=0$的基础解系是$\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_r$，故其解空间就是$V_0$. □

Proposition 4.10

设$\boldsymbol{A}$是秩为r的$m\times n$矩阵，$\boldsymbol{\alpha}_1\sim\boldsymbol{\alpha}_{n-r}$与$\boldsymbol{\beta}_1\sim\boldsymbol{\beta}_{n-r}$是齐次线性方程组$\boldsymbol{A}\boldsymbol{x}=0$的两个基础解系

求证：必存在$n-r$阶可逆矩阵P $st.\left(\boldsymbol{\beta}_1,\cdots\boldsymbol{\beta}_{n-r}\right)=\left(\boldsymbol{\alpha}_1,\cdots\boldsymbol{\alpha}_{n-r}\right)\boldsymbol{P}$

Proof 设U_A为$\boldsymbol{A}\boldsymbol{x}=0$的解空间，则$\boldsymbol{\alpha}_1\sim\boldsymbol{\alpha}_{n-r}$与$\boldsymbol{\beta}_1\sim\boldsymbol{\beta}_{n-r}$是$U_A$的两组基
令P为过渡矩阵即可 □

Theorem 4.8.3 (矩阵方程有解判定)

$\boldsymbol{A}_{m\times n}$; $\boldsymbol{B}_{m\times p}$矩阵; $\boldsymbol{X}_{n\times p}$为未知矩阵

求证：$\boldsymbol{A}\boldsymbol{X}=\boldsymbol{B}$有解的充要条件是$r\left(\boldsymbol{A}|\boldsymbol{B}\right)=r\left(\boldsymbol{A}\right)$

Proof 设$A=\left(\alpha_1\cdots\alpha_{\boldsymbol{n}}\right)$ $\quad \boldsymbol{B}=\left(\beta_1\cdots\beta_p\right)$ $\quad X=\left(\boldsymbol{x}_1\cdots\boldsymbol{x}_p\right)$
设$r\left(A\right)=r$; 且$\alpha_{i_1},\cdots\alpha_{i_r}$是$A$列向量的极大无关组.注意到矩阵方程$\boldsymbol{A}\boldsymbol{X}=\boldsymbol{B}$有解当且仅当$p$个线性方程组$Ax_i=\beta_i\left(1\leqslant i\leqslant p\right)$有解
因此若矩阵方程有解，则每个β_i都是A的列向量的线性组合，从而是$\alpha_{i_1}\cdots\alpha_{i_r}$的线性组合
$\Rightarrow\alpha_{i_1}\cdots\alpha_{i_r}$是$\left(A|B\right)$列向量的极大无关组，$\Rightarrow r\left(A|B\right)=r$.
反之若$r\left(A|B\right)=r$.可知$\alpha_{i_1}\cdots\alpha_{i_r}$是$\left(A|B\right)$列向量的极大无关组于是每个$\beta_i$都是$A$的列向量的线性组合从而矩阵方程有解 □

Proposition 4.11

设$\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\cdots,\boldsymbol{\eta}_t$是齐次线性方程组$\boldsymbol{A}\boldsymbol{x}=\vec{0}$的一个基础解系

证明：与$\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\cdots,\boldsymbol{\eta}_t$等价的线性无关的向量组也是齐次线性方程组的一个基础解系

Proof 设$\boldsymbol{\gamma}_1,\cdots\boldsymbol{\gamma}_m$是等价的向量组因为等价有相同的秩且$\gamma$是线性无关的故$m=t$
故$\boldsymbol{\gamma}_1,\cdots\boldsymbol{\gamma}_m$也是解空间的一组基故也是基础解系

Theorem 4.8.4 (齐次线性方程组有非零公共解)

对于n元齐次线性方程组$(I)\boldsymbol{A}_{m_1\times n}\boldsymbol{X}=0$　与　$(II)\boldsymbol{B}_{m_2\times n}\boldsymbol{X}=0$

(1)(I)与(II)有非零公共解的充要条件是$r\begin{pmatrix}\boldsymbol{A}\\ \boldsymbol{B}\end{pmatrix}<n.$

(2)设$\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\cdots,\boldsymbol{\eta}_s(s=n-r(\boldsymbol{B}))$是$(II)$的基础解系，则$(I)$与$(II)$有非零公共解的充要条件是$\boldsymbol{A}\boldsymbol{\eta}_1,\boldsymbol{A}\eta_2,\cdots,\boldsymbol{A}\boldsymbol{\eta}_s$线性相关.

(3)设$\boldsymbol{\gamma}_1,\boldsymbol{\gamma}_2,\cdots,\boldsymbol{\gamma}_t(t=n-r(\boldsymbol{A}))$为$(I)$的基础解系，$\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\cdots,\boldsymbol{\eta}_s(s=n-r(\boldsymbol{B}))$是$(II)$的基础解系

则(I)与(II)有非零公共解的充要条件是$\gamma_1,\gamma_2,\cdots,\gamma_t,\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\cdots,\boldsymbol{\eta}_s$线性相关.

Theorem 4.8.5 (非齐次线性方程组有公共解)

对于n元非齐次线性方程组$(I)\boldsymbol{A}\boldsymbol{X}=\boldsymbol{b}$　与　$(II)\boldsymbol{B}\boldsymbol{X}=\boldsymbol{d}$.若$(I)$与$(II)$都有解$\Longrightarrow$

(1)(I)与(II)有公共解的充要条件是$r\begin{pmatrix}\boldsymbol{A}\\ \boldsymbol{B}\end{pmatrix}=r\begin{pmatrix}\boldsymbol{A}&\boldsymbol{b}\\ \boldsymbol{B}&\boldsymbol{d}\end{pmatrix}$

(2)若$r(\boldsymbol{B})=s$，且$\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\cdots,\boldsymbol{\eta}_{n-s+1}$是$(II)$的$n-s+1$个线性无关的解

则(I)与(II)有公共解的充要条件是$\boldsymbol{b}$是$\boldsymbol{A}\boldsymbol{\eta}_1,\boldsymbol{A}\boldsymbol{\eta}_2,\cdots,\boldsymbol{A}\boldsymbol{\eta}_{n-s+1}$的凸组合，

即存在数$k_1,k_2,\cdots,k_{n-s+1}$使得$\boldsymbol{b}=k_1\boldsymbol{A}\boldsymbol{\eta}_1+k_2\boldsymbol{A}\boldsymbol{\eta}_2+\cdots+k_{n-s+1}\boldsymbol{A}\boldsymbol{\eta}_{n-s+1}$，其中$k_1+k_2+\cdots+k_{n-s+1}=1.$

(3)若$r(\boldsymbol{A})=t,r(\boldsymbol{B})=s$，且$\gamma_1,\gamma_2,\cdots,\gamma_{n-t+1}$是$(I)$的$n-t+1$个线性无关的解，$\eta_1,\boldsymbol{\eta}_2,\cdots,\boldsymbol{\eta}_{n-s+1}$是$(II)$的$n-s+1$个线性无关的解

则(I)与(II)有公共解的充要条件是存在数$k_1,k_2,\cdots,k_{n-t+1}$与$l_1,l_2,\cdots,l_{n-s+1}$使得

$k_1\gamma_1+k_2\gamma_2+\cdots+k_{n-t+1}\boldsymbol{\gamma}_{n-t+1}-l_1\boldsymbol{\eta}_1-l_2\boldsymbol{\eta}_2-\cdots-l_{n-s+1}\boldsymbol{\eta}_{n-s+1}=0$其中$k_1+k_2+\cdots+k_{n-t+1}=1,l_1+l_2+\cdots+l_{n-s+1}=1.$

Theorem 4.8.6 (齐次线性方程组同解)

n元齐次线性方程组$(I)\boldsymbol{A}_{m_1\times n}\boldsymbol{X}=0$与$(II)\boldsymbol{B}_{m_2\times n}\boldsymbol{X}=0$同解的$\Longleftrightarrow r\begin{pmatrix}\boldsymbol{A}\\ \boldsymbol{B}\end{pmatrix}=r(\boldsymbol{A})=r(\boldsymbol{B}).$

Lemma 4.3

n元非齐次线性方程组（I）$\boldsymbol{A}\boldsymbol{X}=\boldsymbol{b}$与（$II$）$\boldsymbol{B}\boldsymbol{X}=\boldsymbol{d}$都有解，且（$I$）与（$II$）同解

则其导出组$AX=0$与$BX=0$同解

Theorem 4.8.7 (非齐次线性方程组同解)

n元非齐次线性方程组$(I)\boldsymbol{A}_{m_1\times n}\boldsymbol{X}=\boldsymbol{b}$　与　$(II)\boldsymbol{B}_{m_2\times n}\boldsymbol{X}=\boldsymbol{d}$同解$\Longleftrightarrow$

(1)$r(\boldsymbol{A})\neq r(\boldsymbol{A},\boldsymbol{b})$且$r(\boldsymbol{B})\neq r(\boldsymbol{B},\boldsymbol{d})$或 (2)$r\begin{pmatrix}\boldsymbol{A}&\boldsymbol{b}\\ \boldsymbol{B}&\boldsymbol{d}\end{pmatrix}=r\begin{pmatrix}\boldsymbol{A}\\ \boldsymbol{B}\end{pmatrix}=r(\boldsymbol{A})=r(\boldsymbol{A},\boldsymbol{b})=r(\boldsymbol{B})=r(\boldsymbol{B},\boldsymbol{d}).$

Problem 4.1

设$A = \begin{pmatrix} 1 & 1 & 2 & 1 \\ 1 & 2 & 3 & 3 \\ 2 & 3 & 5 & 4 \\ 3 & 5 & 8 & 7 \end{pmatrix}$ $B = \begin{pmatrix} 1 & 1 \\ 5 & -1 \\ 6 & 0 \\ 11 & -1 \end{pmatrix}$ X为4×2未知矩阵

求矩阵方程$AX = B$的解

Proof $\left(\begin{array}{cccc|cc} 1 & 1 & 2 & 1 & 1 & 1 \\ 1 & 2 & 3 & 3 & 5 & -1 \\ 2 & 3 & 5 & 4 & 6 & 0 \\ 3 & 5 & 8 & 7 & 11 & -1 \end{array}\right) \to \left(\begin{array}{cccc|cc} 1 & 0 & 1 & -1 & -3 & 3 \\ 0 & 1 & 1 & 2 & 4 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$

由$r(A|B) = r(A)$知道矩阵方程有解；那么$X = \begin{pmatrix} -1 & 1 \\ -1 & -2 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} k_{11} & k_{12} \\ k_{21} & k_{22} \end{pmatrix} + \begin{pmatrix} -3 & 3 \\ 4 & -2 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$ $k_{ij} \in \mathbb{K}$

Problem 4.2

求解 $Ax = 0$与$Bx = 0$两个齐次线性方程组公共解

求解 $Ax = \beta_1$ $Bx = \beta_2$的公共解

解

$method1$: 若一开始仅知道两个方程那么我们不防直接求解联立

$method2$: 若不知道方程形式但知道方程的基础解系, 不防求解空间的交即可

解

设$\boldsymbol{Ax} = \boldsymbol{\alpha}_1, \boldsymbol{Bx} = \boldsymbol{\beta}_2$是两个含$n$个未知数的非齐次线性方程组.

方程组$\boldsymbol{Ax} = \boldsymbol{\beta}_1$有特解$\boldsymbol{\gamma}$且$\boldsymbol{Ax} = 0$的基础解系为$\boldsymbol{\eta}_1, \cdots, \boldsymbol{\eta}_{n-r}$.方程组$\boldsymbol{Bx} = \boldsymbol{\beta}_2$有特解$\boldsymbol{\delta}$且$\boldsymbol{Bx} = 0$的基础解系为$\boldsymbol{\xi}_1, \cdots, \boldsymbol{\xi}_{n-s}$.

Proof 假设它们的公共解为$\boldsymbol{\zeta}$, 则$\boldsymbol{\zeta} = \boldsymbol{\gamma} + t_1\boldsymbol{\eta}_1 + \cdots + t_{n-r}\boldsymbol{\eta}_{n-r} = \boldsymbol{\delta} + (-u_1)\boldsymbol{\xi}_1 + \cdots + (-u_{n-s})\boldsymbol{\xi}_{n-s}$.

求公共解$\boldsymbol{\zeta}$等价于求解下列关于未定元$t_1, \cdots, t_{n-r}; u_1, \cdots, u_{n-s}$的线性方程组 : $t_1\boldsymbol{\eta}_1 + \cdots + t_{n-r}\boldsymbol{\eta}_{n-r} + u_1\boldsymbol{\xi}_1 + \cdots + u_{n-s}\boldsymbol{\xi}_{n-s} = \boldsymbol{\delta} - \boldsymbol{\gamma}$

Problem 4.3

设有非齐次线性方程组 (I) : $\begin{cases} 7x_1 - 6x_2 + 3x_3 = b, \\ 8x_1 - 9x_2 + ax_4 = 7. \end{cases}$ 又已知方程组 (II)的通解为 $(1, 1, 0, 0)' + t_1(1, 0, -1, 0)' + t_2(2, 3, 0, 1)'$.

若这两个方程组有无穷多组公共解, 求出a, b的值并求出公共解.

Proof 将 (II)的通解写为 $(1 + t_1 + 2t_2, 1 + 3t_2, -t_1, t_2)'$, 代入方程组 (I)后化简得到 $\begin{cases} 4t_1 - 4t_2 = b - 1, \\ 8t_1 + (a - 11)t_2 = 8. \end{cases}$

要使这两个方程组有无穷多组公共解, t_1, t_2必须有无穷多组解, 因此上面方程组的系数矩阵和增广矩阵的秩都应该等于1.

于是$a = 3, b = 5$.解出方程组得到$t_1 = t_2 + 1$.

所以方程组 $(I), (II)$的公共解为 $(1 + t_1 + 2t_2, 1 + 3t_2, -t_1, t_2)' = (2, 1, -1, 0)' + t_2(3, 3, -1, 1)'$, 其中$t_2$为任意数. □

Theorem 4.8.8

$\boldsymbol{A},\boldsymbol{B}$都是数域$\mathbb{F}$上的$m\times n$矩阵，证：方程组$\boldsymbol{Ax}=0,\boldsymbol{Bx}=0$同解的 $\Longleftrightarrow$ 存在可逆矩阵P，使得$B=PA$.
(A与B行向量组等价)

Proof 因为P是可逆矩阵，充分性是显然的.
现通过两种方法来证明必要性.
代数方法由条件可得方程组$Ax=0,Bx=0,\begin{pmatrix} A\\ B\end{pmatrix}x=0$都同解，从而有 $\mathrm{r}(\boldsymbol{A})=\mathrm{r}(\boldsymbol{B})=\mathrm{r}\begin{pmatrix}\boldsymbol{A}\\ \boldsymbol{B}\end{pmatrix}$.
注意到结论$B=PA$就是说$\boldsymbol{A},\boldsymbol{B}$可以通过初等行变换相互转化，因此在证明的过程中，对$\boldsymbol{A}$或$\boldsymbol{B}$实施初等行变换不影响结论的证明.
设$\boldsymbol{A}=\begin{pmatrix}\boldsymbol{\alpha}_1\\ \boldsymbol{\alpha}_2\\ \vdots\\ \boldsymbol{\alpha}_m\end{pmatrix},\boldsymbol{B}=\begin{pmatrix}\boldsymbol{\beta}_1\\ \boldsymbol{\beta}_2\\ \vdots\\ \boldsymbol{\beta}_m\end{pmatrix}$分别为$\boldsymbol{A},\boldsymbol{B}$的行分块.
不妨对$\boldsymbol{A},\boldsymbol{B}$都进行行对换，故可设$\alpha_1,\cdots,\boldsymbol{\alpha}_r$是$\boldsymbol{A}$的行向量的极大无关组，$\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_r$是$\boldsymbol{B}$的行向量的极大无关组.
由于 $\mathrm{r}\begin{pmatrix}\boldsymbol{A}\\ \boldsymbol{B}\end{pmatrix}=r$，故可知，$\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_r$和$\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_r$是向量组$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m,\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_m$的两组极大无关组.
(到此我们即可用过渡矩阵的知识来得到)
设$\boldsymbol{\beta}_i=\sum\limits_{j=1}^{r}c_{ij}\boldsymbol{\alpha}_j(1\le i\le r)$，则容易验证$r$阶方阵$\boldsymbol{C}=(c_{ij})$是非异阵.
设$\boldsymbol{\beta}_i-\boldsymbol{\alpha}_i=\sum\limits_{j=1}^{r}d_{ij}\boldsymbol{\alpha}_j(r+1\le i\le m),\boldsymbol{D}=(d_{ij})$是$(m-r)\times r$矩阵
则容易验证$\boldsymbol{P}=\begin{pmatrix}\boldsymbol{C} & \boldsymbol{O}\\ \boldsymbol{D} & \boldsymbol{I}_{m-r}\end{pmatrix}$是$m$阶非异阵，并且满足$\boldsymbol{B}=\boldsymbol{PA}$.

几何方法
将问题转化成几何的语言即为：
设V是$\mathbb{F}$上的n维线性空间，U是$\mathbb{F}$上的m维线性空间　$\boldsymbol{\varphi},\boldsymbol{\psi}:V\to U$是两个线性映射.
求证：若 $\mathrm{Ker}\varphi=\mathrm{Ker}\boldsymbol{\psi}$，则存在$U$上的自同构$\sigma$，使得$\psi=\sigma\varphi$.
设 $\mathrm{r}(\boldsymbol{\varphi})=r$，则 $\dim\mathrm{Ker}\boldsymbol{\varphi}=\dim\mathrm{Ker}\boldsymbol{\psi}=n-r$.
取 $\mathrm{Ker}\boldsymbol{\varphi}=\mathrm{Ker}\boldsymbol{\psi}$的一组基$e_{r+1},\cdots,e_n$，并将其扩张为$V$的一组基$e_1,\cdots,e_r,e_{r+1},\cdots,e_n$.
根据维数公式的证明方法之一可知，$\varphi(e_1),\cdots,\varphi(e_r)$是 $\mathrm{Im}\varphi$的一组基，故可将其扩张为U的一组基$\boldsymbol{\varphi}(\boldsymbol{e}_1),\cdots,\boldsymbol{\varphi}(\boldsymbol{e}_r),\boldsymbol{f}_{r+1},\cdots,\boldsymbol{f}_m$.
同理可知，$\boldsymbol{\psi}(\boldsymbol{e}_1),\cdots,\boldsymbol{\psi}(\boldsymbol{e}_r)$是 $\mathrm{Im}\boldsymbol{\psi}$的一组基，故可将其扩张为$U$的一组基$\boldsymbol{\psi}(\boldsymbol{e}_1),\cdots,\boldsymbol{\psi}(\boldsymbol{e}_r),\boldsymbol{g}_{r+1},\cdots,\boldsymbol{g}_m$.
定义U上的线性变换σ如下：$\boldsymbol{\sigma}(\boldsymbol{\varphi}(\boldsymbol{e}_i))=\boldsymbol{\psi}(\boldsymbol{e}_i),1\le i\le r;\quad \boldsymbol{\sigma}\left(\boldsymbol{f}_j\right)=\boldsymbol{g}_j,r+1\le j\le m$.
因为σ把U的一组基映射为U的另一组基，故σ是U的自同构.又对$r+1\le j\le n,\sigma\left(\varphi\left(e_j\right)\right)=0=\psi\left(e_j\right)$，故$\sigma\varphi=\psi$成立.

注 但是行向量组等价未必列向量组等价 (亦为：两个方程组同解推不出列向量组等价)
例如：$A=\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}\quad B=\begin{pmatrix}1 & 0\\ 1 & 0\end{pmatrix}$

4.9 向量组与矩阵重要结论

Proposition 4.12 (一组向量线性相关的线性组合充要表述)

$\alpha_1,\alpha_2,\cdots,\alpha_m$线性相关的充分必要条件是其中至少有一个向量可以表示为其余向量的线性组合.

Proof 设$\alpha_1,\alpha_2,\cdots,\alpha_m$线性相关,则存在不全为零的一组数$k_1,k_2,\cdots,k_m$,使$k_1\alpha_1+k_2\alpha_2+\cdots+k_m\alpha_m=0$其中有某个$k_j\neq 0$.
于是$\alpha_i=-k_i^{-1}k_1\alpha_1-\cdots-k_i^{-1}k_{i\cdots1}\alpha_{i-1}-k_i^{-1}k_{i+1}\alpha_{i+1}-\cdots-k_i^{-1}k_m\alpha_m$,即$\alpha_i$是其余$m-1$个向星的线性组合.
反过来,若$\alpha_i=b_1\alpha_1+\cdots+b_{i-1}\alpha_{i-1}+b_{i+1}\alpha_{i+1}+\cdots+b_m\alpha_m$,则
$b_1\alpha_1+\cdots+b_{i-1}\alpha_{i-1}+(-1)\alpha_i+b_{i+1}\alpha_{i+1}+\cdots+b_m\alpha_m=0$,即$\alpha_1,\alpha_2,\cdots,\alpha_m$线性相关.
证毕. □

Proposition 4.13 (表示法唯一与线性相关性的联系)

设$\alpha_1,\alpha_2,\cdots,\alpha_m;\beta$是 K 上线性空间$V$中的向量.已知$\beta$可表示为$\alpha_1,\alpha_2,\cdots,\alpha_m$的线性组合
即$\beta=k_1\alpha_1+k_2\alpha_2+\cdots+k_m\alpha_m$
则表示唯一的充分必要条件是向量$\alpha_1,\alpha_2,\cdots,\alpha_m$线性无关.

Proof $\Leftarrow$ 假定向量$\alpha_1,\alpha_2,\cdots,\alpha_m$线性无关且另外有一个表示：$\beta=b_1\alpha_1+b_2\alpha_2+\cdots+b_m\alpha_m$则将已知的两个表示式相减得到：
$(b_1-k_1)\alpha_1+(b_2-k_2)\alpha_2+\cdots+(b_m-k_m)\alpha_m=0$.因为$\alpha_1,\alpha_2,\cdots,\alpha_m$线性无关,故$b_1-k_1=0,b_2-k_2=0,\cdots,b_m-k_m=0$
即$b_1=k_1,b_2=k_2,\cdots,b_m=k_m$.也就是说$\beta$只能用唯一一种方式表示为$\alpha_1,\alpha_2,\cdots,\alpha_m$的线性组合.
$\Rightarrow$ 反证法,若向量$\alpha_1,\alpha_2,\cdots,\alpha_m$线性相关,即存在不全为零的数$c_1,c_2,\cdots,c_m$,使得$_{c_1}\alpha_1+c_2\alpha_2+\cdots+_{c_m}\alpha_m=0$
则除了已知的表示外,β还有另外$-..$个不同的表示：$\beta=(k_1+c_1)\alpha_1+(k_2+c_2)\alpha_2+\cdots+(k_m+c_m)\alpha_m$. □

Theorem 4.9.1 (一组线性无关向量与空间中另一向量关系)

设$\alpha_1,\alpha_2,\cdots,\alpha_m$是线性空间$V$中一组线性无关的向量,$\beta$是$V$中的向量.
求证：或者$\alpha_1,\alpha_2,\cdots,\alpha_m,\beta$线性无关,或者$\beta$是$\alpha_1,\alpha_2,\cdots,\alpha_m$的线性组合(等价于$\alpha_1,\alpha_2,\cdots,\alpha_m,\beta$线性相关)
上述结论等价于：若$\alpha_1,\alpha_2,\cdots,\alpha_m$线性无关且$\beta\notin L(\alpha_1,\alpha_2,\cdots,\alpha_m)$,则$\alpha_1,\alpha_2,\cdots,\alpha_m,\beta$线性无关.

Proof 若$\alpha_1,\alpha_2,\cdots,\alpha_m,\beta$线性无关,则结论得证.
若$\alpha_1,\alpha_2,\cdots,\alpha_m,\beta$线性相关,则存在不全为零的数$c_1,c_2,\cdots,c_m,d$,使得$c_1\alpha_1+c_2\alpha_2+\cdots+c_m\alpha_m+d\beta=0$.
若$d=0$,则$c_1,c_2,\cdots,c_m$不全为零且$c_1\alpha_1+c_2\alpha_2+\cdots+c_m\alpha_m=0$,这与$\alpha_1,\alpha_2,\cdots,\alpha_m$线性无关矛盾.
因此$d\neq 0$,从而$\beta=-\dfrac{c_1}{d}\alpha_1-\dfrac{c_2}{d}\alpha_2-\cdots-\dfrac{c_m}{d}\alpha_m$,即$\beta$是$\alpha_1,\alpha_2,\cdots,\alpha_m$的线性组合. □

Theorem 4.9.2 (维数缩短向量组)

设$\{\alpha_i=(a_{i1},a_{i2},\cdots,a_{in}),1\leq i\leq m\}$是一组$n$维行向量,$1\leq j_1<j_2<\cdots<j_t\leq n$是给定的$t(t<n)$个指标.
定义$\widetilde{\alpha}_i=\left(a_{ij_1},a_{ij_2},\cdots,a_{ij_t}\right)$,称$\widetilde{\alpha}_i$为$\alpha_i$的$t$维缩短向量.
(1) 若$\alpha_1,\alpha_2,\cdots,\alpha_m$线性相关,则$\widetilde{\alpha}_1,\widetilde{\alpha}_2,\cdots,\widetilde{\alpha}_m$也线性相关.
(2) 如果$\alpha_1,\cdots\alpha_m$线性无关,那么把每个α_i内添上l个分量(所添分量的位置对于每个向量都一样)得到的延伸组也线性无关。
(3) 设$\alpha_0=(t_1,t_2\cdots t_n)$是$\alpha_1,\alpha_2,\cdots,\alpha_m$的线性组合,则$\widetilde{\alpha_0}$也是$\widetilde{\alpha}_1,\widetilde{\alpha}_2,\cdots,\widetilde{\alpha}_m$的线性组合

Proof (1) 由假设存在不全为零的数$c_1,c_2,\cdots,c_m$,使得$0=c_1\alpha_1+c_2\alpha_2+\cdots+c_m\alpha_m=\left(\sum\limits_{i=1}^m c_ia_{i1},\sum\limits_{i=1}^m c_ia_{i2},\cdots,\sum\limits_{i=1}^m c_ia_{in}\right)$.
两边同时取t维缩短向量,可得$0=\left(\sum\limits_{i=1}^m c_ia_{ij_1},\sum\limits_{i=1}^m c_ia_{ij_2},\cdots,\sum\limits_{i=1}^m c_ia_{ij_t}\right)=c_1\widetilde{\alpha}_1+c_2\widetilde{\alpha}_2+\cdots+c_m\widetilde{\alpha}_m$,从而结论得证.

(2) 反证法若加了分量的向量组线性相关, 那么由 (1) 将所加的分量再去掉得到的$\alpha_1 \cdots \alpha_m$应线性相关, 矛盾

(3) 同理 (1) 的过程即可

Proposition 4.14 (个数减少部分向量组)

设向量组$\alpha_1 \cdots \alpha_n$线性无关, 那么其部分组$\alpha_{i_1} \cdots \alpha_{i_m}$也线性无关.

逆否形式：部分组线性相关那么整个向量组线性相关.

Proof 反证法, 若部分组线性相关, 不防设为$k_1\alpha_1 + k_2\alpha_2 + \cdots + k_m\alpha_m = 0\ (k_1 \sim k_m$不全为$0)$

那么对于$\alpha_1 \cdots \alpha_n$便有$k_1\alpha_1 + k_2\alpha_2 + \cdots + k_m\alpha_m + 0 \cdot \alpha_{m+1} + \cdots + 0 \cdot \alpha_n = 0$线性相关矛盾 □

Proposition 4.15

设$\boldsymbol{A}$是$n \times m$矩阵, $\boldsymbol{B}$是$m \times n$矩阵.若$\boldsymbol{AB} = \boldsymbol{C}_{n\times n}$且$\boldsymbol{C}$的秩为$n$

则$\boldsymbol{B}$的列向量组线性无关, $\boldsymbol{A}$的行向量组线性无关

Proof 设$\boldsymbol{B} = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_n)$为列分块, 则$\boldsymbol{AB} = (\boldsymbol{A\beta}_1, \boldsymbol{A\beta}_2, \cdots, \boldsymbol{A\beta}_n)$.由$\boldsymbol{AB} = \boldsymbol{C}$可得$\boldsymbol{A\beta}_i = \boldsymbol{c}_i (1 \le i \le n)$

其中$\boldsymbol{c}_i$是n维列向量.设$k_1\boldsymbol{\beta}_1 + k_2\boldsymbol{\beta}_2 + \cdots + k_n\boldsymbol{\beta}_n = 0$.上式两边同时左乘$\boldsymbol{A}$, 可得

$0 = k_1\boldsymbol{A\beta}_1 + k_2\boldsymbol{A\beta}_2 + \cdots + k_n\boldsymbol{A\beta}_n = k_1c_1 + k_2c_2 + \cdots + k_nc_n = (c_1, c_2, \cdots, c_n)'$

因此$c_1 = c_2 = \cdots = c_n = 0$, 即$\boldsymbol{B}$的$n$个列向量$\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_n$线性无关. □

Proposition 4.16

设$\boldsymbol{A} = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n)$是一个$m \times n$矩阵, $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n$是列向量.

$\boldsymbol{P}$是一个m阶可逆矩阵, $\boldsymbol{B} = \boldsymbol{PA} = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_n)$, 其中$\boldsymbol{\beta}_j = \boldsymbol{P\alpha}_j (j = 1, 2, \cdots, n)$.

若$\boldsymbol{\alpha}_{i_1}, \boldsymbol{\alpha}_{i_2}, \cdots, \boldsymbol{\alpha}_{i_k}$是$\boldsymbol{A}$的列向量的极大无关组, 则$\boldsymbol{\beta}_{i_1}, \boldsymbol{\beta}_{i_2}, \cdots, \boldsymbol{\beta}_{i_k}$是$\boldsymbol{B}$的列向量的极大无关组.

Proof 先证明向量组$\boldsymbol{\beta}_{i_1}, \boldsymbol{\beta}_{i_2}, \cdots, \boldsymbol{\beta}_{i_k}$线性无关.设$c_1\boldsymbol{\beta}_{i_1} + c_2\boldsymbol{\beta}_{i_2} + \cdots + c_k\boldsymbol{\beta}_{i_k} = 0$, 即$c_1\boldsymbol{P\alpha}_{i_1} + c_2\boldsymbol{P\alpha}_{i_2} + \cdots + c_k\boldsymbol{P\alpha}_{i_k} = 0$.

已知$\boldsymbol{P}$是可逆矩阵, 因此$c_1\boldsymbol{\alpha}_{i_1} + c_2\boldsymbol{\alpha}_{i_2} + \cdots + c_k\boldsymbol{\alpha}_{i_k} = 0$.而向量组$\boldsymbol{\alpha}_{i_1}, \boldsymbol{\alpha}_{i_2}, \cdots, \boldsymbol{\alpha}_{i_k}$线性无关, 故$c_1 = c_2 = \cdots = c_k = 0$.

这证明了向量组$\boldsymbol{\beta}_{i_1}, \boldsymbol{\beta}_{i_2}, \cdots, \boldsymbol{\beta}_{i_k}$线性无关.要证这是$\boldsymbol{B}$列向量的极大无关组, 只需证明$\boldsymbol{B}$的任意一个列向量都是这些向量的线性组合.

设$\boldsymbol{\beta}_j$是$\boldsymbol{B}$的任意一个列向量, 则$\boldsymbol{\beta}_j = \boldsymbol{P\alpha}_j$.因为$\boldsymbol{\alpha}_{i_1}, \boldsymbol{\alpha}_{i_2}, \cdots, \boldsymbol{\alpha}_{i_k}$是$\boldsymbol{A}$的列向量的极大无关组

故$\boldsymbol{\alpha}_j$可用$\boldsymbol{\alpha}_{i_1}, \boldsymbol{\alpha}_{i_2}, \cdots, \boldsymbol{\alpha}_{i_k}$线性表示, 不妨设$\boldsymbol{\alpha}_j = b_1\boldsymbol{\alpha}_{i_1} + b_2\boldsymbol{\alpha}_{i_2} + \cdots + b_k\boldsymbol{\alpha}_{i_k}$

则$\boldsymbol{P\alpha}_j = b_1\boldsymbol{P\alpha}_{i_1} + b_2\boldsymbol{P\alpha}_{i_2} + \cdots + b_k\boldsymbol{P\alpha}_{i_k}$, 即$\boldsymbol{\beta}_j = b_1\boldsymbol{\beta}_{i_1} + b_2\boldsymbol{\beta}_{i_2} + \cdots + b_k\boldsymbol{\beta}_{i_k}$. □

Theorem 4.9.3 (向量组等价充要条件)

设有两个向量组$A = \{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m\}$和$B = \{\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_n\}$.

求证：它们等价的充要条件是它们的秩相等且其中一组向量可以用另外一组向量线性表示.

Proof 根据向量组等价的定义可知必要性, 只需证明充分性.

假设向量组A可用向量组B线性表示, 现证明向量组B也可用向量组A线性表示.

不失一般性, 设$\alpha_1, \alpha_2, \cdots, \alpha_r$是向量组$A$的极大无关组, $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_r$是向量组$B$的极大无关组.

考虑向量组$C = \{\alpha_1, \alpha_2, \cdots, \alpha_r; \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_r\}$.因为$\alpha_1, \alpha_2, \cdots, \alpha_r$可用$\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_r$线性表出

故$\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_r$是向量组$C$的极大无关组, 从而向量组$C$的秩等于$r$.

注意到$\alpha_1, \alpha_2, \cdots, \alpha_r$线性无关, 故可知, $\alpha_1, \alpha_2, \cdots, \alpha_r$也是向量组$C$的极大无关组, 从而$\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_r$可用$\alpha_1, \alpha_2, \cdots, \alpha_r$线性表示.

因此向量组B也可用向量组A线性表示, 故A和B等价. □

Theorem 4.9.4 (线性空间的维数承接子关系)

设$\mathbb{K}_1,\mathbb{K}_2,\mathbb{K}_3$是数域且$\mathbb{K}_1\subseteq\mathbb{K}_2\subseteq\mathbb{K}_3$, 若将$\mathbb{K}_2$看成是$\mathbb{K}_1$上的线性空间, 其维数为$m$, 将$\mathbb{K}_3$看成是$\mathbb{K}_2$上的线性空间, 其维数为$n$ 如将$\mathbb{K}_3$看成是$\mathbb{K}_1$上的线性空间, 则其维数为mn.

设$\mathbb{F}\subseteq\mathbb{K}$为数域, $\mathbb{K}$作为$\mathbb{F}$上的线性空间, 一组基为 $\{\alpha_1\cdots\alpha_m\}$;
设V为$\mathbb{K}$上的n维线性空间, 一组基为 $\{\boldsymbol{e}_1\cdots\boldsymbol{e}_n\}$
完全同类似上面的证明过程可知, V是$\mathbb{F}$上的mn维线性空间

Proof $\mathbb{K}_2$作为$\mathbb{K}_1$上的线性空间, 取其一组基为 $\{\alpha_1,\alpha_2,\cdots,\alpha_m\}$; $\mathbb{K}_3$作为$\mathbb{K}_2$上的线性空间, 取其一组基为 $\{\beta_1,\beta_2,\cdots,\beta_n\}$.
注意到α_i,β_j都是数, 现在我们断言: $\mathbb{K}_3$作为$\mathbb{K}_1$上的线性空间, $\left\{\alpha_i\beta_j\mid 1\le i\le m,1\le j\le n\right\}$恰为其一组基.
一方面, 对$\mathbb{K}_3$中任一数a, 存在$\mathbb{K}_2$中的数$b_1,b_2,\cdots,b_n$, 使得$a=b_1\beta_1+b_2\beta_2+\cdots+b_n\beta_n$.
又对$b_j\in\mathbb{K}_2$, 存在$\mathbb{K}_1$中的数$c_{1j},c_{2j},\cdots,c_{mj}$, 使得$b_j=c_{1j}\alpha_1+c_{2j}\alpha_2+\cdots+c_{mj}\alpha_m, j=1,2,\cdots,n$.
将上述两式进行整理, 可得$a=\sum\limits_{j=1}^{n}b_j\beta_j=\sum\limits_{j=1}^{n}\left(\sum\limits_{i=1}^{m}c_{ij}\alpha_i\right)\beta_j=\sum\limits_{j=1}^{n}\sum\limits_{i=1}^{m}c_{ij}\alpha_i\beta_j$,
即$\mathbb{K}_3$中任一数均可由$\left\{\alpha_i\beta_j\mid 1\le i\le m,1\le j\le n\right\}$线性表示.

另一方面, 设有$\mathbb{K}_1$中的数$k_{ij}(1\le i\le m,1\le j\le n)$, 使得$\sum\limits_{j=1}^{n}\sum\limits_{i=1}^{m}k_{ij}\alpha_i\beta_j=0$, 则经过变形可得$\sum\limits_{j=1}^{n}\left(\sum\limits_{i=1}^{m}k_{ij}\alpha_i\right)\beta_j=0$.
注意到$\sum\limits_{i=1}^{m}k_{ij}\alpha_i\in\mathbb{K}_2$且$\beta_1,\beta_2,\cdots,\beta_n$是$\mathbb{K}_3/\mathbb{K}_2$的一组基, 故有$\sum\limits_{i=1}^{m}k_{ij}\alpha_i=0(1\le j\le n)$.
又因为$\{\alpha_1,\alpha_2,\cdots,\alpha_m\}$是$\mathbb{K}_2/\mathbb{K}_1$的一组基, 故有$k_{ij}=0(1\le i\le m,1\le j\le n)$, 即$\left\{\alpha_i\beta_j\mid 1\le i\le m,1\le j\le n\right\}$是$\mathbb{K}_1-$线性无关的.
综上所述, $\left\{\alpha_i\beta_j\mid 1\le i\le m,1\le j\le n\right\}$是$\mathbb{K}_3/\mathbb{K}_1$的一组基, 特别地, $\dim_{\mathbb{K}_1}\mathbb{K}_3=mn=\dim_{\mathbb{K}_1}\mathbb{K}_2\cdot\dim_{\mathbb{K}_2}\mathbb{K}_3$. □

Theorem 4.9.5

设$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$是一组线性无关的向量, 若向量组$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_k$可用$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$线性表示如下:

$$\begin{cases}\boldsymbol{\beta}_1=a_{11}\boldsymbol{\alpha}_1+a_{12}\boldsymbol{\alpha}_2+\cdots+a_{1m}\boldsymbol{\alpha}_m,\\ \boldsymbol{\beta}_2=a_{21}\boldsymbol{\alpha}_1+a_{22}\boldsymbol{\alpha}_2+\cdots+a_{2m}\boldsymbol{\alpha}_m,\\ \qquad\cdots\cdots\cdots\cdots\\ \boldsymbol{\beta}_k=a_{k1}\boldsymbol{\alpha}_1+a_{k2}\boldsymbol{\alpha}_2+\cdots+a_{km}\boldsymbol{\alpha}_m.\end{cases}\quad \text{记表示矩阵}\boldsymbol{A}=\left(a_{ij}\right)_{k\times m}.$$

求证: 向量组$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_k$的秩等于$\mathrm{r}(\boldsymbol{A})$.

Proof

设$\mathrm{r}(\boldsymbol{A})=r$.记$\boldsymbol{A}$的$k$个行向量为$\boldsymbol{\gamma}_1,\boldsymbol{\gamma}_2,\cdots,\boldsymbol{\gamma}_k$.不失一般, 可假定$\boldsymbol{A}$的前$r$个行向量线性无关, 其余行向量均可用前$r$个行向量线性表示.
若$\boldsymbol{\gamma}_i=c_1\boldsymbol{\gamma}_1+c_2\boldsymbol{\gamma}_2+\cdots+c_r\boldsymbol{\gamma}_r$则经过简单计算可得$c_1\boldsymbol{\beta}_1+c_2\boldsymbol{\beta}_2+\cdots+c_r\boldsymbol{\beta}_r=\boldsymbol{\beta}_i$.
另一方面, 若$c_1\boldsymbol{\beta}_1+c_2\boldsymbol{\beta}_2+\cdots+c_r\boldsymbol{\beta}_r=0$, 则$c_1\left(a_{11}\boldsymbol{\alpha}_1+\cdots+a_{1m}\boldsymbol{\alpha}_m\right)+\cdots+c_r\left(a_{r1}\boldsymbol{\alpha}_1+\cdots+a_{rm}\boldsymbol{\alpha}_m\right)=0$,
即$\left(c_1a_{11}+\cdots+c_ra_{r1}\right)\boldsymbol{\alpha}_1+\cdots+\left(c_1a_{1m}+\cdots+c_ra_{rm}\right)\boldsymbol{\alpha}_m=0$.

因为$\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_m$线性无关, 得$\begin{cases}a_{11}c_1+a_{21}c_2+\cdots+a_{r1}c_r=0,\\ a_{12}c_1+a_{22}c_2+\cdots+a_{r2}c_r=0,\\ \qquad\cdots\cdots\cdots\cdots\\ a_{1m}c_1+a_{2m}c_2+\cdots+a_{rm}c_r=0.\end{cases}$

将上述方程组看成是未知数c_i的齐次线性方程组, 其系数矩阵的秩为r, 未知数个数也是r, 因此只有唯一组解, 即零解.
这表明$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_r$是向量组$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_k$的极大无关组, 因此向量组$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_k$的秩等于$r$.

证法2　令V是由$\alpha_1,\alpha_2,\cdots,\alpha_m$生成的向量空间.因为$\alpha_1,\alpha_2,\cdots,\alpha_m$线性无关,故它们组成$V$的一组基,$V$的维数等于$m$.
注意到β_i在这组基下的坐标向量为$(a_{i1},a_{i2},\cdots,a_{im})'$,则列向量组成的矩阵就是$A'$,则向量组$\beta_1,\beta_2,\cdots,\beta_k$的秩等于$\mathrm{r}(A')=\mathrm{r}(A)$.

Theorem 4.9.6 (向量组的表示秩关系)

设$\alpha_1,\alpha_2,\cdots,\alpha_k,\beta_1,\beta_2,\cdots,\beta_m$是向量空间$V$中的向量且

$$\begin{cases} \beta_1=c_{11}\alpha_1+c_{12}\alpha_2+\cdots+c_{1k}\alpha_k,\\ \beta_2=c_{21}\alpha_1+c_{22}\alpha_2+\cdots+c_{2k}\alpha_k,\\ \qquad\cdots\cdots\cdots\cdots\cdots\\ \beta_m=c_{m1}\alpha_1+c_{m2}\alpha_2+\cdots+c_{mk}\alpha_k. \end{cases}$$

记上述表示式中的系数矩阵为$C=\left(c_{ij}\right)_{m\times k}$
(1)若$\mathrm{r}(C)=k$,则这两组向量等价.
(2)若$\mathrm{r}(C)=r$,则向量组$\beta_1,\beta_2,\cdots,\beta_m$的秩不超过$r$.

Proof (1)在V中取定一组基$e_1,e_2,\cdots,e_n$,假定在这组基下α_i的坐标向量是$\widetilde{\alpha}_i(i=1,\cdots,k)$,$\beta_j$的坐标向量是$\widetilde{\beta}_j(j=1,\cdots,m)$,则

$$\begin{cases} \widetilde{\beta}_1=c_{11}\widetilde{\alpha}_1+c_{12}\widetilde{\alpha}_2+\cdots+c_{1k}\widetilde{\alpha}_k,\\ \widetilde{\beta}_2=c_{21}\widetilde{\alpha}_1+c_{22}\widetilde{\alpha}_2+\cdots+c_{2k}\widetilde{\alpha}_k,\\ \qquad\cdots\cdots\cdots\cdots\cdots\\ \widetilde{\beta}_m=c_{m1}\widetilde{\alpha}_1+c_{m2}\widetilde{\alpha}_2+\cdots+c_{mk}\widetilde{\alpha}_k, \end{cases}$$

写成矩阵形式为$\left(\widetilde{\beta}_1,\widetilde{\beta}_2,\cdots,\widetilde{\beta}_m\right)=(\widetilde{\alpha}_1,\widetilde{\alpha}_2,\cdots,\widetilde{\alpha}_k)C'$.
因为C'是一个行满秩$k\times m$矩阵,根据定理,存在列满秩$m\times k$矩阵T,使$C'T=I_k$,于是$\left(\widetilde{\beta}_1,\widetilde{\beta}_2,\cdots,\widetilde{\beta}_m\right)T=(\widetilde{\alpha}_1,\widetilde{\alpha}_2,\cdots,\widetilde{\alpha}_k)$.
这表明$\alpha_1,\alpha_2,\cdots,\alpha_k$可以用$\beta_1,\beta_2,\cdots,\beta_m$来线性表示.于是这两组向量等价.
(2)类似于(1)的讨论,可以用两个矩阵积的秩不超过每个矩阵的秩得到. □

注 这是对两个向量组表示的秩关系的最准确表示,包含了以下两点
(1) $\{\alpha_1\sim\alpha_n\}$能够被$\{\beta_1\sim\beta_m\}$表示则$r(\alpha)\leqslant r(\beta)$
(2) 上一个定理

4.10 秩不等式

Proposition 4.17

(1) $\boldsymbol{C}=\begin{pmatrix}\boldsymbol{A} & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{B}\end{pmatrix}$ 则 $r(\boldsymbol{C})=r(\boldsymbol{A})+r(\boldsymbol{B})$

(2) $r(\boldsymbol{AB})\leqslant\min\{r(\boldsymbol{A}),r(\boldsymbol{B})\}$

(3) $r(k\boldsymbol{A})=r(\boldsymbol{A})\ (k\neq 0)\quad A_{m\times n}$

(4) $r(\boldsymbol{A}\pm\boldsymbol{B})\leqslant r(\boldsymbol{A})+r(\boldsymbol{B})$

(5) $r\begin{pmatrix}\boldsymbol{A} & \boldsymbol{C}\\ \boldsymbol{O} & \boldsymbol{B}\end{pmatrix}\geqslant r(\boldsymbol{A})+r(\boldsymbol{B})\qquad r\begin{pmatrix}\boldsymbol{A} & \boldsymbol{O}\\ \boldsymbol{D} & \boldsymbol{B}\end{pmatrix}\geqslant r(\boldsymbol{A})+r(\boldsymbol{B})$

(6) $\max\{r(\boldsymbol{A}),r(\boldsymbol{B})\}\leqslant r\begin{pmatrix}\boldsymbol{A} & \boldsymbol{B}\end{pmatrix}\leqslant r(\boldsymbol{A})+r(\boldsymbol{B})\qquad\max\{r(\boldsymbol{A}),r(\boldsymbol{B})\}\leqslant r\begin{pmatrix}\boldsymbol{A}\\ \boldsymbol{B}\end{pmatrix}\leqslant r(\boldsymbol{A})+r(\boldsymbol{B})$

(7) $r(\boldsymbol{A}-\boldsymbol{B})\geqslant|r(\boldsymbol{A})-r(\boldsymbol{B})|$

(8) $R(A\pm B)\leqslant R(A,B)\leqslant R(A)+R(B)$

Proof (1) 设 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的秩为 r_1,r_2 由相抵标准型知道存在可逆阵 $\boldsymbol{P}_1,\boldsymbol{P}_2,\boldsymbol{Q}_1,\boldsymbol{Q}_2$

$$\boldsymbol{P}_1\boldsymbol{A}\boldsymbol{Q}_1=\begin{pmatrix}\boldsymbol{I}_{r_1} & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{pmatrix}\qquad \boldsymbol{P}_2\boldsymbol{A}\boldsymbol{Q}_2=\begin{pmatrix}\boldsymbol{I}_{r_2} & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{pmatrix}$$

那么 $\begin{pmatrix}\boldsymbol{P}_1 & 0\\ 0 & \boldsymbol{P}_2\end{pmatrix}\begin{pmatrix}\boldsymbol{A} & 0\\ 0 & \boldsymbol{B}\end{pmatrix}\begin{pmatrix}\boldsymbol{Q}_1 & 0\\ 0 & \boldsymbol{Q}_2\end{pmatrix}=\begin{pmatrix}\boldsymbol{P}_1\boldsymbol{A}\boldsymbol{Q}_1 & 0\\ 0 & \boldsymbol{P}_2\boldsymbol{B}\boldsymbol{Q}_2\end{pmatrix}=\begin{pmatrix}\boldsymbol{I}_{r_1} & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & \boldsymbol{I}_{r_2} & 0\\ 0 & 0 & 0 & 0\end{pmatrix}$. 因此 $r(C)=r_1+r_2$. 证毕.

(2) 设 A 是 $m\times n$ 矩阵, B 是 $n\times s$ 矩阵. 将矩阵 B 按列分块, $B=(\beta_1,\beta_2,\cdots,\beta_s)$, 则 $AB=(A\beta_1,A\beta_2,\cdots,A\beta_s)$.
若 B 列向量的极大无关组为 $\{\beta_{i_1},\beta_{i_2},\cdots,\beta_{i_r}\}$, 则 B 的任一列向量 β_j 均可用 $\{\beta_{i_1},\beta_{i_2},\cdots,\beta_{i_r}\}$ 线性表示.
于是任一 $A\beta_j$ 也可用 $\{A\beta_{i_1},A\beta_{i_2},\cdots,A\beta_{i_r}\}$ 来线性表示. 因此向量组 $\{A\beta_1,A\beta_2,\cdots,A\beta_s\}$ 的秩不超过 r, 即 $r(AB)\leqslant r(B)$.
同理, 对矩阵 A 用行分块的方法可以 $r(AB)\leqslant r(A)$. 证毕.

(3) $k\boldsymbol{A}=P_1(k)P_2(k)\cdots P_m(k)A$ 故 $r(k\boldsymbol{A})=r(A)$

(5) 法1我们只证明第一个不等式, 第二个不等式同理可证. 采用与 (1) 相同的证法和记号, 可得

$$\begin{pmatrix}P_1 & O\\ O & P_2\end{pmatrix}\begin{pmatrix}A & C\\ O & B\end{pmatrix}\begin{pmatrix}Q_1 & O\\ O & Q_2\end{pmatrix}=\begin{pmatrix}P_1AQ_1 & P_1CQ_2\\ O & P_2BQ_2\end{pmatrix}=\begin{pmatrix}I_{r_1} & O & C_{11} & C_{12}\\ O & O & C_{21} & C_{22}\\ O & O & I_{r_2} & O\\ O & O & O & O\end{pmatrix}$$

在上面的分块矩阵中实施第三类分块初等变换

用 $\boldsymbol{I}_{r_1}$ 消去同行的矩阵; 用 $\boldsymbol{I}_{r_2}$ 消去同列的矩阵, 再将 $\boldsymbol{C}_{22}$ 对换到第 (2,2) 位置:

$$\begin{pmatrix}I_{r_1} & O & C_{11} & C_{12}\\ \boldsymbol{O} & O & C_{21} & C_{22}\\ \boldsymbol{O} & O & I_{r_2} & O\\ \boldsymbol{O} & O & O & O\end{pmatrix}\to\begin{pmatrix}I_{r_1} & O & O & O\\ O & O & O & C_{22}\\ O & O & I_{r_2} & O\\ \boldsymbol{O} & O & O & O\end{pmatrix}\to\begin{pmatrix}I_{r_1} & O & O & O\\ O & C_{22} & O & O\\ O & O & I_{r_2} & O\\ O & O & O & O\end{pmatrix}$$

再由 (1) 可得

$$r\begin{pmatrix}A & C\\ O & B\end{pmatrix}=r(I_{r_1})+r(C_{22})+r(I_{r_2})\geqslant r_1+r_2=r(A)+r(B)$$

证法2我们也可用子式法来证明.设 $\mathrm{r}\begin{pmatrix} A & O \\ O & B \end{pmatrix} = r$, 则可知, $\begin{pmatrix} A & O \\ O & B \end{pmatrix}$ 有一个r阶子式不为零

不妨设为 $\begin{vmatrix} A_1 & O \\ O & B_1 \end{vmatrix}$, 其中$\boldsymbol{A}_1, \boldsymbol{B}_1$分别是$A, B$的子阵.注意$A_1$或$B_1$允许是零阶矩阵, 这对应于该子式完全包含在$B$或$\boldsymbol{A}$中

但若$\boldsymbol{A}_1, \boldsymbol{B}_1$的阶数都大于零, 则通过该子式非零容易验证$\boldsymbol{A}_1, \boldsymbol{B}_1$都是方阵.

设在矩阵 $\begin{pmatrix} A & C \\ O & B \end{pmatrix}$ 中对应的r阶子式是 $\begin{vmatrix} A_1 & C_1 \\ O & B_1 \end{vmatrix}$, 则得 $\begin{vmatrix} A_1 & C_1 \\ O & B_1 \end{vmatrix} = |A_1||B_1| = \begin{vmatrix} A_1 & O \\ O & B_1 \end{vmatrix} \neq 0$

那么可得 $\mathrm{r}\begin{pmatrix} A & C \\ O & B \end{pmatrix} \geq r = \mathrm{r}\begin{pmatrix} A & O \\ O & B \end{pmatrix} = \mathrm{r}(A) + \mathrm{r}(B)$

证法3

设$\boldsymbol{A} = (\alpha_1, \alpha_2, \cdots, \alpha_n)$ 是$\boldsymbol{A}$的列分块, $\alpha_{i_1}, \alpha_{i_2}, \cdots, \alpha_{i_r}$是$\boldsymbol{A}$的列向量的极大无关组;

设$\boldsymbol{B} = (\beta_1, \beta_2, \cdots, \beta_l)$, $\boldsymbol{C} = (\gamma_1, \gamma_2, \cdots, \gamma_l)$ 是$\boldsymbol{B}, \boldsymbol{C}$的列分块, $\beta_{j_1}, \beta_{j_2}, \cdots, \beta_{j_s}$是$\boldsymbol{B}$的列向量的极大无关组, 则 $\mathrm{r}(\boldsymbol{A}) = r$且 $\mathrm{r}(\boldsymbol{B}) = s$.

我们接下来证明：作为 $\begin{pmatrix} A & C \\ O & B \end{pmatrix}$ 的列向量, $\begin{pmatrix} \alpha_{i_1} \\ 0 \end{pmatrix}, \cdots, \begin{pmatrix} \alpha_{i_r} \\ 0 \end{pmatrix}, \begin{pmatrix} \gamma_{j_1} \\ \beta_{j_1} \end{pmatrix}, \cdots, \begin{pmatrix} \gamma_{j_s} \\ \beta_{j_s} \end{pmatrix}$ 线性无关.

设$c_1 \begin{pmatrix} \alpha_{i_1} \\ 0 \end{pmatrix} + \cdots + c_r \begin{pmatrix} \alpha_{i_r} \\ 0 \end{pmatrix} + d_1 \begin{pmatrix} \gamma_{j_1} \\ \beta_{j_1} \end{pmatrix} + \cdots + d_s \begin{pmatrix} \gamma_{j_s} \\ \beta_{j_s} \end{pmatrix} = 0$

即$c_1\alpha_{i_1} + \cdots + c_r\alpha_{i_r} + d_1\gamma_{j_1} + \cdots + d_s\gamma_{j_s} = 0, d_1\beta_{j_1} + \cdots + d_s\beta_{j_s} = 0$.

由上面的假设即得$c_1 = \cdots = c_r = d_1 = \cdots = d_s = 0$, 于是上述结论得证.

因为 $\begin{pmatrix} A & C \\ O & B \end{pmatrix}$ 的列向量中有$r + s$个线性无关, 故 $\mathrm{r}\begin{pmatrix} A & C \\ O & B \end{pmatrix} \geq r + s = \mathrm{r}(\boldsymbol{A}) + \mathrm{r}(\boldsymbol{B})$.

(6) 注意到 $(I; I)\begin{pmatrix} A & O \\ O & B \end{pmatrix} = (A : B)$, $\begin{pmatrix} A & O \\ O & B \end{pmatrix}\begin{pmatrix} I \\ I \end{pmatrix} = \begin{pmatrix} A \\ B \end{pmatrix}$, 故由 (2) 即得结论, 另外我们可以转化到向量组的秩考虑, 利用推论 2.4.

(4) 注意到 $(A : B)\begin{pmatrix} I \\ I \end{pmatrix} = A + B$, $(A : B)\begin{pmatrix} I \\ -I \end{pmatrix} = A - B$, 故由 (2) 即得结论.

(8) 设A, B的列向量组分别为$\alpha_1 \sim \alpha_s; \beta_1 \sim \beta_s$; 那么$A \pm B$的列向量组为$\alpha_1 \pm \beta_1 \sim \alpha_s \pm \beta_s$.

其可以被$\alpha_1 \sim \alpha_s, \beta_1 \sim \beta_s$表示 $\Rightarrow R(A+B) \leqslant R(A, B) \leqslant R(A) + R(B)$

(7) 由于 $\mathrm{r}(A - B) = \mathrm{r}(B - A)$, 故不妨设 $\mathrm{r}(A) \geq \mathrm{r}(B)$, 则由 (4) 可得

$\mathrm{r}(A - B) + \mathrm{r}(B) \geq \mathrm{r}(A - B + B) = \mathrm{r}(A)$, 即 $\mathrm{r}(A - B) \geq \mathrm{r}(A) - \mathrm{r}(B)$.

Theorem 4.10.1

($Sylvester$不等式) 设$A \in P^{s\times n}, B \in P^{n\times m}$, 证明 : $r(AB) \geq r(A) + r(B) - n$.

($Frobenius$不等式) 设$A \in P^{m\times n}, B \in P^{n\times s}, C \in P^{s\times t}$, 证明 : $r(ABC) \geq r(AB) + r(BC) - r(B)$.

Proof 因为 $\begin{pmatrix} E_n & 0 \\ 0 & AB \end{pmatrix} \to \begin{pmatrix} E_n & B \\ 0 & AB \end{pmatrix} \to \begin{pmatrix} E_n & B \\ -A & 0 \end{pmatrix} \to \begin{pmatrix} B & E_n \\ 0 & A \end{pmatrix}$.

故$r(AB)+n=r\begin{pmatrix}E_n & 0\\ 0 & AB\end{pmatrix}=r\begin{pmatrix}B & E_n\\ 0 & A\end{pmatrix}\geq r\begin{pmatrix}B & 0\\ 0 & A\end{pmatrix}=r(A)+r(B)$.

$$\begin{pmatrix}B & 0\\ 0 & ABC\end{pmatrix}\to\begin{pmatrix}B & BC\\ 0 & ABC\end{pmatrix}\to\begin{pmatrix}B & BC\\ -AB & 0\end{pmatrix}\to\begin{pmatrix}BC & B\\ 0 & AB\end{pmatrix}.$$

故$r(ABC)+r(B)=r\begin{pmatrix}B & 0\\ 0 & ABC\end{pmatrix}=r\begin{pmatrix}BC & B\\ 0 & AB\end{pmatrix}\geq r\begin{pmatrix}BC & 0\\ 0 & AB\end{pmatrix}=r(AB)+r(BC)$. □

Corollary 4.9

$A\in P^{s\times n}, B\in P^{n\times m}$, 当$AB=O$ 有$r(A)+r(B)\leq n$

Corollary 4.10

设$A_1\sim A_m$为n阶方阵则$r(A_1)+\cdots+r(A_m)\leqslant(m-1)n+r(A_1A_2\cdots A_m)$

Theorem 4.10.2 (秩降阶公式)

$M=\begin{pmatrix}A & B\\ C & D\end{pmatrix}$

(1) 若A可逆, 则$r(M)=r(A)+r\left(D-CA^{-1}B\right)$

(2) 若D可逆, 则$r(M)=r(D)+r\left(A-BD^{-1}C\right)$

(3) 若A,D可逆, 则$r(A)+r\left(D-CA^{-1}B\right)=r(D)+r\left(A-BD^{-1}C\right)$

4.11 满秩分解与 LU 分解

Lemma 4.4 (行列满秩矩阵的可逆性)

设$\boldsymbol{A}$是$m\times n$矩阵

(1)若 $\mathrm{r}(\boldsymbol{A})=n$, 即$\boldsymbol{A}$是列满秩阵, 则必存在秩等于$n$的$n\times m$矩阵$\boldsymbol{B}$, 使$\boldsymbol{BA}=\boldsymbol{I}_n$;

B称为A的左逆; 那么列满秩矩阵满足左消去律即 : A列满秩, $AD=AE \quad \Rightarrow D=E$

(2)若 $\mathrm{r}(\boldsymbol{A})=m$, 即$\boldsymbol{A}$是行满秩阵, 则必存在秩等于$m$的$n\times m$矩阵$\boldsymbol{C}$, 使$\boldsymbol{AC}=\boldsymbol{I}_m$.

C称为A的右逆; 那么行满秩矩阵满足右消去律即 : A行满秩.$DA=EA \quad \Rightarrow D=E$

Proof (1)存在可逆矩阵$\boldsymbol{P}$和$\boldsymbol{Q}$, 使$\boldsymbol{PAQ}=\begin{pmatrix}\boldsymbol{I}_n\\ \boldsymbol{O}\end{pmatrix}$, 因此 $(\boldsymbol{I}_n,\boldsymbol{O})\boldsymbol{PAQ}=\boldsymbol{I}_n$, 即 $(\boldsymbol{I}_n,\boldsymbol{O})\boldsymbol{PA}=\boldsymbol{Q}^{-1}$, 也就是$\boldsymbol{Q}(\boldsymbol{I}_n,\boldsymbol{O})\boldsymbol{PA}=\boldsymbol{I}_n$.
令$\boldsymbol{B}=\boldsymbol{Q}(\boldsymbol{I}_n,\boldsymbol{O})\boldsymbol{P}$即可.
(2) 同理

Theorem 4.11.1 (满秩分解)

设$\boldsymbol{A}$是秩为r的$m\times n$矩阵

(1) : $\boldsymbol{A}=\boldsymbol{BC}$, 其中$\boldsymbol{B}$是$m\times r$矩阵且 $\mathrm{r}(\boldsymbol{B})=r$, $\boldsymbol{C}$是$r\times n$矩阵且 $\mathrm{r}(\boldsymbol{C})=r$.即$\boldsymbol{A}$可分解为$r$阶列满秩与$r$阶行满秩矩阵的乘积

(2) : 若$\boldsymbol{A}$有两个满秩分解即$\boldsymbol{A}=\boldsymbol{B}_1\boldsymbol{C}_1=\boldsymbol{B}_2\boldsymbol{C}_2$; 则存在$r$阶可逆阵$\boldsymbol{P}$; $st.\boldsymbol{B}_2=\boldsymbol{B}_1\boldsymbol{P} \quad \boldsymbol{C}_2=\boldsymbol{P}^{-1}\boldsymbol{C}_1$

(3)若有两个满秩分解$C=AB=A_1B_1$那么BA与B_1A_1是相似的因为$A_1=AP$; $B_1=P^{-1}B$那么$P^{-1}BAP=B_1A_1$

Proof (1) 存在可逆矩阵$\boldsymbol{P}$和$\boldsymbol{Q}$, 使$\boldsymbol{A}=\boldsymbol{P}_{m\times m}\begin{pmatrix}\boldsymbol{I}_r & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{pmatrix}_{m\times n}\boldsymbol{Q}_{n\times n}=\boldsymbol{P}\begin{pmatrix}\boldsymbol{I}_r\\ \boldsymbol{O}\end{pmatrix}_{m\times r}(\boldsymbol{I}_r,\boldsymbol{O})_{r\times n}\boldsymbol{Q}$. 令$\boldsymbol{B}_{m\times r}=\boldsymbol{P}\begin{pmatrix}\boldsymbol{I}_r\\ \boldsymbol{O}\end{pmatrix}$, $\boldsymbol{C}=(\boldsymbol{I}_r,\boldsymbol{O})\boldsymbol{Q}$即可.
但是我们仍要说明B的秩为r, C的秩为r, 显然$\begin{pmatrix}\boldsymbol{I}_r\\ \boldsymbol{O}\end{pmatrix}$秩为$r$左乘一个可逆阵秩不变即可
值得注意的是我们并没有用到B,C的内部结构这也就是说我们只要找到$B^{m\times r}$与$C^{r\times n}$ $st.A=BC$那么就为满秩分解这一点后续会用到

(2) 由上题引理知道存在$\boldsymbol{S}^2_{r\times m}$与$\boldsymbol{T}^2_{n\times r}$行与列满秩矩阵 $st.\boldsymbol{S}_2\boldsymbol{B}_2=\boldsymbol{I}_r \quad \boldsymbol{C}_2\boldsymbol{T}_2=\boldsymbol{I}_r$

$$\Rightarrow\begin{cases}\boldsymbol{B}_2=\boldsymbol{B}_2(\boldsymbol{C}_2\boldsymbol{T}_2)=(\boldsymbol{B}_2\boldsymbol{C}_2)\boldsymbol{T}_2=(\boldsymbol{B}_1\boldsymbol{C}_1)\boldsymbol{T}_2=\boldsymbol{B}_1(\boldsymbol{C}_1\boldsymbol{T}_2)\\ \boldsymbol{C}_2=(\boldsymbol{S}_2\boldsymbol{B}_2)\boldsymbol{C}_2=\boldsymbol{S}_2(\boldsymbol{B}_2\boldsymbol{C}_2)=\boldsymbol{S}_2(\boldsymbol{B}_1\boldsymbol{C}_1)=(\boldsymbol{S}_2\boldsymbol{B}_1)\boldsymbol{C}_1\end{cases}$$

至此我们得到了跟要证式子相似的结论但是结论中是共用一个P那么
$(\boldsymbol{S}_2\boldsymbol{B}_1)(\boldsymbol{C}_1\boldsymbol{T}_2)=\boldsymbol{S}_2(\boldsymbol{B}_1\boldsymbol{C}_1)\boldsymbol{T}_2=\boldsymbol{S}_2(\boldsymbol{B}_2\boldsymbol{C}_2)\boldsymbol{T}_2=(\boldsymbol{S}_2\boldsymbol{B}_2)(\boldsymbol{C}_2\boldsymbol{T}_2)=\boldsymbol{I}_r$
那么令$\boldsymbol{P}=\boldsymbol{S}_2\boldsymbol{B}_1$即可 □

Proposition 4.18

1.设$\boldsymbol{A}$是数域K上的一个$s\times n$矩阵, 且$\boldsymbol{A}\neq 0$

证明 : $\mathrm{rank}(\boldsymbol{A})=1$当且仅当$\boldsymbol{A}$能表示成一个$s$维列向量与一个$n$维行向量的乘积

2.设$\boldsymbol{A}$是数域K上的一个n级矩阵.证明 : 若 $\mathrm{rank}(\boldsymbol{A})=1$, 则存在唯一的$k\in K$使得$\boldsymbol{A}^2=k\boldsymbol{A}$

Proof 1.由满秩分解技术因为$r(A)=1$故存在$\alpha_{s\times 1},\beta^T_{1\times n}$使得$A=\alpha\beta^T$其中$r(\alpha)=r(\beta^T)=1$
若A可以表示为$A=\alpha\beta^T$那么$r(A)\leqslant r(\alpha)\leqslant 1$又$\boldsymbol{A}\neq 0\Longrightarrow r(A)=1$
2.由1.问知道$A=\alpha_{n\times 1},\beta^T_{1\times n}\Longrightarrow A^2=(\beta^T\alpha)A$故存在$k=(\beta^T\alpha)$, 至此证明了存在性
若还存在一个k^*那么$A^2-A^2=(k-k^*)A=0$而$\boldsymbol{A}\neq 0\Longrightarrow(k-k^*)=0\Longrightarrow k=k^*$

Proposition 4.19

设A是数域K上的一个$s\times n$行满秩矩阵, 证明：对于K上任意一个$s\times m$矩阵$\boldsymbol{B}$, 矩阵方程$\boldsymbol{A}\boldsymbol{X}=\boldsymbol{B}$都有解

Proof 由题知道存在$T_{n\times s}$列满秩矩阵使得$AT=I_s$
故对于$\forall \boldsymbol{B}_{s\times m}$我们取$\boldsymbol{X}=T\boldsymbol{B}$即可

Proposition 4.20

设$\boldsymbol{A}$是数域K上的n级矩阵，且 $\operatorname{rank}(\boldsymbol{A})=1$。试问：$\boldsymbol{I}+\boldsymbol{A}$是否可逆?当$\boldsymbol{I}+\boldsymbol{A}$可逆时, 求 $(\boldsymbol{I}+\boldsymbol{A})^{-1}$

Proof 由于 $\operatorname{rank}(\boldsymbol{A})=1$, 因此之前的结论知道得, 存在唯一的$k\in K$使得$\boldsymbol{A}^2=k\boldsymbol{A}$
此时根据$A^2-kA=O$待定系数
由此得出有 $(\boldsymbol{I}+\boldsymbol{A})\left(\boldsymbol{I}-\dfrac{1}{1+k}\boldsymbol{A}\right)=\boldsymbol{I}$
故$k\neq -1$时, 此时$\boldsymbol{I}+\boldsymbol{A}$可逆, 且 $(\boldsymbol{I}+\boldsymbol{A})^{-1}=\boldsymbol{I}-\dfrac{1}{1+k}\boldsymbol{A}$
当$k=-1$时, $\boldsymbol{A}^2=-\boldsymbol{A}$, 从而$0=\boldsymbol{A}^2+\boldsymbol{A}=\boldsymbol{A}(\boldsymbol{A}+\boldsymbol{I})$
于是根据$Slverster$不等式得, $\operatorname{rank}(\boldsymbol{A})+\operatorname{rank}(\boldsymbol{A}+\boldsymbol{I})\leqslant n$.由于 $\operatorname{rank}(\boldsymbol{A})=1$, 因此 $\operatorname{rank}(\boldsymbol{A}+I)\leqslant n-1<n$, 从而$\boldsymbol{A}+\boldsymbol{I}$不可逆

笔记

在满秩分解A = BC中，rank(A) = rank(B)，并且B是列满秩矩阵, 于是可以考虑将B直接设为为A的列向量的最大线性无关向量组
不妨设B = $\begin{pmatrix}\alpha_1 & \alpha_2 & \cdots & \alpha_r\end{pmatrix}$，其中 $\begin{Bmatrix}l\alpha_1 & \alpha_2 & \cdots & \alpha_r\end{Bmatrix}$ 为A的列向量的最大线性无关向量组
则A的剩下列向量可以表示为列向量极大线性无关向量组的组合，不妨记剩下列向量为A_1, 则有$\mathrm{A}_1=\mathrm{BS}$.
为了让问题再次简化，不妨假设B恰好为A的前r列，则$\mathrm{A}=(\mathrm{B}\mid\mathrm{A}_1)=\mathrm{B}\,(\mathrm{I}_r\mid\mathrm{S})$，取$\mathrm{C}=(\mathrm{I}_r\mid\mathrm{S})$即得到了A的满秩分解方案.
为了得到$\mathrm{C}=(\mathrm{I}_r\mid\mathrm{S})$的结构，在初等行变换时需满足得到的阶梯型矩阵每一行第一个非零元素为1，该元素所在列其他元素为0
这种结构也称为$Hermite$标准形.
但事情通常并不会那么顺利，倘若B并不对应A的前r列呢?答案是简单的，直接让A进行初等列变化
让A的列向量最大线性无关向量组位于A的前r列即可，也即右乘一个可逆矩阵P，使得$\mathrm{AP}=\mathrm{B}\,(\mathrm{I}_r\mid\mathrm{S})$，
此时B仍然是A列向量最大线性无关组，$\mathrm{C}=(\mathrm{I}_r\mid\mathrm{S})\,\mathrm{P}^{-1}$.P的作用是将A的列向量最大线性无关向量组移动到A的前r列
P^{-1}的作用自然是再移回去，也就意味着C各列向量的顺序和A是一致的，至此，我们就得到了满秩分解的另一种方法.

Problem 4.4

$$\boldsymbol{B}=\begin{bmatrix}2&-1&-1&1&2\\1&1&-2&1&4\\4&-6&2&-2&4\\3&6&-9&7&9\end{bmatrix}=(\boldsymbol{b}_1,\boldsymbol{b}_2,\boldsymbol{b}_3,\boldsymbol{b}_4,\boldsymbol{b}_5)$$

$\boldsymbol{B}$的行最简形矩阵为$\boldsymbol{B}\sim\begin{bmatrix}1&0&-1&0&4\\0&1&-1&0&3\\0&0&0&1&-3\\0&0&0&0&0\end{bmatrix}=\boldsymbol{B}_1$所以矩阵$\boldsymbol{B}$的极大线性无关组为$\boldsymbol{b}_1,\boldsymbol{b}_2,\boldsymbol{b}_4$。

同时式也可以看出，矩阵$\boldsymbol{B}$的每个列向量写成极大线性无关组的线性表示的形式分别为：

$$\begin{cases} \boldsymbol{b}_1 = 1\boldsymbol{b}_1 + 0\boldsymbol{b}_2 + 0\boldsymbol{b}_4 \\ \boldsymbol{b}_2 = 0\boldsymbol{b}_1 + 1\boldsymbol{b}_2 + 0\boldsymbol{b}_4 \\ \boldsymbol{b}_3 = -1\boldsymbol{b}_1 - 1\boldsymbol{b}_2 + 0\boldsymbol{b}_4 \\ \boldsymbol{b}_4 = 0\boldsymbol{b}_1 + 0\boldsymbol{b}_2 + 1\boldsymbol{b}_4 \\ \boldsymbol{b}_5 = 4\boldsymbol{b}_1 + 3\boldsymbol{b}_2 - 3\boldsymbol{b}_4 \end{cases}$$

写成矩阵乘法的形式：

$$(\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3, \boldsymbol{b}_4, \boldsymbol{b}_5) = (\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_4)\begin{bmatrix} 1 & 0 & -1 & 0 & 4 \\ 0 & 1 & -1 & 0 & 3 \\ 0 & 0 & 0 & 1 & -3 \end{bmatrix}$$

$$= \begin{bmatrix} 2 & -1 & 1 \\ 1 & 1 & 1 \\ 4 & -6 & -2 \\ 3 & 6 & 7 \end{bmatrix}\begin{bmatrix} 1 & 0 & -1 & 0 & 4 \\ 0 & 1 & -1 & 0 & 3 \\ 0 & 0 & 0 & 1 & -3 \end{bmatrix}$$

就得到满秩分解了

Theorem 4.11.2 (LU 分解定理)

证明：数域K上的n级矩阵$\boldsymbol{A}$能够分解成一个主对角元都为1的下三角矩阵$\boldsymbol{L}$与可逆上三角矩阵$\boldsymbol{U}$的乘积$\boldsymbol{A} = \boldsymbol{LU}$当且仅当$\boldsymbol{A}$的各阶顺序主子式全不为0, 并且$\boldsymbol{A}$的这种分解是唯一的。

Proof 先来证充分性

显然 $|A| \neq 0$

当$n = 1$时显然成立

假设$n-1$成立.此时对n我们来考虑$A = \begin{pmatrix} A_1 & \alpha \\ \beta & a_{nn} \end{pmatrix}$

$$\begin{pmatrix} I_{n-1} & O \\ -\beta A_1^{-1} & 1 \end{pmatrix}\begin{pmatrix} A_1 & \alpha \\ \beta & a_{nn} \end{pmatrix}\begin{pmatrix} I_{n-1} & -A_1^{-1}\alpha \\ O & 1 \end{pmatrix} = \begin{pmatrix} A_1 & O \\ O & a_{nn} - \beta A_1^{-1}\alpha \end{pmatrix}$$

此时 $|A| = \left|\begin{pmatrix} A_1 & O \\ O & a_{nn} - \beta A_1^{-1}\alpha \end{pmatrix}\right| \neq 0 \Longrightarrow a_{nn} - \beta A_1^{-1}\alpha \neq 0$我们同样可以对$\begin{pmatrix} A_1 & O \\ O & a_{nn} - \beta A_1^{-1}\alpha \end{pmatrix}$用归纳假设

故存在$L_1U_1 = A_1$

那么 $\begin{pmatrix} L_1 & O \\ O & 1 \end{pmatrix}\begin{pmatrix} U_1 & O \\ O & a_{nn} - \beta A_1^{-1}\alpha \end{pmatrix} = \begin{pmatrix} A_1 & O \\ O & a_{nn} - \beta A_1^{-1}\alpha \end{pmatrix}$

此时 $\begin{pmatrix} U_1 & O \\ O & a_{nn} - \beta A_1^{-1}\alpha \end{pmatrix}$为可逆的上三角矩阵 $\begin{pmatrix} L_1 & O \\ O & 1 \end{pmatrix}$为主对角线全为1的下三角矩阵

那么 $\begin{pmatrix} I_{n-1} & O \\ -\beta A_1^{-1} & 1 \end{pmatrix}\begin{pmatrix} A_1 & \alpha \\ \beta & a_{nn} \end{pmatrix}\begin{pmatrix} I_{n-1} & -A_1^{-1}\alpha \\ O & 1 \end{pmatrix} = \begin{pmatrix} A_1 & O \\ O & a_{nn} - \beta A_1^{-1}\alpha \end{pmatrix} = \begin{pmatrix} L_1 & O \\ O & 1 \end{pmatrix}\begin{pmatrix} U_1 & O \\ O & a_{nn} - \beta A_1^{-1}\alpha \end{pmatrix}$

而上三角矩阵的逆矩阵也为上三角矩阵故$U^* = \begin{pmatrix} U_1 & O \\ O & a_{nn} - \beta A_1^{-1}\alpha \end{pmatrix}\begin{pmatrix} I_{n-1} & -A_1^{-1}\alpha \\ O & 1 \end{pmatrix}^{-1}$是可逆的上三角矩阵

而下三角矩阵的逆矩阵也为下三角矩阵故$L^* = \begin{pmatrix} I_{n-1} & O \\ -\beta A_1^{-1} & 1 \end{pmatrix}^{-1}\begin{pmatrix} L_1 & O \\ O & 1 \end{pmatrix}$是主对角线全为1的下三角矩阵

$$\Longrightarrow A = \begin{pmatrix} A_1 & \alpha \\ \beta & a_{nn} \end{pmatrix} = \underbrace{\begin{pmatrix} I_{n-1} & O \\ -\beta A_1^{-1} & 1 \end{pmatrix}^{-1}\begin{pmatrix} L_1 & O \\ O & 1 \end{pmatrix}}_{L^*}\underbrace{\begin{pmatrix} U_1 & O \\ O & a_{nn} - \beta A_1^{-1}\alpha \end{pmatrix}\begin{pmatrix} I_{n-1} & -A_1^{-1}\alpha \\ O & 1 \end{pmatrix}^{-1}}_{U^*} = L^*U^*$$

再证必要性：

必要性, 从$\boldsymbol{A}=\boldsymbol{L}\boldsymbol{U}$得$\boldsymbol{A}=\begin{pmatrix}\boldsymbol{B}_1 & 0\\ \boldsymbol{B}_2 & \boldsymbol{B}_3\end{pmatrix}\begin{pmatrix}\boldsymbol{C}_1 & \boldsymbol{C}_2\\ 0 & \boldsymbol{C}_3\end{pmatrix}$, 其中$\boldsymbol{B}_1,\boldsymbol{C}_1$都是$k$级矩阵, $|\boldsymbol{B}_1\boldsymbol{C}_1|$是$\boldsymbol{A}$的$k$阶顺序主子式, $k\in\{1,2,\cdots,n-1\}$

再证明唯一性：

若有$A=LU=L^*U^*$, 此时显然L,U,L^*,U^*都可逆那么

$\Longrightarrow L^{-1}L^*=U\left(U^*\right)^{-1}\Longrightarrow L^{-1}L^*$此时仍然为主对角线全为1的下三角矩阵而$U\left(U^*\right)^{-1}$为上三角矩阵

$\Longrightarrow L^{-1}L^*=U\left(U^*\right)^{-1}=I\Longrightarrow L=L^*$且$U=U^*$

4.12 线性空间重要结论

Theorem 4.12.1 (真子空间无法覆盖全空间)

设$V_1, V_2, \cdots, V_m$是数域$\mathbb{F}$上向量空间V的m个真子空间, 证明：在V中必存在一个向量α, 它不属于任何一个V_i.

Proof 对m用归纳法. 当$m=1$时结论显然成立. 设$m=k$时结论成立, 现要证明$m=k+1$时结论也成立.
由归纳假设, 存在向量α, 它不属于任何一个$V_i (i=1, \cdots, k)$. 若α也不属于V_{k+1}, 则结论已成立
因此可设$\alpha \in V_{k+1}$. 在V_{k+1}外选一个向量β. 作集合$M=\{t\alpha+\beta \mid t \in \mathbb{F}\}$.
事实上, 我们可将M看成是通过β的终点且平行于α的一根直线, 现要证明它和每个V_i最多只有一个交点.
首先, M和V_{k+1}无交点, 因为若$t\alpha+\beta \in V_{k+1}$, 则从$t\alpha \in V_{k+1}$可推出$\beta \in V_{k+1}$, 与假设矛盾.
又若$t_1\alpha+\beta \in V_i, t_2\alpha+\beta \in V_i (i<k+1)$, 则$(t_1-t_2)\alpha \in V_i$. 若$t_1 \neq t_2$, 将导致$\alpha \in V_i$, 与假设矛盾.
由此可以看到, M中只有有限个向量属于V_i的并集, 而t有无穷多个选择, 由此即得结论. □

注 证明中用到任意一个数域都有无穷个元素这一事实, 因此, 对有限域 (以后可能会学到) 上的向量空间, 结论不一定成立.

Proposition 4.21

设$V_1, V_2, \cdots, V_m$是数域$\mathbb{F}$上向量空间V的m个真子空间, 证明：V中必有一组基, 使得其中每个基向量都不在诸V_i的并中.

Proof 由上个问题可知, 存在非零向量$\boldsymbol{e}_1 \in V$, 使得$\boldsymbol{e}_1 \notin \bigcup\limits_{i=1}^{m} V_i$. 定义$V_{m+1}=L(\boldsymbol{e}_1)$
同理可知, 存在向量$\boldsymbol{e}_2 \in V$, 使得$\boldsymbol{e}_2 \notin \bigcup\limits_{i=1}^{m+1} V_i$.
由向量组重要论断3可知, $\boldsymbol{e}_2 \notin L(\boldsymbol{e}_1)$ 意味着$\boldsymbol{e}_1, \boldsymbol{e}_2$线性无关.
重新定义$V_{m+1}=L(\boldsymbol{e}_1, \boldsymbol{e}_2)$, 同理可知, 存在向量$\boldsymbol{e}_3 \in V$, 使得$\boldsymbol{e}_3 \notin \bigcup\limits_{i=1}^{m+1} V_i$.
再由论断可知, $\boldsymbol{e}_3 \notin L(\boldsymbol{e}_1, \boldsymbol{e}_2)$ 意味着$\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$线性无关.
不断重复上述讨论, 即添加线性无关的向量重新定义V_{m+1}, 并反复利用结论
最后可以得到n个线性无关的向量$\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$, 它们构成$V$的一组基, 且满足$\boldsymbol{e}_j \notin \bigcup\limits_{i=1}^{m} V_i, j=1,2,\cdots,n$. □

注 利用几何问题代数化这一技巧, 我们可以给出上述两道例题的一个统一解法.
任取V的一组基$\{\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n\}$.
对任意的正整数k, 构造V中向量$\boldsymbol{\alpha}_k=\boldsymbol{e}_1+k\boldsymbol{e}_2+\cdots+k^{n-1}\boldsymbol{e}_n$设向量族$S=\{\boldsymbol{\alpha}_k \mid k=1,2,\cdots\}$.
可知, S中任意n个不同的向量都构成V的一组基. (利用坐标向量和范德蒙德行列式)
因为V_i都是V的真子空间, 所以每个V_i至多包含S中$n-1$个向量.
由于S是无限集合, 故存在某个向量$\boldsymbol{\alpha}_k$, 使得$\boldsymbol{\alpha}_k$不属于任何一个V_i, 这就证明了第一问
进一步, 在S中还存在n个不同的向量$\boldsymbol{\alpha}_{k_1}, \boldsymbol{\alpha}_{k_2}, \cdots, \boldsymbol{\alpha}_{k_n}$, 使得每个$\boldsymbol{\alpha}_{k_j}$都不属于任何一个$V_i$, 此时$\{\boldsymbol{\alpha}_{k_1}, \boldsymbol{\alpha}_{k_2}, \cdots, \boldsymbol{\alpha}_{k_n}\}$构成了$V$的一组基
证毕

Theorem 4.12.2 (互素多项式诱导了方程组解空间的直和)

若 $(f(x), g(x))$ 在数域$\mathbb{K}$上互素，A为n阶方阵.设$f(A)x=0$与$g(A)x=0$的解空间为V_1，V_2则$V_1+V_2=V_1\oplus V_2$

此外若还有$f(A)g(A)=O$那么$\mathbb{K}^n=V_1+V_2=V_1\oplus V_2$

Proof 由互素性质知道$\exists u(x)$与$v(x)$使得$u(x)f(x)+v(x)g(x)=1$

$\Longrightarrow u(A)f(A)+v(A)g(A)=I_n$

此时$\forall x\in V_1\cap V_2$那么$u(A)f(A)\vec{x}+v(A)g(A)\vec{x}=\vec{x}\Longrightarrow x=0$

故$V_1+V_2=V_1\oplus V_2$

此外,$\forall x\in\mathbb{K}^n$那么$\underbrace{u(A)f(A)x}_{\in V_2}+\underbrace{v(A)g(A)x}_{\in V_1}=x\Longrightarrow\mathbb{K}^n=V_1+V_2=V_1\oplus V_2$

$\left(\text{这是因为}g(A)\cdot u(A)f(A)x=u(A)g(A)f(A)x=0\right)$

Theorem 4.12.3 (互素多项式诱导矩阵秩不等式)

设$f(x),g(x)$是数域K上互素的多项式,A是数域K上的n阶矩阵

证明：$f(A)g(A)=O\Longleftrightarrow\operatorname{rank}(f(A))+\operatorname{rank}(g(A))=n$，其中 rank 表示矩阵的秩

Proof 由于$(f(x),g(x))=1$,所以存在多项式$u(x),v(x)\in K[x]$,满足$u(x)f(x)+v(x)g(x)=1$

$\Longrightarrow u(A)f(A)+v(A)g(A)=E_n$

现在对矩阵$\operatorname{diag}\{f(A),g(A)\}$作广义初等变换,有

$$\begin{pmatrix}f(A)&O\\O&g(A)\end{pmatrix}\longrightarrow\begin{pmatrix}f(A)&f(A)u(A)\\O&g(A)\end{pmatrix}\longrightarrow\begin{pmatrix}f(A)&f(A)u(A)+v(A)g(A)\\O&g(A)\end{pmatrix}$$
$$==\begin{pmatrix}f(A)&E_n\\O&g(A)\end{pmatrix}\longrightarrow\begin{pmatrix}O&E_n\\-g(A)f(A)&g(A)\end{pmatrix}\longrightarrow\begin{pmatrix}O&E_n\\-g(A)f(A)&O\end{pmatrix}\longrightarrow\begin{pmatrix}E_n&O\\O&f(A)g(A)\end{pmatrix}.$$

由此可知 $\operatorname{rank}(f(A))+\operatorname{rank}(g(A))=\operatorname{rank}(E_n)+\operatorname{rank}(f(A)g(A))=n+\operatorname{rank}(f(A)g(A))$那么

$\operatorname{rank}(f(A))+\operatorname{rank}(g(A))=n$的充要条件是 $\operatorname{rank}(f(A)g(A))=0$,即$f(A)g(A)=O$.

Corollary 4.11

设数域$\mathbb{K}$上的多项式$f(x),g(x)$且$(f(x),g(x))=d(x)$.设$m(x)$为$f(x)$与$g(x)$的最小公倍式

A为n阶方阵,$m(A)=O$.若设$f(A)x=0$与$g(A)x=0$与$d(A)x=0$的解空间为V_1,V_2,V_3.

证明：$rank(f(A))+rank(g(A))=rank(d(A))$或$dimV_1+dimV_2=n+dimV_3$

Proof 通过方程组系数矩阵秩与解空间维数关系即可得到两个结论实际上是等价的

设$f(x)=f_1(x)d(x)$且$g(x)=g_1(x)d(x)$那么$m(A)=O\Longleftrightarrow f_1(A)g_1(A)d(A)=O$

$\exists u(x)$与$v(x)$使得$u(x)f(x)+v(x)g(x)=d(x)\Longrightarrow u(A)f(A)+v(A)g(A)=d(A)$

此时$\forall x\in V_1\cap V_2$,则$u(A)f(A)\boldsymbol{x}+v(A)g(A)\boldsymbol{x}=d(A)\boldsymbol{x}\Longrightarrow d(A)\boldsymbol{x}=0\Longrightarrow\boldsymbol{x}\in V_3$

$\Longrightarrow V_1\cap V_2\subset V_3$

此外$\forall x\in V_3$即$d(A)\boldsymbol{x}=0\Longrightarrow f(A)\boldsymbol{x}=f_1(A)d(A)\boldsymbol{x}=0$同理$g(A)\boldsymbol{x}=0\Longrightarrow\boldsymbol{x}\in V_1$且$x\in V_2\Longrightarrow x\in V_1\cap V_2$

$\Longrightarrow V_3\subset V_1\cap V_2$

故$V_3=V_1\cap V_2$

此外$u(x)f_1(x)+v(x)g_1(x)=1\Longrightarrow u(A)f_1(A)+v(A)g_1(A)=1$

故$\forall x\in\mathbb{K}^n$则$\underbrace{u(A)f_1(A)\boldsymbol{x}}_{\in V_2}+\underbrace{v(A)g_1(A)\boldsymbol{x}}_{\in V_1}=\boldsymbol{x}\Longrightarrow V_1+V_2=\mathbb{K}^n$ $\left(\text{这是因为}g(A)u(A)f_1(A)\boldsymbol{x}=0\right)$

故由维数公式知道$dim(V_1+V_2)+dim(V_1\cap V_2)=dim(V_1)+dim(V_2)$

$\Longrightarrow n+dimV_3=dim(V_1)+dim(V_2)$

代数证法：一方面$u(A)f(A)+v(A)g(A)=d(A)\Longrightarrow r(d(A))\leqslant r(u(A)f(A))+r(v(A)g(A))\leqslant r(f(A))+r(g(A))$

另一方面：

$$\begin{pmatrix} d(A) & \\ & O \end{pmatrix}=\begin{pmatrix} d(A) & \\ & f_1(A)g_1(A)d(A) \end{pmatrix}\longrightarrow\begin{pmatrix} d(A) & d(A)f_1(A) \\ & f_1(A)g_1(A)d(A) \end{pmatrix}\longrightarrow\begin{pmatrix} d(A) & d(A)f_1(A) \\ -g_1(A)d(A) & O \end{pmatrix}$$

$$\longrightarrow\begin{pmatrix} d(A) & d(A)f_1(A) \\ -g_1(A)d(A) & O \end{pmatrix}=\begin{pmatrix} d(A) & f(A) \\ g(A) & O \end{pmatrix}$$

$\Longrightarrow r(d(A))=r\begin{pmatrix} d(A) & f(A) \\ g(A) & O \end{pmatrix}\geqslant r(f(A))+r(g(A))$ (实际上由$Frobenius$不等式即可)

$\Longrightarrow rank(f(A))+rank(g(A))=rank(d(A))$

Proposition 4.22

令$m\times n$阶实矩阵M全体构成的线性空间为$\mathbb{R}^{m\times n}$, 给定$p\times m$阶实矩阵A, $n\times q$阶实矩阵B

定义线性映射$\varphi:\mathbb{R}^{m\times n}\longrightarrow\mathbb{R}^{p\times q}$, $\varphi(M)=AMB$

$\Longrightarrow$ 给出φ为可逆线性映射的充要条件, 并求φ^{-1}.

Proof 断言：φ可逆的充要条件是$p=m, q=n$，且A,B均为可逆矩阵

其中充分性是显然的, 当$p=m, q=n$, 且A,B均为可逆矩阵, φ就是$\mathbb{R}^{m\times n}$上的线性变换, 其逆变换φ^{-1}为$\varphi^{-1}(M)=A^{-1}MB^{-1}, M\in\mathbb{R}^{m\times n}$

下面证明必要性, 设$r(A)=r, r(B)=s$, 那么存在阶数分别为p,m,n,q的可逆矩阵P,Q,P_1,Q_1

使得$A=P\begin{pmatrix} E_r & O \\ O & O \end{pmatrix}Q \quad B=P_1\begin{pmatrix} E_s & O \\ O & O \end{pmatrix}Q_1$

对任意的$M\in\mathbb{R}^{m\times n}$, 记$QMP_1=\begin{pmatrix} M_1 & M_2 \\ M_3 & M_4 \end{pmatrix}$，其中$M_1$为$r\times s$矩阵

那么$M\in Ker\varphi\Longleftrightarrow\varphi(M)=O\Longleftrightarrow AMB=P\begin{pmatrix} E_r & O \\ O & O \end{pmatrix}\begin{pmatrix} M_1 & M_2 \\ M_3 & M_4 \end{pmatrix}\begin{pmatrix} E_s & O \\ O & O \end{pmatrix}Q_1=P\begin{pmatrix} M_1 & O \\ O & O \end{pmatrix}Q_1=O\Longleftrightarrow M_1=O$

也就是$M=Q^{-1}\begin{pmatrix} O & M_2 \\ M_3 & M_4 \end{pmatrix}P_1^{-1}$其中$M_2,M_3,M_4$为任意的$r\times(n-s), (m-r)\times s, (m-r)\times(n-s)$矩阵

此时φ可逆自然需要φ为单射需要$Ker\varphi=\{0\}$则$M\Longrightarrow$ 则$\begin{pmatrix} O & M_2 \\ M_3 & M_4 \end{pmatrix}=O$

因此上式中不会出现M_2,M_3,M_4, 也就是$r=m, s=n$

即$A^{p\times m}$为列满秩矩阵，$B^{n\times q}$为行满秩矩阵, 这也意味着$p\geqslant m, q\geqslant n$.

另外, 根据φ可逆根据维数公式就有 $\dim\mathbb{R}^{m\times n}=\dim\mathbb{R}^{p\times q}$, 也就是$mn=pq$, 因此$p=m, q=n$, 即$A,B$分别为$m$阶和$n$阶的可逆矩阵.

第 5 章 线性映射

5.1 集合与映射

Definition 5.1

设M到M'为两个集合,所谓$M \to M'$的一个映射指的是一个对应法则,$st.M$中的每一个元素都有M'的一个确定的元素与之对应.特殊的$M \to M$的映射我们称之为变换

另外我们定义两个映射相等的概念 : 对于σ和τ $(M \to M')$,$\forall a \in M$，有$\sigma(a) = \tau(a)$; 特别的若有一个映射将M映射到M; 且$\tau(a) = a$我们称之为恒等映射记Id_M；显然$Id_M\sigma = \sigma Id_M = \sigma$

再定义映射的乘法, 设σ和τ都为集合$M \to M'$的映射乘积$\tau\sigma$定义为$\tau\sigma(a) = \tau(\sigma(a))$ 其中$a \in M$

若映射保持了向量的加法和数乘我们则称为线性映射

性质 映射的乘法满足乘法结合律, 但是一般并不满足交换律显然的$\sigma : S \to S'$和$\tau : S' \to S''$那么其交换甚至可能没有意义因为$S'' \cap S' = \emptyset$

Definition 5.2

设σ是M到M'的一个映射，我们用$\sigma(M)$来表示M在映射σ下的像的全体。显然 $\sigma(M) \subseteq M'$, 若 $\sigma(M) = M'$ 我们称之为满射 (映上的)。若有$\forall a_1 \neq a_2$，$\sigma(a_1) \neq \sigma(a_2)$ 我们称该映射是单射 如果一个映射不仅是单射也是满射我们称之为双射,

性质 已知$f : A \to B$，$g : B \to C$; $\begin{cases} 若f单射, g单射 \Rightarrow gf也单射 \\ 若f满射, g满射 \Rightarrow gf也满射 \\ 若gf单射 \Rightarrow f单射 \\ 若gf满射 \Rightarrow g满射 \end{cases}$

Definition 5.3

对于σ满射自然可以定义其逆映射σ^{-1}显然也为满射$(M' \to M)$。显然若σ为双射 $\Leftrightarrow \exists\tau$ 映射 $(M' \to M)$ st.$\tau\sigma = 1_M$ and $\sigma\tau = 1'_M$

性质 显然若 f,g 都为满射则 gf 也为满射

性质 线性映射 (同构) 的复合仍然是线性映射 (同构)

Theorem 5.1.1

设$f : A \to B$的映射(map) ; 则f存在逆映射 $\Leftrightarrow$ f是双射

Proof 先证充分性，设f为双射，那么$\forall b \in B$, 因为f满射所以$\exists a \in A, st; f(a) = b$;

再由单性, 则$f(a) = b$的a是唯一的。这就证明了，$\forall b \in B$, 存在唯一的$a \in A; st f(a) = b$

构造g：$B \to A\ (b \to a)$ (满足$f(a) = b$)

容易发现$f \circ g = 1_B$; $g \circ f = 1_A$; 所以$g = f^{-1}$

再证必要性，设$g : B \to A\ (b \to a)$ 是f的逆映射则$f \circ g = 1_B$; $g \circ f = 1_A$

假设$a_1, a_2 \in A, st.f(a_1) = f(a_2)$; 用g复合，则 $\Rightarrow a_1 = a_2$

假设$\forall b \in B$, 带入$f \circ g = 1_B$; $\Rightarrow f(g(b)) = b$; 则b至少有一个原像$g(b)$

上述就证明了单满性。

综上证毕

Theorem 5.1.2 (像集和原像集的关系)

设f是X到Y的一个映射.设$A \subset X, B \subset Y$.
令$f(A) = \{f(x) \in Y \mid x \in A\}$和$f^{-1}(B) = \{x \in X \mid f(x) \in B\}$.
称$f(A)$为集合A在映射f下的像集, 称$f^{-1}(B)$为集合B在映射f下的原像集.

设$f: X \longrightarrow Y$为从X到Y的一个映射, $\{A_\alpha\}_{\alpha\in\Lambda_1}$ 为X中子集族, $\{B_\alpha\}_{\alpha\in\Lambda_2}$ 为Y中子集族, $B \subset Y$, 则

$(i) f\left(\bigcup_{\alpha\in\Lambda_1} A_\alpha\right) = \bigcup_{\alpha\in\Lambda_1} f(A_\alpha);$

$(ii) f\left(\bigcap_{\alpha\in\Lambda_1} A_\alpha\right) \subset \bigcap_{\alpha\in\Lambda_1} f(A_\alpha);$

(iii)若$B_1 \subset B_2 \subset Y$, 则$f^{-1}(B_1) \subset f^{-1}(B_2);$

$(iv) f^{-1}\left(\bigcup_{\alpha\in\Lambda_2} B_\alpha\right) = \bigcup_{\alpha\in\Lambda_2} f^{-1}(B_\alpha);$

$(v) f^{-1}\left(\bigcap_{\alpha\in\Lambda_2} B_\alpha\right) = \bigcap_{\alpha\in\Lambda_2} f^{-1}(B_\alpha);$

$(vi) f^{-1}(B^c) = \left(f^{-1}(B)\right)^c.$

Proof $(i) y \in f\left(\bigcup_{\alpha\in\Lambda} A_\alpha\right) \Leftrightarrow \exists x \in \bigcup_{\alpha\in\Lambda} A_\alpha, y = f(x) \Leftrightarrow \exists \alpha \in \Lambda, x \in A_\alpha, y = f(x) \Leftrightarrow y \in \bigcup_{\alpha\in\Lambda} f(A_\alpha);$

$(ii)\ y \in f\left(\bigcap_{\alpha\in\Lambda_1} A_\alpha\right) \Rightarrow \exists x \in \bigcap_{\alpha\in\Lambda_1} A_\alpha, y = f(x) \Rightarrow y \in \bigcap_{\alpha\in\Lambda_1} f(A_\alpha)$

但反之例如$A_1 = \{1, 2\}; A_2 = \{2, 3\}; f(1) = a; f(2) = b; f(3) = a$并不对

(iii) 易知

$(iv) x \in f^{-1}\left(\bigcup_{\alpha\in\Lambda} B_\alpha\right) \Leftrightarrow f(x) \in \bigcup_{\alpha\in\Lambda} B_\alpha \Leftrightarrow \exists \alpha \in \Lambda, f(x) \in B_\alpha \Leftrightarrow \exists \alpha \in \Lambda, x \in f^{-1}(B_\alpha) \Leftrightarrow x \in \bigcup_{\alpha\in\Lambda} f^{-1}(B_\alpha).$

$(v)\ x \in f^{-1}\left(\bigcap_{\alpha\in\Lambda} B_\alpha\right) \Leftrightarrow f(x) \in \bigcap_{\alpha\in\Lambda} B_\alpha \Rightarrow x \in \bigcap_{\alpha\in\Lambda_2} f^{-1}(B_\alpha)$

$(vi)\ x \in f^{-1}(B^c) \Leftrightarrow f(x) \in B^c \Leftrightarrow x \in \left(f^{-1}(B)\right)^c$

5.2 线性空间的同构

Definition 5.4

数域 P 上两个线性空间 V和V' 称为 (线性) 同构的 $\Leftrightarrow$ 存在一个双射 σ 并且满足① $\sigma(\alpha+\beta)=\sigma(\alpha)+\sigma(\beta)$ ② $\sigma(k\alpha)=k\sigma(\alpha)$ 这样的的映射 σ 我们称之为同构映射, 显然数域 P 上的任何一个 n 维线性空间都与 P^n 同构, 显然有（实际上这一点由线性映射就保证了并不需要双射）$\sigma(0)=0; \sigma(-\alpha)=-\sigma(\alpha)$显然线性空间$V\to V$的恒等映射$Id_V$；也是线性同构。

Lemma 5.1

若$\varphi: V\to U$的线性映射，那么若φ是同构 $\Rightarrow \varphi^{-1}$是$U\to V$的同构

Proof 显然φ^{-1}为双射，因为双射等价条件是存在逆映射，那么 $(\varphi^{-1})^{-1}=\varphi$; 则$\varphi^{-1}$为双射
其次证明其线性。要证$\varphi^{-1}(k\alpha+l\beta)=k\varphi^{-1}(\alpha)+l\varphi^{-1}(\beta)$。即证$\varphi^{-1}(k\alpha+l\beta)-k\varphi^{-1}(\alpha)-l\varphi^{-1}(\beta)=0$。
我们用φ作用。那么得到$\varphi(LHS)=0_U$； 而$\varphi(0_V)=0_U$; 但是φ为双射，那么只有$LHS=0_V$证毕 □

Theorem 5.2.1

(I) V中向量组$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r$线性相关的充分必要条件是, 它们的像$\sigma(\boldsymbol{\alpha}_1),\sigma(\boldsymbol{\alpha}_2),\cdots,\sigma(\boldsymbol{\alpha}_r)$ 线性相关.
(II) 因此得到同构的线性空间具有相同的维数
(III) 已知$V\cong V'$显然若$V_1\subseteq V\Rightarrow V_1'\subseteq V'$且$V_1\cong V_1'$
(VI) : 显然同构映射的逆映射和乘积仍然为同构映射
(V) 显然同构作为线性空间中的一种等价关系具有自反性，对称性，传递性
(VI) 数域P上两个有限维线性空间同构 $\Leftrightarrow$ dim相等

Proof (I)***Proof*** : 因为由$k_1\alpha_1+k_2\alpha_2+\cdots+k_r\alpha_r=0$可得$k_1\sigma(\alpha_1)+k_2\sigma(\alpha_2)+\cdots+k_r\sigma(\alpha_r)=0$
反过来, 由$k_1\sigma(\alpha_1)+k_2\sigma(\alpha_2)+\cdots+k_r\sigma(\alpha_r)=0$有$\sigma(k_1\alpha_1+k_2\alpha_2+\cdots+k_r\alpha_r)=0$. 因为$\sigma$是$1-1$的, 只有$\sigma(0)=0$,
所以$k_1\alpha_1+k_2\alpha_2+\cdots+k_r\alpha_r=0$. □

Proof (V) : 自反性 : 显然$V\to V$的恒等映射就是一个同构, 所以$V\cong V$
对称性 : 若$\varphi: V\to U$, 为同构 $(V\cong U)$, 那么根据性质。φ^{-1}就是$U\to V$的同构。
传递性 : 根据上一节的性质即可

Proof (VI) 先证必要性，$\varphi: V\to U$
取V的一组基$e_1\cdots e_n$； 由 (I) 知, $\varphi(e_1)\ \varphi(e_n)$ 也是线性无关的. 此时我们只需要证对于$\forall x\in U$. 其均可以被$\varphi(e_1)\ \varphi(e_n)$ 表示。
因为是同构是双射, 所以必定存在$\alpha\in V$; $st.\varphi(\alpha)=x$. 此时还可以设$\alpha=\alpha_1e_1+\cdots+\alpha_ne_n$
用φ作用, 得$x=\alpha_1\varphi(e_1)+\cdots+\alpha_n\alpha(e_n)$; 从而推出$\varphi(e_1)\ \varphi(e_n)$ 也是U的一组基。那么$dimV = dimU$
再证充分性。即若$dimV=dimU$。那么二者之间一定存在一个线性同构
取V基 $\{e_1\ e_n\}$; 取U基 $\{f_1\ f_n\}$. 知道$\varphi_V: V\to K_n$; $\varphi_U: U\to K_n$都是线性同构
那么$\varphi_U^{-1}: K_n\to U$那么$\varphi_V\varphi_U^{-1}: V\to U$的线性同构。
注 注意 : 若$dimV=dimU$说明二者之间存在线性同构，也许有很多种。
其次若$V\cong U$，那么$\varphi: V\to U$不可以说是都是线性同构 (eg : 零映射)。所以我们记$V\cong U(\varphi)$ 这是最好的。

5.3 线性映射构成空间

Definition 5.5

设φ,ψ是K上的线性空间$V\to U$的线性映射。若$k\in$ K,

定义$\varphi+\psi$为$V\to U$的映射 : $(\varphi+\psi)(\alpha)=\varphi(\alpha)+\psi(\alpha),\alpha\in V$.

定义$k\varphi$为$V\to U$的映射 : $(k\varphi)(\alpha)=k\varphi(\alpha),\alpha\in V$

设$\mathscr{L}(V,U)$指的是$V\to U$的全体线性映射构成的全体集合。

Theorem 5.3.1

由上述定义的$\varphi+\psi$与$k\varphi$都仍然是$V\to U$的线性映射

Theorem 5.3.2

$\mathscr{L}(V,U)$在我们所定义的加法和数乘下仍然是数域K上的线性空间。

注 当$U=K$时，$\mathscr{L}(V,K)$称为V上线性函数全体构成的线性空间。记为V^*。称为V的共轭空间(对偶空间)

当$U=V$时，$\mathscr{L}(V,V)$就变成了V上全体线性变换全体构成的线性空间，简记为$\mathscr{L}(V)$；显然在$\mathscr{L}(V)$上还有复合运算

Definition 5.6

设A是数域K上的线性空间,如果在A上定义了一个乘法.(通常可以省略),使对任意A中元素a,b,c及K中元素k,适合下列条件

(1) 乘法结合律 : $a\cdot(b\cdot c)=(a\cdot b)\cdot c$;

(2) 恒等元存在 : 存在A中元e,使对一切$a\in A$,均有 : $e\cdot a=a\cdot e=a$

(3) 分配律 : $a\cdot(b+c)=a\cdot b+a\cdot c \quad (b+c)\cdot a=b\cdot a+c\cdot a$

(4) 乘法相容性 : $(ka)\cdot b=k(a\cdot b)=a\cdot(kb)$

则称A是数域 K 上的代数,元素e称为A的恒等元.

Theorem 5.3.3

设V是数域K 上的线性空间,则$\mathcal{L}(V)$是K上的代数.(此时乘法定义为复合)

Proof 乘法结合律实际上就是映射的结合律,因此自然成立

设1_V是V上的恒等映射,由上节可知它是线性变换显然对任意的$\varphi\in\mathscr{L}(V)$有$1_V\cdot\varphi=\varphi\cdot 1_V=\varphi$因此$1_V$是$\mathcal{L}(V)$的恒等元

设$\varphi_1,\varphi_2,\varphi_3$都是$V$上线性变换,对任意的$\alpha\in V$,有

$$\begin{aligned}(\varphi_1\circ(\varphi_2+\varphi_3))(\alpha)&=\varphi_1((\varphi_2+\varphi_3)(\alpha))\\&=\varphi_1(\varphi_2(\alpha)+\varphi_3(\alpha))\\&=\varphi_1(\varphi_2(\alpha))+\varphi_1(\varphi_3(\alpha))\\&=(\varphi_1\circ\varphi_2)(\alpha)+(\varphi_1\circ\varphi_3)(\alpha)\\&=(\varphi_1\circ\varphi_2+\varphi_1\circ\varphi_3)(\alpha).\end{aligned}$$

因此$\varphi_1\circ(\varphi_2+\varphi_3)=\varphi_1\circ\varphi_2+\varphi_1\circ\varphi_3$

同理可以证明另一个分配律

设k是K中任一数,φ,ψ为V上线性变换,则对V中任意的α,有

$$\begin{aligned}[(k\varphi)\cdot\psi](\alpha)&=(k\varphi)(\psi(\alpha))=k[\varphi(\psi(\alpha))]\\&=k((\varphi\cdot\psi)(\alpha))=(k(\varphi\cdot\psi))(\alpha)\end{aligned}$$

从而$(k\varphi)\cdot\psi=k(\varphi\cdot\psi)$同理可证明$\varphi\cdot(k\psi)=k(\varphi\cdot\psi)$证毕.

性质 在$\mathscr{L}(V)$中,定义线性变换φ的n次幂为n个φ之积,则不难验证 : $\varphi^n\cdot\varphi^m=\varphi^{n+m},(\varphi^n)^m=\varphi^{nm}$ (*)

若φ是双射,即为V上的自同构,则φ^{-1}也是V上的线性映射(也是自同构),称φ^{-1}为φ的逆变换

如定义$\varphi^{-n}=(\varphi^{-1})^n$ 则不难验证 : $\varphi^{-n}=(\varphi^n)^{-1}$ 这时定义$\varphi^0=I_V$

则 (*) 式对一切整数均成立.但需注意φ的负数次幂仅对自同构 (又称可逆变换或非异变换) 有意义.

如果φ与ψ都是可逆线性映射, 则$\varphi \cdot \psi$也是可逆线性映射, 且 $(\varphi \cdot \psi)^{-1} = \psi^{-1} \cdot \varphi^{-1}$

对任一非零数k, 若φ可逆, 则$k\varphi$也可逆, 且 $(k\varphi)^{-1} = k^{-1}\varphi^{-1}$

5.4 线性映射与矩阵

下面两种方式给出了如何构造两个线性空间中的线性映射

Theorem 5.4.1 (线性扩张定理)

设有 K 上线性空间V和U, $\{e_1, e_2, \cdots, e_m\}$是$V$的一组基有如下结论

(1) 如有V到U的线性映射φ和ψ, 对任意的i, 都有$\psi(\boldsymbol{e}_i) = \varphi(\boldsymbol{e}_i)$, 则$\psi = \varphi$

(2) 给定U中m个向量$\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_m$, 有且只有一个从$V$到$U$的线性映射$\varphi$, 满足$\varphi(\boldsymbol{e}_i) = \boldsymbol{\beta}_i (i = 1, 2, \cdots, m)$

该定理表明研究两个线性空间的映射的性质没必要研究所有向量只需要研究基向量即可

换一句话说，在两个线性空间中构造一个映射只需要考虑基向量，剩余的用扩张定理即可

Proof (1)对任意的$\alpha \in V$, 设$\alpha = \lambda_1 e_1 + \lambda_2 e_2 + \cdots + \lambda_m e_m$, 则

$$
\begin{aligned}
\psi(\alpha) &= \psi(\lambda_1 \boldsymbol{e}_1 + \lambda_2 \boldsymbol{e}_2 + \cdots + \lambda_m \boldsymbol{e}_m) \\
&= \lambda_1 \psi(\boldsymbol{e}_1) + \lambda_2 \psi(\boldsymbol{e}_2) + \cdots + \lambda_m \psi(\boldsymbol{e}_m) \\
&= \lambda_1 \varphi(\boldsymbol{e}_1) + \lambda_2 \varphi(\boldsymbol{e}_2) + \cdots + \lambda_m \varphi(\boldsymbol{e}_m) \\
&= \varphi(\lambda_1 \boldsymbol{e}_1 + \lambda_2 \boldsymbol{e}_2 + \cdots + \lambda_m \boldsymbol{e}_m) \\
&= \varphi(\alpha),
\end{aligned}
$$

因此$\psi = \varphi$

(2)定义$V \longrightarrow U$的映射φ如下：对V中任意的$\boldsymbol{\alpha} = \lambda_1 \boldsymbol{e}_1 + \lambda_2 \boldsymbol{e}_2 + \cdots + \lambda_m \boldsymbol{e}_m$, $\varphi(\alpha) = \lambda_1 \boldsymbol{\beta}_1 + \lambda_2 \boldsymbol{\beta}_2 + \cdots + \lambda_m \boldsymbol{\beta}_m$

容易验证这是V到U的线性映射, 且$\varphi(\boldsymbol{e}_i) = \boldsymbol{\beta}_i (i = 1, 2, \cdots, m)$.唯一性由 (1)即得.证毕.

Theorem 5.4.2 (直和扩张定理)

设线性空间$V = V_1 \oplus V_2$, 已知φ_1及φ_2分别是V_1, V_2到U的线性映射, 则存在从V到U唯一的线性映射φ, 当φ限制在V_i上时等于φ_i.

Proof 对任意的$\boldsymbol{\alpha} \in V$, 因为$V = V_1 \oplus V_2$, $\boldsymbol{\alpha}$可唯一地写为$\boldsymbol{\alpha} = \boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_2$, $\alpha_1 \in V_1$, $\alpha_2 \in V_2$.

令$\varphi(\alpha) = \varphi_1(\alpha_1) + \varphi_2(\alpha_2)$, 则$\varphi$是$V$到$U$的映射.不难验证$\varphi$保持加法和数乘, 因此$\varphi$是线性映射.

若另有线性映射ψ, 它在V_i上的限制等于φ_i, 则$\psi(\alpha) = \psi(\alpha_1) + \psi(\alpha_2) = \varphi_1(\alpha_1) + \varphi_2(\alpha_2) = \varphi(\alpha)$.

因此$\psi = \varphi$, 唯一性得证. □

Proposition 5.1

设有数域$\mathbb{F}$上线性空间V及V', 又U是V的子空间, φ是线性空间U到V'的线性映射.

求证：必存在V到V'的线性映射ψ, 它在U上的限制就是φ.

Proof 令W是子空间U在V中的补空间, 即$V = U \oplus W$.

定义V到V'的映射$\boldsymbol{\psi}$如下：其在U上的限制是φ, 在W上的限制为零线性映射

不难验证ψ是所需之线性映射. □

Definition 5.7

下面考虑若V和U是数域$\mathbb{K}$上的n维与m维线性空间；设$\{\boldsymbol{e}_1\sim\boldsymbol{e}_n\};\{\boldsymbol{f}_1\sim\boldsymbol{f}_m\}$为$V$和$U$的基。设$\boldsymbol{\varphi}:V\to U$

$$\text{已知}\begin{cases}\boldsymbol{\varphi}(\boldsymbol{e}_1)=a_{11}\boldsymbol{f}_1+a_{21}\boldsymbol{f}_2+\cdots+a_{1m}\boldsymbol{f}_m\\ \boldsymbol{\varphi}(\boldsymbol{e}_2)=a_{21}\boldsymbol{f}_1+a_{22}\boldsymbol{f}_2+\cdots+a_{2m}\boldsymbol{f}_m\\ \cdots\cdots\cdots\\ \boldsymbol{\varphi}(\boldsymbol{e}_n)=a_{n1}\boldsymbol{f}_1+a_{n2}\boldsymbol{f}_2+\cdots+a_{nm}\boldsymbol{f}_m\end{cases}\Leftrightarrow(\varphi(e_1),\varphi(e_2)\cdots\varphi(e_n))=(f_1,f_2\cdots f_m)A$$

其中$A=(a_{ij})_{m\times n}$称为$\varphi:V\to U$在给定基下的表示矩阵，为系数矩阵的转置

Lemma 5.2

那么任一$\alpha\in V;\alpha=\sum_{i=1}^{n}\lambda_ie_i$ 故$\varphi(\alpha)=\left(\sum_{i=1}^{n}\lambda_ia_{i1}\right)\boldsymbol{f}_1+\left(\sum_{i=1}^{n}\lambda_ia_{i2}\right)\boldsymbol{f}_2+\cdots+\left(\sum_{i=1}^{n}\lambda_ia_{im}\right)\boldsymbol{f}_m$

表明$\varphi(\alpha)$在$\{f_1\cdots f_m\}$下的坐标向量

$$\begin{pmatrix}\mu_1\\ \mu_2\\ \vdots\\ \mu_m\end{pmatrix}=\begin{pmatrix}a_{11}&a_{21}&\cdots&a_{n1}\\ a_{12}&a_{22}&\cdots&a_{n2}\\ \vdots&\vdots&&\vdots\\ a_{1m}&a_{2m}&\cdots&a_{nm}\end{pmatrix}\begin{pmatrix}\lambda_1\\ \lambda_2\\ \vdots\\ \lambda_n\end{pmatrix}$$

$$\text{即}\begin{pmatrix}\mu_1\\ \mu_2\\ \vdots\\ \mu_m\end{pmatrix}=A_{m\times n}\begin{pmatrix}\lambda_1\\ \lambda_2\\ \vdots\\ \lambda_n\end{pmatrix}$$

Definition 5.8

于上方讨论我们得到了一个映射

$$T:\begin{array}{cc}\mathscr{L}(V^n,U^m)\to & M_{m\times n}(\mathbb{K})\\ \varphi\longmapsto & A\end{array}$$

Theorem 5.4.3

上述定义的T,线性映射空间$\to M_{m\times n}$ 该映射是两个空间的同构。

Proof 为了证明该映射是线性同构。我们不仅证明是双射而且满足线性即可

规定记号$T:\begin{array}{l}\mathscr{L}(V^n,U^m)\to M_{m\times n}(\mathbb{K})\\ \varphi\longmapsto A\end{array}$；其中$\{e_1\sim e_n\},\{f_1\sim f_m\}$为$V$，$U$的基

①证明满性：

任取$A_{m\times n}\in M_{m\times n}(\mathbb{K})$。

我们根据线性扩张定理来构造一个线性映射$\varphi\ st.\begin{cases}\varphi(e_1)=a_{11}f_1+a_{21}f_2+\cdots+a_{m1}f_m\\ \varphi(e_2)=a_{12}f_1+a_{22}f_2+\cdots+a_{m2}f_m\\ \cdots\cdots\cdots\\ \varphi(e_n)=a_{1n}f_1+a_{2n}f_2+\cdots+a_{mn}f_m\end{cases}$

那么$\varphi:V\to U$;由定义知道φ的表示矩阵就为A。

②证明单性

设$\varphi,\psi\in\mathscr{L}(V,U)$。若其表示矩阵均为$A$；从而$\varphi,\psi$都满足$\begin{cases}\varphi(e_1)=a_{11}f_1+a_{21}f_2+\cdots+a_{m1}f_m\\ \varphi(e_2)=a_{12}f_1+a_{22}f_2+\cdots+a_{m2}f_m\\ \cdots\cdots\cdots\\ \varphi(e_n)=a_{1n}f_1+a_{2n}f_2+\cdots+a_{mn}f_m\end{cases}$

再由线性扩张定理得到$\varphi=\psi$

③证明线性加法

设$\varphi,\psi\in\mathscr{L}(V,U)$；$T(\varphi)=A=\left(a_{ij}\right)_{m\times n}$;$T(\psi)=B=\left(b_{ij}\right)_{m\times n}$

即$\varphi\left(e_j\right)=\sum\limits_{i=1}^{m}a_{ij}f_i$；$\psi\left(e_j\right)=\sum\limits_{i=1}^{m}b_{ij}f_i$对于$1\leqslant j\leqslant n$都成立

$(\varphi+\psi)\left(e_j\right)=\varphi\left(e_j\right)+\psi\left(e_j\right)=\sum\limits_{i=1}^{m}\left(a_{ij}+b_{ij}\right)f_i$即$T(\varphi+\psi)=\left(a_{ij}+b_{ij}\right)_{m\times n}=(A+B)=T(\varphi)+T(\psi)$

④证明线性数乘

$\forall k\in\mathbb{K},(k\varphi)\left(e_j\right)=k\varphi\left(e_j\right)=\sum\limits_{i=1}^{m}ka_{ij}f_i\Rightarrow T(k\varphi)=\left(ka_{ij}\right)_{m\times n}=kA=kT(\varphi)$

以上证毕该映射为两个空间的线性同构

Theorem 5.4.4

规定如下记号。证明图为交换图

V^n $\{e_1 \sim e_n\}$ 基，$\eta_V : V \to K^n$坐标向量线性同构

U^m $\{f_1 \sim f_m\}$ 基，$\eta_U : U \to K^m$坐标向量线性同构

$A \in M_{m\times n}(K)$ 为表示矩阵 $\varphi(A) : \begin{array}{c} K^n \to K^m \\ x \mapsto Ax \end{array}$ $\varphi : V \to U$

$$
\begin{tikzcd}
V \arrow[r, "\varphi"] \arrow[d, "\eta_V"'] & U \arrow[d, "\eta_U"] \\
K_n \arrow[r, "\varphi_A"'] & K_m
\end{tikzcd}
$$

Proof

$$
\begin{tikzcd}
\forall \alpha \in V \arrow[rrr, "\varphi"] \arrow[d, "\eta_V"'] & & & \varphi(\alpha) \in U \arrow[d, "\eta_U"] \\
\begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix} \in K^n \arrow[r, "\varphi_A"'] & A \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix} & = & \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_m \end{pmatrix} \in K^m
\end{tikzcd}
$$

特别的我们还要考虑线性变换的情形

Definition 5.9

设$\varphi \in \mathscr{L}(V)$, 取定$V$的一组基$\{e_1 \sim e_n\}$; (此时与映射不同，我们取的基仅仅是一组$\{e_1 \sim e_n\}$)

此时 $\begin{cases} \varphi(e_1) = a_{11}e_1 + a_{21}e_2 + \cdots + e_{n1}e_n \\ \varphi(e_2) = a_{12}e_1 + a_{22}e_2 + \cdots + e_{n2}e_n \\ \cdots\cdots\cdots \\ \varphi(e_n) = a_{1n}e_1 + a_{2n}e_2 + \cdots + e_{nn}e_n \end{cases}$ $\Leftrightarrow (\varphi(e_1), \varphi(e_2) \cdots \varphi(e_n)) = (e_1, e_2 \cdots e_n)A \ \ A = (a_{ij})_{n\times n}$

称A为线性变换下的表示矩阵

那么自然满足上面的一般定理即

$T : \begin{array}{c} \mathscr{L}(V) \to M_{n\times n}(K) \\ \varphi \longmapsto \quad A \end{array}$ 是线性同构, 更一般的更是 K 代数同构如下证明

Corollary 5.1

(1) $T(I_V) = I_n$

(2) $\varphi \in \mathscr{L}(V)$; 则$\varphi$是自同构 $\Leftrightarrow T(\varphi)$ 是可逆阵且$T(\varphi^{-1}) = T(\varphi)^{-1}$

我们知道线性变换全体是 K 上代数，且 n 阶方阵的全体也是代数。那么上述的映射是否保持了该乘法呢？为此我们先证明一个更一般的映射的结论如下

Theorem 5.4.5

设$\varphi : V^n \to U^m$;$\psi : U^m \to W^p$线性映射。$\psi \circ \varphi : V^n \to W$线性映射

$\Rightarrow$ 则$T(\psi \circ \varphi) = T(\varphi) \cdot T(\psi)$

Proof 设V的基$\{e_1, e_2 \cdots e_n\}$;U的基$\{f_1, f_2 \cdots f_m\}$;W的基$\{g_1, g_2 \cdots g_p\}$

$T(\varphi) = A = (a_{ij})_{m\times n}$;$T(\psi) = B = (b_{ij})_{p\times m}$

只要证 : $T(\psi \circ \varphi) = BA$

由题干 : $(\varphi(e_1), \varphi(e_2), \cdots \varphi(e_n)) = (f_1, f_2 \cdots f_m) A \Leftrightarrow \varphi(e_j) = \sum_{i=1}^{m} a_{ij} f_i \ (\forall 1 \leqslant j \leqslant n)$

我们用ψ来作用上面，那么$\psi(\varphi(e_j)) = \psi\left(\sum_{i=1}^{m} a_{ij} f_j\right) = \sum_{i=1}^{m} a_{ij}\psi(f_j)$ $(\forall 1 \leqslant j \leqslant n)$

$\Leftrightarrow [\psi(\varphi(e_1)), \psi(\varphi(e_2)), \cdots \psi(\varphi(e_n))] = [\psi(f_1), \psi(f_2) \cdots \psi(f_m)] A$(这也就是说对形式行向量作用线性映射其系数不改变)

且$[\psi(f_1), \psi(f_2) \cdots \psi(f_m)] = (g_1, g_2 \cdots g_p) B$

$\Rightarrow [\psi(\varphi(e_1)), \psi(\varphi(e_2)), \cdots \psi(\varphi(e_n))] = (g_1, g_2 \cdots g_p) BA$

那么根据表示矩阵的定义得到$T(\psi \circ \varphi) = BA = T(\psi) T(\varphi)$

证毕

下面给出上面的性质的证明

Proof

(1) : V的一组基$\{e_1 \sim e_n\}$; $I_V(e_i) = e_i \Rightarrow T(I_V) = I_n$

(2) : 必要性

假设φ是自同构(即线性变换 + 双射)。其一定有逆映射φ^{-1}也为同构(参照线性空间的同构一节)

$I_V = \varphi\varphi^{-1} = \varphi^{-1}\varphi$用$T$作用那么 $\Rightarrow I_n = T(I_V) = T(\varphi\varphi^{-1}) = T(\varphi) T(\varphi^{-1})$

这就得到了$I_n = T(\varphi) T(\varphi^{-1})$又根据$T(\varphi)$是$n$阶方阵，所以$T(\varphi)$为可逆阵且逆阵为$T(\varphi^{-1})$

再证充分性

设$T(\varphi)$可逆；即$T(\varphi)^{-1}$存在

而$T : \mathscr{L}(V) \to M_n(\mathbb{K})$是一个双射那么$\exists \psi \in \mathscr{L}(V)$ $st.\psi \longmapsto T(\varphi)^{-1}$

即$T(\psi) = T(\varphi)^{-1}$ 即$I_n = T(\psi) T(\varphi) = T(\varphi) T(\psi)$ 即$T(I_V) = T(\psi\varphi) = T(\varphi\psi)$

由单性$I_V = \varphi\psi = \psi\varphi$所以$\psi = \varphi^{-1}$； φ是自同构

在上文中我们讲述了线性映射和矩阵具有一一对应的关系我们既可以吧线性映射化为矩阵问题来解决也可以把矩阵化为线性映射来解决

反之, 我们也可将矩阵问题转化为线性映射 (线性变换) 问题来处理.

设$\boldsymbol{A}$是数域$\mathbb{F}$上的$m \times n$矩阵, 定义列向量空间$\mathbb{F}^n$到$\mathbb{F}^m$的线性映射 : $\varphi(\alpha) = A\alpha$

容易验证在$\mathbb{F}^n$和$\mathbb{F}^m$的标准单位列向量构成的基下, φ的表示矩阵就是$\boldsymbol{A}$.

同理, 若$\boldsymbol{A}$是$\mathbb{F}$上的n阶矩阵, 定义$\mathbb{F}^n$上的线性变换 : $\varphi(\alpha) = A\alpha$, 容易验证在$\mathbb{F}^n$的标准单位列向量构成的基下, $\boldsymbol{\varphi}$的表示矩阵就是$\boldsymbol{A}$.

因此, 我们有时就把这个线性映射 (线性变换) 写为$\boldsymbol{A}$.

5.5 线性变换即相似关系构建

性质 设$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$是数域$P$上$n$维线性空间$V$的一组基, 在这组基下, 每个线性变换按公式对应一个$n \times n$矩阵.
这个对应具有以下的性质 :
1) 线性变换的和对应于矩阵的和;
2) 线性变换的乘积对应于矩阵的乘积;
3) 线性变换的数量乘积对应于矩阵的数量乘积
4) 可逆的线性变换与可逆矩阵对应, 且逆变换对应于逆矩阵.

Theorem 5.5.1

设线性空间V中线性变换$\mathscr{A}$在两组基 $\begin{matrix} \boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \cdots, \boldsymbol{\varepsilon}_n, \\ \boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \cdots, \boldsymbol{\eta}_n \end{matrix}$ 下的矩阵分别为$\boldsymbol{A}$和$\boldsymbol{B}$,
从基ε到η的过渡矩阵是$\boldsymbol{X}$, 于是$\boldsymbol{B} = \boldsymbol{X}^{-1}\boldsymbol{A}\boldsymbol{X}$.

Proof 证明已知 $(\mathscr{A}\boldsymbol{\varepsilon}_1, \mathscr{A}\boldsymbol{\varepsilon}_2, \cdots, \mathscr{A}\boldsymbol{\varepsilon}_n) = (\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \cdots, \boldsymbol{\varepsilon}_n)\boldsymbol{A}$, $(\mathscr{A}\boldsymbol{\eta}_1, \mathscr{A}\boldsymbol{\eta}_2, \cdots, \mathscr{A}\boldsymbol{\eta}_n) = (\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \cdots, \boldsymbol{\eta}_n)\boldsymbol{B}$,
$(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \cdots, \boldsymbol{\eta}_n) = (\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \cdots, \boldsymbol{\varepsilon}_n)X$.

于是
$$\begin{aligned} &(\mathscr{A}\boldsymbol{\eta}_1, \mathscr{A}\boldsymbol{\eta}_2, \cdots, \mathscr{A}\boldsymbol{\eta}_n) \\ =&\mathscr{A}(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \cdots, \boldsymbol{\eta}_n) = \mathscr{A}[(\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \cdots, \boldsymbol{\varepsilon}_n)X] \\ =&[\mathscr{A}(\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \cdots, \boldsymbol{\varepsilon}_n)]X = (\mathscr{A}\boldsymbol{\varepsilon}_1, \mathscr{A}\boldsymbol{\varepsilon}_2, \cdots, \mathscr{A}\boldsymbol{\varepsilon}_n)X \\ =&(\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \cdots, \boldsymbol{\varepsilon}_n)AX = (\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \cdots, \boldsymbol{\eta}_n)X^{-1}AX. \end{aligned}$$
由此即得$B = X^{-1}AX$.

Definition 5.10

设A，B是n阶方阵，若存在可逆矩阵P；$st.B = P^{-1}AP$
则称A，B相似。记作$A \approx B$

性质 相似是一种等价关系

5.6 线性映射的像与核

5.6.1 有限维空间线性映射像与核

Definition 5.11

设φ是数域K上向量空间V到U的线性映射,φ的全体像元素组成U的子集称为φ的像, 记为 $\mathrm{Im}\varphi$.
又V中在φ下映射为零向量的全体向量构成V的子集, 称为φ的核, 记为 $\mathrm{Ker}\varphi$.
即 $\mathrm{Im}\varphi = \left\{\varphi(v) \middle| v \in V\right\} \subseteq U \quad Ker\varphi = \left\{v \in V \middle| \varphi(v) = 0_U\right\} \subseteq V$
显然二者都非空都有零向量元素在

Theorem 5.6.1 (有限维一般空间线性映射单满射的维数判别法)

$\varphi : V^n \to U^m$为线性映射
(1) : φ为满射 $\Leftrightarrow dim\mathrm{Im}\varphi \;=\; dimU$
(2) : φ为单射 $\Leftrightarrow Ker\varphi = \{0\} \Leftrightarrow dim\ Ker\varphi = 0$

Proof (1) 根据满射定义有,φ是满射 $\Leftrightarrow \mathrm{Im}\varphi = U \underbrace{\Leftrightarrow dim\mathrm{Im}\varphi \;=\; dimU}_{\text{因为 }\mathrm{Im}\varphi\text{是}U\text{子空间}}$
(2) $\Rightarrow$: 任取$\alpha \in Ker\varphi$; 那么有$\varphi(\alpha) = 0_U = \varphi(0_V)$ 又φ为单射
$\Rightarrow \alpha = 0_V$; 从而$Ker\varphi = \{0\}$
$\Leftarrow$: 设$\varphi(\alpha) = \varphi(\beta)$; 其中$\alpha, \beta \in V$;
得到$\varphi(\alpha - \beta) = 0 \Rightarrow \alpha - \beta \in Ker\varphi \Rightarrow \alpha - \beta = 0 \Rightarrow \alpha = \beta \Rightarrow$ 为单射

Definition 5.12 (线性映射的秩与零度)

设φ是$V \to U$的线性映射.称像空间 $\mathrm{Im}\varphi$的维数为φ的秩, 记作$r(\varphi)$.核空间 $\mathrm{Ker}\varphi$的维数称为φ的零度.
即$dim\mathrm{Im}\varphi \;=\; r(\varphi) \quad ; \; dim\ Ker\varphi =$ 零度

Lemma 5.3 (线性映射的限制与保持单射)

设$\varphi : V^n \to U^m$为线性映射; V'为V子空间; U'为U子空间; 满足$\varphi(V') \subseteq U'$
则可通过限制得到线性映射$\varphi' : V' \to U'$
使得φ'与φ有相同的映射法则, 称φ'是φ在V'上的限制。
且若φ为单射那么φ'也是单射

Proof 定义$\varphi' : \begin{array}{l} V' \to U' \\ v' \longmapsto \varphi(v') \end{array} \quad \varphi(v') \in U'$这样的$\varphi'$与$\varphi$有相同的映射法则。
下面我们来验证φ'是线性映射
任取$v_1'; v_2' \in V'; k \in K$
$\varphi'\left(v_1' + v_2'\right) = \varphi\left(v_1' + v_2'\right) = \varphi\left(v_1'\right) + \varphi\left(v_2'\right) = \varphi'\left(v_1'\right) + \varphi'\left(v_2'\right)$
$\varphi'\left(kv_1'\right) = \varphi\left(kv_1'\right) = k\varphi\left(v_1'\right) = k\varphi'\left(v_1'\right)$
这就证明了φ'是线性映射
又若φ是单射我们下证φ'也是单射。所以只要证 : $Ker\varphi' = \{0\}$
φ单射 $\Rightarrow Ker\varphi = \{0\}$
而$Ker\ \varphi' = \{v' \in V' | \varphi'(v') = 0\} = \{v' \in V' | \varphi(v') = 0\} = Ker\ \varphi \cap V'$
$\Rightarrow Ker\ \varphi' = \{0\} \Rightarrow \varphi'$一定是单射

Theorem 5.6.2 (核空间与像空间同构定理)

设$\varphi : V^n \to U^m$；设$\{e_1 \sim e_n\}\{f_1 \sim f_m\}$为$V$和$U$的基。$\varphi$在给定基下的矩阵为$A_{m\times n}$

$\Rightarrow$ 则$rank\left(\varphi\right) = dim\mathrm{Im}\varphi = rank\left(A\right)$ $\quad dim\ Ker\varphi = n - rank\left(A\right)$

$Ker(\varphi)$与$Ax = 0$解空间同构 $\qquad Im(\varphi)$与A列向量张成的空间同构

$$\begin{tikzcd} V \arrow[r, "\varphi"] \arrow[d, "\eta_V"'] & U \arrow[d, "\eta_U"] \\ K_n \arrow[r, "\varphi_A"'] & K_m \end{tikzcd}$$

Proof

$Step1 \quad \eta_V\left(Ker\varphi\right) \subseteq Ker\left(\varphi_A\right) \quad \eta_U\left(\mathrm{Im}\varphi\right) \subseteq \mathrm{Im}\left(\varphi_A\right)$

$1^\circ\forall v \in Ker\varphi,$ 只要证$\eta_V\left(v\right) \in Ker\left(\varphi_A\right) \Leftrightarrow \varphi_A\left(\eta_V\left(v\right)\right) = 0$

通过交换图的性质我们有$\eta_U \circ \varphi = \varphi_A \circ \eta_V$

所以$\varphi_A\left(\eta_V\left(v\right)\right) = \eta_U\left(\varphi\left(v\right)\right) = \eta_U\left(0_U\right) = 0$

$2^\circ\forall\varphi\left(v\right) \in \mathrm{Im}\left(\varphi\right), v \in V$；只要证$\eta_U\left(\varphi\left(v\right)\right) \in \mathrm{Im}\left(\varphi_A\right)$

$\eta_U\left(\varphi\left(v\right)\right) = \varphi_A\left(\eta_V\left(v\right)\right) \in \mathrm{Im}\left(\varphi_A\right)$

$Step2$ 　做限制$\eta_V : Ker\varphi \to Ker\varphi_A \quad \eta_U : \mathrm{Im}\varphi \to \mathrm{Im}\varphi_A$

那么做了限制之后的η_V，η_U也是线性映射且是单的根据上文引理保持单性

那么接下来证明这个线性映射还是满性。

$1^\circ\eta_V : Ker\varphi \to Ker\varphi_A$是满射

任取$x \in Ker\varphi_A$; 即$\varphi_A\left(x\right) = 0$; $\ x \in Ker\varphi_A \subseteq K^n$根据$\eta_V$是线性同构具有满性

所以一定$\exists v \in V; st.\eta_V\left(v\right) = x$; 那么接下来只要证$v \in Ker\varphi$即可

由交换图有$\eta_U\left(\varphi\left(v\right)\right) = \varphi_A\left(\eta_V\left(v\right)\right) = \varphi_A\left(x\right) = 0$又$\eta_U$是线性同构那么$\varphi\left(v\right) = 0$

$\Rightarrow v \in Ker\varphi$证毕

$2^\circ\eta_U : \mathrm{Im}\varphi \to \mathrm{Im}\varphi_A$

任取$\varphi_A\left(x\right) \in \mathrm{Im}\varphi_A, x \in K^n$根据$\eta_V$是线性同构具有满性

所以一定有$v \in V\ \ st.\eta_V\left(v\right) = x$; 那么$\varphi\left(v\right) \in U$

$\eta_U\left(\varphi\left(v\right)\right) = \varphi_A\left(\eta_V\left(v\right)\right) = \varphi_A\left(x\right)$ 从而也是满射

$Step3$

$Ker\varphi_A = \{x \in K^n | Ax = 0\}$ 即$Ker\varphi_A$为$Ax = 0$的解空间

根据$Ker\varphi \cong Ker\varphi_A \Rightarrow dim\ Ker\varphi\ =\ dim\ Ker\varphi_A\ =\ n - rank\left(A\right)$

做列分块$A_{m\times n} = \left(\alpha_1, \alpha_2 \cdots \alpha_n\right)$ 其中$\alpha_i \in K^m \quad \mathrm{Im}\varphi_A = \{Ax | x \in K^n\} \ \ x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in K^n$

$\Rightarrow \mathrm{Im}\varphi_A = \{x_1\alpha_1 + x_2\alpha_2 + \cdots x_n\alpha_n\}$ 这就是A的列向量张成的子空间

$\mathrm{Im}\varphi \cong \mathrm{Im}\varphi_A \Rightarrow r\left(\varphi\right) = dim\mathrm{Im}\varphi = dim\mathrm{Im}\varphi_A = dimL\left(\alpha_1, \alpha_2 \cdots \alpha_n\right) = rank\left(A\right)$

下面给出几个重要推论定理

Theorem 5.6.3 (有限维一般空间线性映射的维数公式)

设$\varphi: V^n \to U^m \Rightarrow$ 则$dim\ Ker\varphi + dim\mathrm{Im}\varphi \ = \ dimV$

Theorem 5.6.4 (有限维一般空间线性映射单满射的矩阵判定法)

$\varphi: V^n \to U^m$ (线性映射), 在给定基下的矩阵为$A_{m\times n}$

$$\Rightarrow \text{则} \begin{cases} \varphi\text{满射} \Leftrightarrow A\text{行满秩} \Leftrightarrow r(A) = m \\ \varphi\text{单射} \Leftrightarrow A\text{列满秩} \Rightarrow r(A) = n \end{cases}$$

Corollary 5.2 (有限维维数相同空间的单满射判定（线性变换）)

$\varphi: V^n \to U^n$为线性映射且$dimV \ = \ dimU$

则φ为同构 $\Leftrightarrow \varphi$单射 $\Leftrightarrow \varphi$满射

$\varphi \in \mathscr{L}(V); \Rightarrow$ 则φ单(满) $\Leftrightarrow \varphi$在某一组(任一组)基下的表示矩阵可逆

特别地，若$\varphi \in \mathscr{L}(V)$ 那么φ是自同构 $\Leftrightarrow \varphi$满射 $\Leftrightarrow \varphi$单射

Proof 利用维数公式和单满射的维数判别法得到

Theorem 5.6.5

设V, U是$\mathbb{F}$上的有限维线性空间, φ是V到U的线性映射

(1)$\boldsymbol{\varphi}$是单映射 $\Leftrightarrow$ 存在U到V的线性映射$\boldsymbol{\psi}$, 使$\boldsymbol{\psi\varphi} = \boldsymbol{Id}_V$, 这里$Id_V$表示$V$上的恒等映射;

(2)$\boldsymbol{\varphi}$是满映射 $\Leftrightarrow$ 是存在U到V的线性映射$\boldsymbol{\eta}$, 使$\boldsymbol{\varphi\eta} = \boldsymbol{Id}_U$, 这里$Id_U$表示$U$上的恒等映射.

Proof (1)若$\boldsymbol{\psi}\varphi = \boldsymbol{Id}_V$, 则对任意的$\boldsymbol{v} \in \mathrm{Ker}\boldsymbol{\varphi}$, 那么$\psi\varphi(v) = \boldsymbol{v} = 0$, 即 $\mathrm{Ker}\boldsymbol{\varphi} = 0$, 从而$\varphi$是单映射.

反之, 若φ是单映射, 则 $\dim\mathrm{Im}\varphi = \dim V$, 且$\varphi$是$V$到 $\mathrm{Im}\varphi$上的线性同构.设U_0是 $\mathrm{Im}\varphi$在U中的补空间, 即$U = \mathrm{Im}\varphi \oplus U_0$.

定义ψ为U到V的线性映射, 它在 $\mathrm{Im}\varphi$上的限制为φ^{-1}, 它在U_0上的限制是零线性映射, 则容易验证$\psi\varphi = Id_U$成立.

(2)若$\boldsymbol{\varphi\eta} = \boldsymbol{Id}_U$, 则对任意的$\boldsymbol{u} \in U, \boldsymbol{u} = \boldsymbol{\varphi}(\boldsymbol{\eta}(\boldsymbol{u}))$, 从而$\boldsymbol{\varphi}$是满映射.

反之, 若$\boldsymbol{\varphi}$是满映射, 则可取U的一组基$\boldsymbol{f}_1, \boldsymbol{f}_2, \cdots, \boldsymbol{f}_m$以及$V$中的向量$\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_m$, 使得$\boldsymbol{\varphi}(\boldsymbol{v}_i) = \boldsymbol{f}_i (i = 1, 2, \cdots, m)$.

定义$\boldsymbol{\eta}$为U到V的线性映射, 它在基上的作用为$\boldsymbol{\eta}(\boldsymbol{f}_i) = \boldsymbol{v}_i (i = 1, 2, \cdots, m)$, 则容易验证$\boldsymbol{\varphi\eta} = \boldsymbol{Id}_U$成立.

充分性同上述证法, 另一个方面用矩阵说法$\exists B_{n\times m}$使得$BA = I_n$那么$n = r(BA) \leqslant r(A) \leqslant n$

现只证必要性.取定V和U的两组基, 设$\boldsymbol{\varphi}$在这两组基下的表示矩阵为$m \times n$矩阵$\boldsymbol{A}$.

(1)若$\boldsymbol{\varphi}$是单映射, 则可知$\boldsymbol{A}$是列满秩矩阵.再由满秩分解可知, 存在$n \times m$矩阵$\boldsymbol{B}$, 使得$\boldsymbol{BA} = \boldsymbol{I}_n$.

由矩阵$\boldsymbol{B}$可定义从U到V的线性映射$\boldsymbol{\psi}$, 它适合$\psi\varphi = \mathrm{Id}_V$.

(2)若$\boldsymbol{\varphi}$是满映射, 则可知$\boldsymbol{A}$是行满秩矩阵.再由满秩分解可知, 存在$n \times m$矩阵$\boldsymbol{C}$, 使得$\boldsymbol{AC} = \boldsymbol{I}_m$.

由矩阵$\boldsymbol{C}$可定义从U到V的线性映射$\boldsymbol{\eta}$, 它适合$\varphi\boldsymbol{\eta} = \mathrm{Id}_U$. □

Lemma 5.4

设φ是线性空间V到U的线性映射

$\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\}$和$\{\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n\}$是$V$的两组基,$\{\boldsymbol{g}_1,\boldsymbol{g}_2,\cdots,\boldsymbol{g}_m\}$和$\{\boldsymbol{h}_1,\boldsymbol{h}_2,\cdots,\boldsymbol{h}_m\}$是$U$的两组基,

$\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\}$到$\{\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n\}$的过渡矩阵为$\boldsymbol{P}$. $\{\boldsymbol{g}_1,\boldsymbol{g}_2,\cdots,\boldsymbol{g}_m\}$到$\{\boldsymbol{h}_1,\boldsymbol{h}_2,\cdots,\boldsymbol{h}_m\}$的过渡矩阵为$\boldsymbol{Q}$.

φ在$\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\}$和$\{\boldsymbol{g}_1,\boldsymbol{g}_2,\cdots,\boldsymbol{g}_m\}$下的表示矩阵为$\boldsymbol{A}$,在$\{\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n\}$和$\{\boldsymbol{h}_1,\boldsymbol{h}_2,\cdots,\boldsymbol{h}_m\}$下的表示矩阵为$\boldsymbol{B}$.

求证:$\boldsymbol{B}=\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{P}$.

$$\begin{tikzcd}
V \arrow[rr, "\varphi"] & & U \\
\{e_1\cdots e_n\} \arrow[rr, leftrightarrow, "A"'] \arrow[d, "P"'] & & \{g_1\cdots g_m\} \arrow[d, "Q"] \\
\{f_1\cdots f_n\} \arrow[rr, leftrightarrow, "B"'] & & \{h_1\cdots h_m\}
\end{tikzcd}$$

Proof

设$\boldsymbol{v}\in V$,它在基$\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\}$下的坐标向量为$(x_1,x_2,\cdots,x_n)'$,则它在基$\{\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n\}$下的坐标向量为$\boldsymbol{P}^{-1}(x_1,x_2,\cdots,x_n)'$.

$\varphi(\boldsymbol{v})$在基$\{\boldsymbol{g}_1,\boldsymbol{g}_2,\cdots,\boldsymbol{g}_m\}$下的坐标向量为$\boldsymbol{A}(x_1,x_2,\cdots,x_n)'$,而在基$\{\boldsymbol{h}_1,\boldsymbol{h}_2,\cdots,\boldsymbol{h}_m\}$下的坐标向量为$\boldsymbol{B}\boldsymbol{P}^{-1}(x_1,x_2,\cdots,x_n)'$.

由于从$\{\boldsymbol{g}_1,\boldsymbol{g}_2,\cdots,\boldsymbol{g}_m\}$到$\{\boldsymbol{h}_1,\boldsymbol{h}_2,\cdots,\boldsymbol{h}_m\}$的过渡矩阵为$\boldsymbol{Q}$,故

$\boldsymbol{A}(x_1,x_2,\cdots,x_n)'=\boldsymbol{Q}\boldsymbol{B}\boldsymbol{P}^{-1}(x_1,x_2,\cdots,x_n)'$.因为$(x_1,x_2,\cdots,x_n)'$是任意的,故$A=QBP^{-1}$,即$B=Q^{-1}AP$. □

注 另一方面也反映出线性映射所反映出的矩阵的秩是一个定值

Corollary 5.3

设φ是有限维线性空间V到U的线性映射

求证:必存在V和U的两组基,使线性映射φ在两组基下的表示矩阵为$\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$.

Proof 设$\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\}$是$V$的一组基,$\{\boldsymbol{g}_1,\boldsymbol{g}_2,\cdots,\boldsymbol{g}_m\}$是$U$的一组基,$\varphi$在这两组基下的表示矩阵为$\boldsymbol{A}$.由相抵标准型理论可知存在$m$阶非异阵$\boldsymbol{Q}$,$n$阶非异阵$P$,使得$Q^{-1}AP=\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$.

设$\{f_1,f_2,\cdots,f_n\}$是V的一组新基,使得从$\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\}$到$\{\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n\}$的过渡矩阵为$\boldsymbol{P}$;

设$\{\boldsymbol{h}_1,\boldsymbol{h}_2,\cdots,\boldsymbol{h}_m\}$是$U$的一组新基,使得从$\{\boldsymbol{g}_1,\boldsymbol{g}_2,\cdots,\boldsymbol{g}_m\}$到$\{\boldsymbol{h}_1,\boldsymbol{h}_2,\cdots,\boldsymbol{h}_m\}$的过渡矩阵为$\boldsymbol{Q}$,

则由上个引理可知,φ在两组新基下的表示矩阵为$Q^{-1}\boldsymbol{A}\boldsymbol{P}=\begin{pmatrix} \boldsymbol{I}_r & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} \end{pmatrix}$.

Proposition 5.2 (维数公式证法二)

设φ是有限维线性空间$V\longrightarrow U$的线性映射

Proof.必存在V和U的两组基.使得φ在两组基下的表示矩阵为$\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$

Proof 设$\{e_1\cdots e_n\}$是V^n的一组基.设$\{g_1\cdots g_m\}$是U^m的一组基;φ在这两组基下的表示矩阵为$A^{m\times n}$

即有$\varphi(e_1\cdots e_n)=(g_1\cdots g_m)A$

由相抵标准型理论知:$\exists P^{m\times m};Q^{n\times n}\quad st.PAQ=\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$

接下来我们需要另外找两组基 $\{f_1\cdots f_n\}$ 与 $\{h_1\cdots h_m\}$ 使得φ在这两组基下表示矩阵为$\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$

我们不防假设已经找到了.设 $\{e_1\cdots e_n\}\to\{f_1\cdots f_n\}$ 的过渡矩阵为M ，$\{g_1\cdots g_m\}\to\{h_1\cdots h_m\}$ 的过渡矩阵为N

则$\varphi(f_1\cdots f_n)=(h_1\cdots h_m)\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$

$LHS=\varphi(f_1\cdots f_n)=(\varphi(e_1\cdots e_n))M=(g_1\cdots g_m)AM \qquad RHS=(g_1\cdots g_m)N\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$

那么$AM=N\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}\Rightarrow N^{-1}AM=PAQ=\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$

所以我们令$M=Q;N=P^{-1}$

反过来重新得到新基 $\{f_1\cdots f_n\}=\{e_1\cdots e_n\}M=\{e_1\cdots e_n\}Q \qquad \{h_1\cdots h_m\}=\{g_1\cdots g_m\}N=\{g_1\cdots g_m\}P^{-1}$即可 □

注 通过此法,很轻易的可以得到 $\mathrm{Ker}\varphi=L(f_{r+1}\cdots f_n)\quad \mathrm{Im}\varphi=L(h_1\cdots h_r)$ 立马也得到维数公式

Proposition 5.3 (维数公式证法三)

设$\varphi:V\longrightarrow U$为线性映射, 求证 : $\dim\mathrm{Ker}\varphi+\dim\mathrm{Im}\varphi=\dim V$.

Proof 设 $\dim V=n,\dim\mathrm{Ker}\varphi=k$, 我们只要证明 $\dim\mathrm{Im}\varphi=n-k$即可.

取 Kerφ的一组基$\boldsymbol{e}_1,\cdots,\boldsymbol{e}_k$, 并将其扩张为$V$的一组基$\boldsymbol{e}_1,\cdots,\boldsymbol{e}_k,\boldsymbol{e}_{k+1},\cdots,\boldsymbol{e}_n$.

任取$\boldsymbol{\alpha}\in V$, 设$\boldsymbol{\alpha}=c_1\boldsymbol{e}_1+\cdots+c_k\boldsymbol{e}_k+c_{k+1}\boldsymbol{e}_{k+1}+\cdots+c_n\boldsymbol{e}_n$, 则$\boldsymbol{\varphi}(\boldsymbol{\alpha})=c_{k+1}\boldsymbol{\varphi}(\boldsymbol{e}_{k+1})+\cdots+c_n\boldsymbol{\varphi}(\boldsymbol{e}_n)$

即 Imφ中任一向量都是$\boldsymbol{\varphi}(\boldsymbol{e}_{k+1}),\cdots,\boldsymbol{\varphi}(\boldsymbol{e}_n)$ 的线性组合.

下证$\boldsymbol{\varphi}(e_{k+1}),\cdots,\varphi(e_n)$ 线性无关.设$\lambda_{k+1}\varphi(e_{k+1})+\cdots+\lambda_n\varphi(e_n)=0$, 则$\boldsymbol{\varphi}(\lambda_{k+1}e_{k+1}+\cdots+\lambda_n e_n)=0$

即$\lambda_{k+1}e_{k+1}+\cdots+\lambda_n e_n\in\mathrm{Ker}\varphi$, 故可设$\lambda_{k+1}e_{k+1}+\cdots+\lambda_n e_n=\lambda_1 e_1+\cdots+\lambda_k e_k$

再由$e_1,\cdots,e_k,e_{k+1},\cdots,e_n$线性无关可知$\lambda_1=\cdots=\lambda_k=\lambda_{k+1}=\cdots=\lambda_n=0$.

因此$\varphi(e_{k+1}),\cdots,\varphi(e_n)$ 是 Imφ的一组基, 从而 $\dim\mathrm{Im}\varphi=n-k$, 结论得证. □

Proposition 5.4 (维数公式证法四)

下面我们从商空间的角度给出线性映射维数公式的第四种证法.

设V,U是数域$\mathbb{K}$上的有限维线性空间, $\varphi:V\longrightarrow U$是线性映射

证明 : 由φ诱导的线性映射$\bar{\varphi}:V/\mathrm{Ker}\varphi\longrightarrow\mathrm{Im}\varphi,\bar{\varphi}(v+\mathrm{Ker}\varphi)=\varphi(v)$是线性同构.特别地, $\dim V=\dim\mathrm{Ker}\varphi+\dim\mathrm{Im}\varphi$.

Proof 首先, $\bar{\varphi}$的定义不依赖于 Kerφ − 陪集代表元的选取.

事实上, 若$v_1+\mathrm{Ker}\varphi=v_2+\mathrm{Ker}\varphi$, 即$v_1-v_2\in\mathrm{Ker}\varphi$, 则$0=\varphi(v_1-v_2)=\varphi(v_1)-\varphi(v_2)$, 即$\varphi(v_1)=\varphi(v_2)$.

其次, 容易验证$\bar{\varphi}$是一个线性映射 (留给读者完成).

再次, 由$\bar{\varphi}$的定义不难看出它是满射.

最后, 若$\varphi(v_1)=\varphi(v_2)\in\mathrm{Im}\varphi$那么$v_1-v_2\in\mathrm{Ker}\varphi$那么很容易证明$v_1+\mathrm{Ker}\varphi=v_2+\mathrm{Ker}\varphi$

由商空间的维数公式可得 $\dim\mathrm{Im}\varphi=\dim(V/\mathrm{Ker}\varphi)=\dim V-\dim\mathrm{Ker}\varphi$, 由此即得线性映射的维数公式.

Theorem 5.6.6 (线性变换下-像核空间与全空间直和定理)

设V是数域$\mathbb{K}$上的n维线性空间,φ是V上的线性变换,证明以下9个结论等价:

(1)$V = \mathrm{Ker}\varphi \oplus \mathrm{Im}\varphi$

(2)$V = \mathrm{Ker}\varphi + \mathrm{Im}\varphi$;

(3)$\mathrm{Ker}\varphi \cap \mathrm{Im}\varphi = 0$;

(4)$\mathrm{Ker}\varphi = \mathrm{Ker}\boldsymbol{\varphi}^2$, 或等价地, $\dim\mathrm{Ker}\varphi = \dim\mathrm{Ker}\boldsymbol{\varphi}^2$;

(5)$\mathrm{Ker}\varphi = \mathrm{Ker}\varphi^2 = \mathrm{Ker}\varphi^3 = \cdots$, 或等价地, $\dim\mathrm{Ker}\varphi = \dim\mathrm{Ker}\varphi^2 = \dim\mathrm{Ker}\varphi^3 = \cdots$;

(6)$\mathrm{Im}\varphi = \mathrm{Im}\varphi^2$, 或等价地, $\mathrm{r}(\varphi) = \mathrm{r}\left(\varphi^2\right)$;

(7)$\mathrm{Im}\varphi = \mathrm{Im}\varphi^2 = \mathrm{Im}\varphi^3 = \cdots$, 或等价地, $\mathrm{r}(\varphi) = \mathrm{r}\left(\varphi^2\right) = \mathrm{r}\left(\varphi^3\right) = \cdots$

(8)$\mathrm{Ker}\varphi$存在$\boldsymbol{\varphi}-$不变补空间, 即存在$\boldsymbol{\varphi}-$不变子空间U, 使得$V = \mathrm{Ker}\varphi \oplus U$

(9)$\mathrm{Im}\varphi$存在$\boldsymbol{\varphi}-$不变补空间, 即存在$\boldsymbol{\varphi}-$不变子空间W, 使得$V = \mathrm{Im}\varphi \oplus W$.

Proof 由直和的定义可知 (1) $\Leftrightarrow$ (2) + (3), 于是 (1) $\Rightarrow$ (2)和 (1) $\Rightarrow$ (3)都是显然的.

根据交和空间维数公式和线性映射维数公式可知

$$\begin{aligned}\dim(\mathrm{Ker}\varphi + \mathrm{Im}\varphi) &= \dim\mathrm{Ker}\varphi + \dim\mathrm{Im}\varphi - \dim(\mathrm{Ker}\varphi \cap \mathrm{Im}\varphi)\\ &= \dim V - \dim(\mathrm{Ker}\varphi \cap \mathrm{Im}\varphi),\end{aligned}$$

于是 (2) $\Leftrightarrow$ (3)成立, 从而前3个结论两两等价.

(3) $\Rightarrow$ (4): 显然 $\mathrm{Ker}\varphi \subseteq \mathrm{Ker}\varphi^2$成立.

在取$\alpha \in \mathrm{Ker}\varphi^2$, 则$\varphi(\alpha) \in \mathrm{Ker}\varphi \cap \mathrm{Im}\varphi = 0$, 于是$\varphi(\alpha) = 0$, 即$\alpha \in \mathrm{Ker}\varphi$, 从而 $\mathrm{Ker}\varphi^2 \subseteq \mathrm{Ker}\varphi$也成立, 故 (4)成立.

(4) $\Rightarrow$ (3): 任取$\alpha \in \mathrm{Ker}\varphi \cap \mathrm{Im}\varphi$, 则存在$\beta \in V$, 使得$\alpha = \varphi(\beta)$, 于是$0 = \boldsymbol{\varphi}(\alpha) = \boldsymbol{\varphi}^2(\boldsymbol{\beta})$, 即$\beta \in \mathrm{Ker}\varphi^2 = \mathrm{Ker}\varphi$, 从而$\alpha = \boldsymbol{\varphi}(\boldsymbol{\beta}) = 0$, 即 (3)成立.

(5) $\Rightarrow$ (4): 显然

(4) $\Rightarrow$ (5): 已知 $\mathrm{Ker}\varphi^2 = \mathrm{Ker}\varphi$下证 $\mathrm{Ker}\varphi^3 = \mathrm{Ker}\varphi^2$

显然我们已经有了 $\mathrm{Ker}\varphi^2 \subset \mathrm{Ker}\varphi^3$那么另一边, 任取$\alpha \in \mathrm{Ker}\varphi^3$那么想证$\alpha \in \mathrm{Ker}\varphi^2$

有$\varphi^3(\alpha) = 0$那么$\varphi(\alpha) \in \mathrm{Ker}\varphi^2 = \mathrm{Ker}\varphi \Rightarrow \varphi(\varphi(\alpha)) = 0 = \varphi^2(\alpha) \Rightarrow \alpha \in \mathrm{Ker}\varphi^2$

类似的继续可以往下做

即: 假设已经证明了 $\mathrm{Ker}\varphi = \mathrm{Ker}\varphi^2 = \cdots \mathrm{Ker}\varphi^k$下证 $\mathrm{Ker}\varphi^k = \mathrm{Ker}\varphi^{k+1}$

$\mathrm{Ker}\varphi^k \subset \mathrm{Ker}\varphi^{k+1}$显然, 任取$\alpha \in \mathrm{Ker}\varphi^{k+1}$那么$\varphi^{k-1}(\alpha) \in \mathrm{Ker}\varphi^2 = \mathrm{Ker}\varphi \Rightarrow \varphi^k(\alpha) = 0 \Rightarrow \alpha \in \mathrm{Ker}\varphi^k \Rightarrow \mathrm{Ker}\varphi^{k+1} \subset \mathrm{Ker}\varphi^k$

至此一到五全部等价

显然 (4) $\Leftrightarrow$ (6)　(5) $\Leftrightarrow$ (7)

至此一到七全部等价

(1) $\Rightarrow$ (8)是显然的

下证 (8) $\Rightarrow$ (1).我们先证 $\mathrm{Im}\varphi \subseteq U$: 任取$\varphi(v) \in \mathrm{Im}\varphi$, 由直和分解可设$v = v_1 + u$, 其中$v_1 \in \mathrm{Ker}\varphi, u \in U$

则由U的$\varphi-$不变性可得$\varphi(v) = \varphi(v_1) + \varphi(u) = \varphi(u) \in U$.

考虑不等式 $\dim V = \dim(\mathrm{Ker}\varphi \oplus U) = \dim\mathrm{Ker}\varphi + \dim U \geq \dim\mathrm{Ker}\varphi + \dim\mathrm{Im}\varphi = \dim V$, 从而只能是$U = \mathrm{Im}\varphi$, 于是 (1)成立.

(1) $\Rightarrow$ (9)是显然的

下证 (9) $\Rightarrow$ (1).我们先证$W \subseteq \mathrm{Ker}\varphi$: 任取$\boldsymbol{w} \in W$, 则由$W$的$\boldsymbol{\varphi}-$不变性可得$\boldsymbol{\varphi}(\boldsymbol{w}) \in \mathrm{Im}\varphi \cap W = 0$, 即有$\boldsymbol{w} \in \mathrm{Ker}\boldsymbol{\varphi}$.

考虑不等式 $\dim V=\dim(\mathrm{Im}\varphi\oplus W)=\dim\mathrm{Im}\varphi+\dim W\le\dim\mathrm{Im}\varphi+\dim\mathrm{Ker}\varphi=\dim V$, 从而只能是$W=\mathrm{Ker}\boldsymbol{\varphi}$, 于是 (1)成立.

至此全部等价!

Example 5.1 $V=P[x]_n$;$\mathscr{A}f(x)=f'(x)$;$\mathscr{A}V=P[x]_{n-1}$;$Ker\mathscr{A}\ =\ P$ 易验证 $V\neq\mathrm{Im}\mathscr{A}\oplus Ker\mathscr{A}$

Proposition 5.5

设$\boldsymbol{A}$是n阶方阵, 求证 : $\mathrm{r}(\boldsymbol{A}^n)=\mathrm{r}\left(\boldsymbol{A}^{n+1}\right)=\mathrm{r}\left(\boldsymbol{A}^{n+2}\right)=\cdots\cdots$

Proof 证法1(代数方法)

由秩的不等式可得$n=\mathrm{r}(\boldsymbol{I}_n)\ge\mathrm{r}(\boldsymbol{A})\ge\mathrm{r}\left(\boldsymbol{A}^2\right)\ge\cdots\ge\mathrm{r}(\boldsymbol{A}^n)\ge\mathrm{r}\left(\boldsymbol{A}^{n+1}\right)\ge 0$.

上述$n+2$个整数都在 $[0,n]$之间, 故由抽屉原理可知, 存在某个整数$m\in[0,n]$, 使得 $\mathrm{r}(\boldsymbol{A}^m)=\mathrm{r}\left(\boldsymbol{A}^{m+1}\right)$.

对任意的$k\ge m$, 由矩阵秩的$Frobenius$不等式可得

$\mathrm{r}\left(\boldsymbol{A}^{k+1}\right)=\mathrm{r}\left(\boldsymbol{A}^{k-m}\boldsymbol{A}^m\boldsymbol{A}\right)\ge\mathrm{r}\left(\boldsymbol{A}^{k-m}\boldsymbol{A}^m\right)+\mathrm{r}(\boldsymbol{A}^m\boldsymbol{A})-\mathrm{r}(\boldsymbol{A}^m)=\mathrm{r}\left(\boldsymbol{A}^k\right)$, 又 $\mathrm{r}\left(\boldsymbol{A}^{k+1}\right)\le\mathrm{r}\left(\boldsymbol{A}^k\right)$, 故 $\mathrm{r}\left(\boldsymbol{A}^{k+1}\right)=\mathrm{r}\left(\boldsymbol{A}^k\right)$

对任意的$k\ge m$成立, 结论得证.

证法2(几何方法)

将$\boldsymbol{A}$看成是n维列向量空间上的线性变换, 记为$\boldsymbol{\varphi}$, 注意下列子空间链 : $V\supseteq\mathrm{Im}\varphi\supseteq\mathrm{Im}\varphi^2\supseteq\cdots\supseteq\mathrm{Im}\varphi^n\supseteq\mathrm{Im}\varphi^{n+1}$

上述$n+2$个子空间的维数都在 $[0,n]$之间, 故由抽屉原理可知, 存在某个整数$m\in[0,n]$, 使得 $\mathrm{Im}\varphi^m=\mathrm{Im}\varphi^{m+1}$.

现要证明对任意的$k\ge m$, $\mathrm{Im}\varphi^k=\mathrm{Im}\varphi^{k+1}$.一方面, $\mathrm{Im}\varphi^{k+1}\subseteq\mathrm{Im}\varphi^k$是显然的.

另一方面, 任取$\alpha\in\mathrm{Im}\varphi^k$, 则存在$\beta\in V$, 使得$\alpha=\varphi^k(\beta)$.

由于$\varphi^m(\beta)\in\mathrm{Im}\varphi^m=\mathrm{Im}\varphi^{m+1}$, 故存在$\gamma\in V$, 使得$\varphi^m(\beta)=\varphi^{m+1}(\gamma)$

从而$\alpha=\varphi^k(\beta)=\varphi^{k-m}\left(\varphi^m(\beta)\right)=\varphi^{k-m}\left(\varphi^{m+1}(\gamma)\right)=\varphi^{k+1}(\gamma)\in\mathrm{Im}\varphi^{k+1}$

故 $\mathrm{Im}\varphi^k=\mathrm{Im}\varphi^{k+1}$对任意的$k\ge m$成立, 取维数后即得结论.

Theorem 5.6.7 (线性变换幂次下像与核的性质)

设φ是n维线性空间V上的线性变换

求证 : 必存在整数$m\in[0,n]$, 使得 $\mathrm{Im}\varphi^m=\mathrm{Im}\varphi^{m+1}$, $\mathrm{Ker}\varphi^m=\mathrm{Ker}\varphi^{m+1}$, $\quad V=\mathrm{Im}\varphi^m\oplus\mathrm{Ker}\varphi^m$.

一方面从下面证明我们就可以知道从m往后都满足上式另一方面, 体现出 $\mathrm{Im}\varphi$不会一直变小, $Ker\varphi$不会一直变大

Proof 根据上个命题的证明可知, 存在整数$m\in[0,n]$, 使得 $\mathrm{Im}\varphi^m=\mathrm{Im}\varphi^{m+1}=\mathrm{Im}\varphi^{m+2}=\cdots$.

注意到对任意的正整数i, $\mathrm{Ker}\varphi^i\subseteq\mathrm{Ker}\varphi^{i+1}$.

再由维数公式可知, 对任意的$i\ge m$, $\dim\mathrm{Ker}\varphi^i=\dim V-\dim\mathrm{Im}\varphi^i=n-\dim\mathrm{Im}\varphi^m$是一个不依赖于$i$的常数

因此 $\mathrm{Ker}\varphi^m=\mathrm{Ker}\varphi^{m+1}=\mathrm{Ker}\varphi^{m+2}=\cdots$.

若$\alpha\in\mathrm{Im}\varphi^m\cap\mathrm{Ker}\varphi^m$, 则$\exists\beta$使得$\alpha=\varphi^m(\beta)$, $\varphi^m(\alpha)=0$.于是$0=\varphi^m(\alpha)=\varphi^{2m}(\beta)$, 即$\beta\in\mathrm{Ker}\varphi^{2m}=\mathrm{Ker}\varphi^m$, 从而$\alpha=\varphi^m(\beta)=0$

这证明了 $\mathrm{Im}\varphi^m\cap\mathrm{Ker}\varphi^m=0$.

又对V中任一向量α, 因为$\varphi^m(\alpha)\in\mathrm{Im}\varphi^m=\mathrm{Im}\varphi^{2m}$, 所以$\exists\beta$使得$\varphi^m(\alpha)=\varphi^{2m}(\beta)$, 其中$\beta\in V$.

我们有分解式$\alpha=\varphi^m(\beta)+(\alpha-\varphi^m(\beta))$.注意到$\varphi^m(\alpha-\varphi^m(\beta))=0$, 即$\alpha-\varphi^m(\beta)\in\mathrm{Ker}\varphi^m$, 这就证明了$V=\mathrm{Im}\varphi^m+\mathrm{Ker}\varphi^m$.

因此$V=\mathrm{Im}\varphi^m\oplus\mathrm{Ker}\varphi^m$.

注也可不证明$V=\mathrm{Im}\varphi^m+\mathrm{Ker}\varphi^m$, 改由维数公式 $\dim\mathrm{Im}\varphi^m+\dim\mathrm{Ker}\varphi^m=n$直接得到$V=\mathrm{Im}\varphi^m\oplus\mathrm{Ker}\varphi^m$

Problem 5.1 我们现在来考虑在 $\mathscr{L}(V^n)$ 上如何来判断自同构 (可逆线性变换) 的问题

解 我们如何判断一个映射是自同构无外乎两点.

第一点是否是线性映射, 这一点是比较简单地

第二点是否可逆, 也就是是否为双射那么就无外乎是否能直接找到逆映射, 是否能判断为单射, 是否能判断为满射. 因为在有限维空间上这两点只需证明一点

其中后面两种已经证毕前面一种略作说明在此

$(1):\varphi\in\mathscr{L}(V^n)$, 若存在$\psi\in\mathscr{L}(V^n)\quad st.\varphi\psi=I_V\Rightarrow\psi\varphi=I_V$亦即$\varphi^{-1}=\psi$

$(2):\varphi\in\mathscr{L}(V^n)$, 证明$\varphi$为单射 $\Leftrightarrow$ 证 $\mathrm{Ker}\varphi=\{0\}\Leftrightarrow$ 证 $\dim(\mathrm{Ker}\varphi)=0$

$(3):\varphi\in\mathscr{L}(V^n)$, 证明$\varphi$为满射 $\Leftrightarrow$ 证 $\mathrm{Im}\varphi=V\Leftrightarrow$ 证 $\dim(\mathrm{Im}\varphi)=n$

$$(1):\text{若}\varphi\psi=I_V\text{由于}\begin{cases}I_V\text{单射}\Rightarrow\psi\text{单射}\Rightarrow\psi\text{双射}\\ I_V\text{满射}\Rightarrow\varphi\text{满射}\Rightarrow\varphi\text{双射}\end{cases}\Rightarrow\varphi\psi\text{都可逆}$$

那么$\varphi=\psi^{-1}\Rightarrow\varphi\psi=I_V$反映出了二者互为逆映射

笔记

下面来叙述如何计算像空间与核空间的方法

$\varphi:V^n\rightarrow U^m$; $\{e_1\sim e_n\}\{f_1\sim f_m\}$ 为V与U的基

$Ker\varphi$如何用e_i表示; $\mathrm{Im}\varphi$如何用f_i表示

$$\eta_V:Ker\varphi\rightarrow Ker\varphi_A\text{线性同构}\quad \eta_V^{-1}:\begin{array}{c}Ker\varphi_A\rightarrow Ker\varphi\\ \begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}\longmapsto x_1e_1+\cdots+x_ne_n\end{array}$$

$$\eta_U:\mathrm{Im}\varphi\rightarrow\mathrm{Im}\varphi_A\text{线性同构}\quad \eta_U^{-1}:\begin{array}{c}\mathrm{Im}\varphi_A\rightarrow\mathrm{Im}\varphi\\ \begin{pmatrix}f_1\\f_2\\\vdots\\f_m\end{pmatrix}\longmapsto y_1f_1+\cdots+y_mf_m\end{array}$$

那么只要求出$Ker\varphi_A$和 $\mathrm{Im}\varphi_A$如何表示即可

$Ker\varphi_A$为$Ax=0$的解空间 $\quad\mathrm{Im}\varphi_A$即为$L(\alpha_1\cdots\alpha_n)$

$$\text{对}A\text{进行行变换}\Rightarrow\begin{cases}A\text{的列向量组的极大无关组}\\ Ax=0\text{的基础解系}\end{cases}\Rightarrow\begin{cases}\mathrm{Im}\varphi_A\text{的一组基通过线性同构得到 }\mathrm{Im}\varphi\text{的一组基}\\ Ker\varphi_A\text{的一组基通过线性同构得到}Ker\varphi\text{的一组基}\end{cases}$$

Example 5.2 设V是K上五维空间, U是K上四维空间, $\{\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3,\boldsymbol{e}_1,\boldsymbol{e}_5\}$ 是V的基, $\{f_1,f_2,f_3,f_4\}$ 是U的基

$$V\rightarrow U\text{的线性变换}\varphi\text{在上述基下的表示矩阵为}A=\begin{pmatrix}1&2&1&-3&2\\2&1&1&1&-3\\1&1&2&2&-2\\2&3&-5&-17&10\end{pmatrix}\text{求 }\mathrm{Im}\varphi\text{和 }\mathrm{Ker}\varphi.$$

笔记 解对矩阵A进行行初等变换: $A=\begin{pmatrix}1&2&1&-3&2\\2&1&1&1&-3\\1&1&2&2&-2\\2&3&-5&-17&10\end{pmatrix}\rightarrow\begin{pmatrix}1&0&0&1&-\frac{9}{4}\\0&1&0&-3&\frac{11}{4}\\0&0&1&2&-\frac{5}{4}\\0&0&0&0&0\end{pmatrix}$

因此$r(A)=3$, 即 $\dim\mathrm{Im}\varphi=3$.从上面可以看出A的前3个列向量线性无关, 因此它们可以组成 $\mathrm{Im}\varphi$的一组基, 故

$$\begin{aligned}\mathrm{Im}\varphi = &k_1\left(f_1+2f_2+f_3+2f_4\right)\\ &+k_2\left(2f_1+f_2+f_3+3f_4\right) \quad k_i\text{可取K中的任意数.}\\ &+k_3\left(f_1+f_2+2f_3-5f_4\right).\end{aligned}$$

方程$Ax=0$的基础解系为$\alpha_1=\begin{pmatrix}-1\\3\\-2\\1\\0\end{pmatrix}, \alpha_2=\begin{pmatrix}9\\-11\\5\\0\\4\end{pmatrix}$.

因此 $\mathrm{Ker}\varphi = k_1\left(-e_1+3e_2-2e_3+e_4\right)+k_2\left(9e_1-11e_2+5e_3+4e_5\right)$. k_i可取K中的任意数.

5.6.2 无限维线性空间的线性映射

笔记 对于无限维线性空间之间的线性映射, 我们并没有定义表示矩阵这一概念, 也没有维数公式等结论, 因此研究线性映射或线性变换, 无限维线性空间的情形远比有限维线性空间的情形难得多, 也常出现对有限维线性空间成立的结论在无限维线性空间却不成立的情况. 例如, 要证明无限维线性空间上的线性变换是自同构, 只能按照定义证明它既是单映射又是满映射, 而不能像有限维线性空间上的线性变换那样, 只验证它是单映射或满映射即可.

Problem 5.2 设V是实系数多项式全体构成的实线性空间, 定义V上的变换D, S如下 : $\boldsymbol{D}(f(x))=\dfrac{\mathrm{d}}{\mathrm{d}x}f(x), \boldsymbol{S}(f(x))=\displaystyle\int_0^x f(t)\mathrm{d}t$. $\boldsymbol{D},\boldsymbol{S}$均为$V$上的线性变换且$\boldsymbol{DS}=\boldsymbol{I}_V$, 但$\boldsymbol{SD}\neq\boldsymbol{I}_V$.

Proof

简单验证即得结论. 由$DS=I_V$可知, S是单线性映射, D是满线性映射. 容易看出$\boldsymbol{S}$不是满映射, $\boldsymbol{D}$不是单映射, 从而它们都不是自同构.

Theorem 5.6.8

设V是$\mathbb{K}$上的无限维线性空间, φ,ψ是V上的线性变换.

(1)φ和ψ都是可逆变换的 $\Leftrightarrow$ $\varphi\psi$和$\psi\varphi$都是可逆变换;

(2)若$\psi\boldsymbol{\varphi}=\boldsymbol{I}_V$, 则称$\psi$是$\boldsymbol{\varphi}$的左逆变换, $\boldsymbol{\varphi}$是ψ的右逆变换. 证明 : φ是可逆变换的 $\Leftrightarrow$ φ有且仅有一个左逆变换 (右逆变换).

Proof (1)若$\boldsymbol{\varphi}$和$\boldsymbol{\psi}$都是可逆变换, 则

$\left(\boldsymbol{\psi}^{-1}\boldsymbol{\varphi}^{-1}\right)(\boldsymbol{\varphi\psi})=(\boldsymbol{\varphi\psi})\left(\boldsymbol{\psi}^{-1}\boldsymbol{\varphi}^{-1}\right)=\boldsymbol{I}_V, \left(\boldsymbol{\varphi}^{-1}\boldsymbol{\psi}^{-1}\right)(\boldsymbol{\psi\varphi})=(\boldsymbol{\psi\varphi})\left(\boldsymbol{\varphi}^{-1}\boldsymbol{\psi}^{-1}\right)=\boldsymbol{I}_V$, 因此$\boldsymbol{\varphi\psi}$和$\boldsymbol{\psi\varphi}$都是可逆变换.

反之, 若$\varphi\psi$和$\psi\varphi$都是可逆变换, 则存在V上的线性变换$\boldsymbol{\xi},\boldsymbol{\eta}$, 使$\boldsymbol{\varphi\psi\xi}=\boldsymbol{\xi\varphi\psi}=\boldsymbol{I}_V, \psi\varphi\eta=\eta\psi\varphi=I_V$.

由$\varphi\psi\xi=I_V$可得φ是满映射, 由$\eta\psi\varphi=I_V$可得φ是单映射, 从而φ是可逆变换. 同理可证ψ也是可逆变换.

(2)若φ是可逆变换, 任取$\boldsymbol{\varphi}$的一个左逆变换$\boldsymbol{\psi}$, 则$\psi=\psi I_V=\psi\varphi\varphi^{-1}=I_V\varphi^{-1}=\varphi^{-1}$, 即$\varphi$的任一左逆变换都是逆变换$\varphi^{-1}$.

由逆变换的唯一性可知, φ有且仅有一个左逆变换.

反之, 若φ有且仅有一个左逆变换ψ, 则$\psi\varphi=I_V$, 且有 $\left(\psi+\varphi\psi-I_V\right)\varphi=\psi\varphi+\varphi\psi\varphi-\varphi=I_V+\varphi-\varphi=I_V$

即$\psi+\varphi\psi-I_V$也是φ的左逆变换, 从而$\psi+\varphi\psi-I_V=\psi$, 即$\varphi\psi=I_V$.

因此ψ也是φ的右逆变换, 从而φ是可逆变换. 同理可证关于右逆变换的结论. □

Problem 5.3 试构造无限维线性空间V以及V上的线性变换φ,ψ, 使得$\varphi\psi-\psi\varphi=I_V$.

Proof 设V是实系数多项式全体构成的实线性空间
线性变换φ,ψ定义为：对任意的$f(x)\in V,\boldsymbol{\varphi}(f(x))=f'(x),\boldsymbol{\psi}(f(x))=xf(x)$.
容易验证$\boldsymbol{\varphi\psi}-\boldsymbol{\psi\varphi}=\boldsymbol{I}_V$成立.注事实上,满足上述性质的线性变换$\varphi,\psi$决不可能存在于有限维线性空间$V$上.
若存在,取V的一组基并设φ,ψ的表示矩阵为$\boldsymbol{A},\boldsymbol{B}$,则有$\boldsymbol{AB}-\boldsymbol{BA}=\boldsymbol{I}$成立.
上式两边同时取迹,可得$0=\mathrm{tr}(\boldsymbol{AB}-\boldsymbol{BA})=\mathrm{tr}(\boldsymbol{I})=\mathrm{dim}V$,导出矛盾. □

5.7 不变子空间

Definition 5.13

设φ是线性空间V上的线性变换,U是V的子空间,若U适合条件:$\varphi(U)\subseteq U$

则称U是φ的不变子空间(或φ—子空间)

这时把φ限制在U上,则φ在U上定义了一个线性变换,称为由φ诱导出的线性变换,或称为φ在U上的限制,记为 $\varphi|_U$.

Proposition 5.6

不难验证零子空间，全子空间，值域，核空间都是不变子空间。把零子空间和全子空间成为平凡不变子空间,此外交空间与和空间也是不变子空间

Lemma 5.5

设$\varphi\in\mathscr{L}(V)$;$U=L(\alpha_1\cdots\alpha_m)$(其中$\alpha_i\in V$),

$\Rightarrow$ 则U为$\varphi-$不变子空间 $\Leftrightarrow\varphi(\alpha_i)\in U$

Proof 这是显然的我们在此略去证明

Theorem 5.7.1

设$\varphi\in\mathscr{L}(V^n)$; U为线性变换φ的不变子空间

且设U的基为$\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_r\}$.将$\{e_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_r\}$扩充为$V$的一组基$\{e_1,\boldsymbol{e}_2,\cdots,e_r,\boldsymbol{e}_{r+1},\cdots,e_n\}$

则φ在这组基下的矩阵具有下列形状:$\begin{pmatrix} a_{11} & \cdots & a_{r1} & a_{r+1,1} & \cdots & a_{n1}\\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\\ a_{1r} & \cdots & a_{rr} & a_{r+1,,} & \cdots & a_{nr}\\ 0 & \cdots & 0 & a_{r+1,r+1} & \cdots & a_{n,r+1}\\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\\ 0 & \cdots & 0 & a_{r+1,n} & \cdots & a_{nn}\end{pmatrix}$.(分块上三角矩阵)

Proof 上述定理较为简单我们略去证明

注 该定理的逆定理也成立

Corollary 5.4

设$V=V_1\oplus V_2$且V_1,V_2都是线性变换φ的不变子空间.又$\{\boldsymbol{e}_1,\cdots,\boldsymbol{e}_r\}$是$V_1$的基,$\{\boldsymbol{e}_{r+1},\cdots,\boldsymbol{e}_n\}$是$V_2$的基

则φ在基$\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\}$下的矩阵为分块对角阵$\begin{pmatrix} A_1 & 0\\ 0 & A_2\end{pmatrix}$,其中$A_1$为$r$阶方阵,$A_2$为$n-r$阶方阵

显然推论还可以进一步推广.若$V=V_1\oplus V_2\oplus\cdots\oplus V_m$

设每个V_i都是线性变换φ的不变子空间那么在V中存在一组基(这组基可由V_i的基合并而成)

使φ在这组基下的矩阵为分块对角阵:$\begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_m\end{pmatrix}$其中每块都是$r_i$阶方阵,$r_i=\dim V_i$

如果n维线性空间的线性变换φ有足够小的不变子空间

比如有n个一维不变子空间,其直和正好组成全空间,那么上式中的表示矩阵就是一个对角阵.

Example 5.3 设V是数域 K 上的三维空间,$\{\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3\}$是$V$的基,$\varphi$是$V$上线性变换,它在这组基下的表示矩阵为$\begin{pmatrix} 3 & 1 & -1\\ 2 & 2 & -1\\ 2 & 2 & 0\end{pmatrix}$

求证：由向量$\{\boldsymbol{e}_3,e_1+e_2+2\boldsymbol{e}_3\}$生成的子空间$U$是$\varphi$的不变子空间

Proof 证明$\varphi\left(e_3\right)$的坐标向量$\begin{pmatrix}3 & 1 & -1\\ 2 & 2 & -1\\ 2 & 2 & 0\end{pmatrix}\begin{pmatrix}0\\ 0\\ 1\end{pmatrix}=\begin{pmatrix}-1\\ -1\\ 0\end{pmatrix}$.容易求出向量组$\{(-1,-1,0)',(0,0,1)',(1,1,2)'\}$的秩为2

而U显然是二维子空间,因此$\varphi\left(e_3\right)\in U$.同理可证$\varphi\left(e_1+e_2+2e_3\right)\in U$.于是$U$是$\varphi-$不变子空间.证毕.

(核心思路是寻找$\varphi\left(e_3\right)$与$e_3,e_1+e_2+2e_3$的线性关系，对此我们利用线性同构来找线性关系。然后利用坐标向量公式即可)

更一般的化简到坐标向量的线性关系判断即是判断线性方程组的有无解的情形这就化为了系数矩阵和增广矩阵秩的相不相等判断

第 6 章 多项式

6.1 整除

Lemma 6.1

1.若$f(x),g(x)\in\mathbb{K}[x]$, 则 $\deg(f(x)\cdot g(x))=\deg f(x)+\deg g(x)$.

2.若$f(x),g(x)\in\mathbb{K}[x]$且$f(x)\neq 0,g(x)\neq 0$, 则$f(x)g(x)\neq 0$.

3.若$f(x)\neq 0$且$f(x)g(x)=f(x)h(x)$, 则$g(x)=h(x)$.

4.设$f(x),g(x)\in\mathbb{K}[x]$, 则 $\deg(cf(x))=\deg f(x), 0\neq c\in\mathbb{K}$

5.设$f(x),g(x)\in\mathbb{K}[x], \deg(f(x)+g(x))\leq\max\{\deg f(x),\deg g(x)\}$.

Definition 6.1

设$f(x),g(x)\in\mathbb{K}[x]$, 若存在$h(x)\in\mathbb{K}[x]$, 使$f(x)=g(x)h(x)$, 则称$g(x)$是$f(x)$的因式, 或$g(x)$可以整除$f(x)$, 或$f(x)$可以被$g(x)$整除记为$g(x)\mid f(x)$.否则称$g(x)$不能整除$f(x)$或$f(x)$不能被$g(x)$整除.

Proposition 6.1 (整除性质)

设$f(x),g(x),h(x)\in\mathbb{K}[x], 0\neq c\in\mathbb{K}$, 则

(1)若$f(x)\mid g(x)$, 则$cf(x)\mid g(x)$, 因此非零常数多项式c是任一非零多项式的因子;

(2)$f(x)\mid f(x)$;

(3)若$f(x)|g(x), g(x)|h(x)$, 则$f(x)\mid h(x)$;

(4)若$f(x)|g(x), f(x)|h(x)$, 则对任意的多项式$u(x),v(x)$, 有$f(x)\mid g(x)u(x)+h(x)v(x)$;

(5)设$f(x)|g(x), g(x)|f(x)$且$f(x),g(x)$都是非零多项式, 则存在$\mathbb{K}$中非零元c, 使$f(x)=cg(x)$.

Theorem 6.1.1 (带余除法)

设$f(x),g(x)\in\mathbb{K}[x], g(x)\neq 0$, 则必存在唯一的$q(x),r(x)\in\mathbb{K}[x]$, 使得$f(x)=g(x)q(x)+r(x)$, 且 $\deg r(x)<\deg g(x)$.

Proof 若$f(x)=0$或 $\deg f(x)<\deg g(x)$, 只需令$q(x)=0, r(x)=f(x)$即可.
现设 $\deg f(x)\geq\deg g(x)$, 对$f(x)$的次数用数学归纳法.
若 $\deg f(x)=0$, 则 $\deg g(x)=0$.因此可设$f(x)=a, g(x)=b(a\neq 0, b\neq 0)$.这时令$q(x)=ab^{-1}, r(x)=0$即可.
作为归纳假设, 我们设结论对小于n次的多项式均成立.
设$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0, a_n\neq 0$,
$g(x)=b_mx^m+b_{m-1}x^{m-1}+\cdots+b_1x+b_0, b_m\neq 0$,
由于$n\geq m$, 可令$f_1(x)=f(x)-a_nb_m^{-1}x^{n-m}g(x)$, 则 $\deg f_1(x)<n$.
由归纳假设, 有
$f_1(x)=g(x)q_1(x)+r(x)$, 且 $\deg r(x)<\deg g(x)$, 于是
$f(x)-a_nb_m^{-1}x^{n-m}g(x)=g(x)q_1(x)+r(x)$.
因此$f(x)=g(x)\left(a_nb_m^{-1}x^{n-m}+q_1(x)\right)+r(x)$.
于是我们可以令$q(x)=a_nb_m^{-1}x^{n-m}+q_1(x)$, 即得所求式.

再证明唯一性.设另有$p(x),t(x)$, 使$f(x)=g(x)p(x)+t(x)$, 且 $\deg t(x)<\deg g(x)$,
则$g(x)(q(x)-p(x))=t(x)-r(x)$.
这是若上式左边若$q(x)-p(x)\neq 0$, 便有 $\deg g(x)(q(x)-p(x))\geq\deg g(x)>\deg(t(x)-r(x))$

引出矛盾.因此只可能$p(x) = q(x), t(x) = r(x)$.

Corollary 6.1

设$f(x), g(x) \in \mathbb{K}[x], g(x) \neq 0$,则$g(x) \mid f(x)$的充分必要条件是$g(x)$除$f(x)$后的余式为零.

6.2 最大公因式

Definition 6.2 (最大公因式与最小公倍式)

设$f(x),g(x)\in\mathbb{K}[x]$,若$d(x)$是$f(x)$与$g(x)$的公因式,且对$f(x)$与$g(x)$的任一公因式$h(x)$均有$h(x)\mid d(x)$
则称$d(x)$为$f(x)$与$g(x)$的最大公因式(或称$d(x)$为$f(x),g(x)$的$g.c.d.$),记为$d(x)=(f(x),g(x))$.
同理,若$m(x)$是$f(x)$与$g(x)$的公倍式,且对$f(x)$与$g(x)$的任一公倍式$l(x)$均有$m(x)\mid l(x)$
则称$m(x)$为$f(x)$与$g(x)$的最小公倍式(或称$m(x)$为$f(x),g(x)$的$l.c.m.$),记为$m(x)=[f(x),g(x)]$.

Proposition 6.2

$(f,g)=f\iff f|g$

Theorem 6.2.1 (多项式中若干概念不随数域扩大)

1.带余除法的结果不随数域增大而改变
2.最大公因式,互素性,整除性,无重因式不随数域增大而改变

Proof 1.注意到带余除法的唯一性(反证法利用次数关系即可)
2.最大公因式可以看成带余除法的产品,互素性可以看成最大公因式的产品,整除性可以看成带余除法的产品
3.无重因式$\iff(f,f')=1$故利用互素即可
也可以另外证最大公因式方法如下
设$K_1\subset K_2$,设f与g在K_1中最大公因式为h_1,f与g在K_2中最大公因式为h_2
那么$h_1|h_2$自然有
存在$u_1,v_1\in K_1$使得$u_1f+v_1g=h_1$自然$u_1,v_1\in K_2\Longrightarrow h_2|h_1$
$\Longrightarrow h_1=h_2$
至此说明了最大公因式不随数域变大而变大,那么互素性是有最大公因式是否为1进而也不改变

Theorem 6.2.2 (欧几里得辗转相除法)

设$f(x),g(x)\in\mathbb{K}[x]$,则存在$f(x)$与$g(x)$的最大公因式$d(x)$,且有$u(x),v(x)\in\mathbb{K}[x]$,使得$f(x)u(x)+g(x)v(x)=d(x)$.

Proof 若$f(x)=0$,则显然$(f(x),g(x))=g(x)$;若$g(x)=0$,则$(f(x),g(x))=f(x)$.
故不妨设$f(x)\neq0,g(x)\neq0$.由带余除法,我们有下列等式:

$$\begin{aligned}
f(x)&=g(x)q_1(x)+r_1(x),\\
g(x)&=r_1(x)q_2(x)+r_2(x),\\
r_1(x)&=r_2(x)q_3(x)+r_3(x),\\
&\cdots\cdots\cdots\cdots\\
r_{s-2}(x)&=r_{s-1}(x)q_s(x)+r_s(x),\\
&\cdots\cdots\cdots\cdots
\end{aligned}$$

余式的次数是严格递减的,因此经过有限步后,必有一个等式其余式为零.
不妨设$r_{s+1}(x)=0$,于是$r_{s-1}(x)=r_s(x)q_{s+1}(x)$.
现在要证明$r_s(x)$即为$f(x)$与$g(x)$的最大公因式.
由上式知$r_s(x)\mid r_{s-1}(x)$,但$r_{s-2}(x)=r_{s-1}(x)q_s(x)+r_s(x)$,因此$r_s(x)\mid r_{s-2}(x)$.
这样可一直推下去,得$r_s(x)\,|g(x),r_s(x)|\,f(x)$.这表明$r_s(x)$是$f(x)$与$g(x)$的公因式.

又设$h(x)$是$f(x)$与$g(x)$的公因式,则$h(x)\mid r_1(x)$,于是$h(x)\mid r_2(x)$,不断往下推

容易看出有$h(x) \mid r_s(x)$.因此$r_s(x)$是最大公因式.
再证明$f(x)u(x)+g(x)v(x)=d(x)$.式.
从$r_{s-2}(x)=r_{s-1}(x)q_s(x)+r_s(x)$式得$r_s(x)=r_{s-2}(x)-r_{s-1}(x)q_s(x)$
但我们有$r_{s-3}(x)=r_{s-2}(x)q_{s-1}(x)+r_{s-1}(x)$
从中解出$r_{s-1}(x)$代入式,得$r_s(x)=r_{s-2}(x)\,(1+q_{s-1}(x)q_s(x))-r_{s-3}(x)q_s(x)$.
用类似的方法逐步将$r_i(x)$用$r_{i-1}(x),r_{i-2}(x)$代入,从而得到$r_s(x)=f(x)u(x)+g(x)v(x)$.
显然$u(x),v(x)\in\mathbb{K}[x]$.

注 若$d_1(x)$与$d_2(x)$都是$f(x),g(x)$的最大公因式,则$d_1(x)\mid d_2(x), d_2(x)\mid d_1(x)$
因此必存在$c\in\mathbb{K}$,使$d_2(x)=cd_1(x)$,即$f(x)$与$g(x)$的两个最大公因式最多相差一个非零常数.
若规定$f(x),g(x)$的最大公因式首项系数为1(首项系数等于1的多项式称为首一多项式),则$f(x)$与$g(x)$的最大公因式就唯一确定了.
今后我们规定,凡最大公因式均指首一多项式.

注 $f(x),g(x)$的最大公因式可以表示为$d(x)=u(x)f(x)+v(x)g(x)$,那其中的$u(x),v(x)$唯一吗?
答案是否定的!例如对任意的多项式$k(x)$有
$d(x)=u(x)f(x)+[k(x)f(x)g(x)-k(x)f(x)g(x)]+v(x)g(x)=(u(x)+k(x)g(x))f(x)+(v(x)-k(x)f(x))g(x)$.
所以$u_1(x)=u(x)+k(x)g(x), v_1(x)=v(x)-k(x)f(x)$也满足$d(x)=u_1(x)f(x)+v_1(x)g(x)$
所以$u(x),v(x)$取法是有无穷多种的.

Proposition 6.3 (多个多项式的最大公因式)

若$f(x),g(x),h(x)\in\mathbb{K}[x]$,则 $((f(x),g(x)),h(x))=(f(x),(g(x),h(x)))=(f(x),g(x),h(x))$.

Proof 设$d_1(x)=(f(x),g(x)), d_2(x)=(g(x),h(x)), d(x)=(f(x),g(x),h(x))$.则$d(x)\mid d_1(x), d(x)\mid h(x)$.
另一方面,若$u(x)\mid d_1(x), u(x)\mid h(x)$,则$u(x)\mid f(x), u(x)\mid g(x)$,从而$u(x)\mid d(x)$.
这说明$d(x)=(d_1(x),h(x))=((f(x),g(x)),h(x))$.同理,$d(x)=(f(x),(g(x),h(x)))$.

Definition 6.3 (互素定义)

设$f(x),g(x)\in\mathbb{K}[x]$,若 $(f(x),g(x))=1$,则称$f(x)$与$g(x)$互素.

Theorem 6.2.3 (互素充分必要条件)

设$f(x),g(x)\in\mathbb{K}[x]$,则$f(x)$与$g(x)$互素的充分必要条件是存在$u(x),v(x)\in\mathbb{K}[x]$,使$f(x)u(x)+g(x)v(x)=1$.

Proof 若 $(f(x),g(x))=1$,由定理知道,存在$u(x),v(x)$,使式成立.
反之,若式成立,假定 $(f(x),g(x))=d(x)$,由线性组合式可知$d(x)\mid 1$,于是$d(x)=1$.

Proposition 6.4 (互素的性质)

1.若$f_1(x)\mid g(x), f_2(x)\mid g(x)$,且 $(f_1(x),f_2(x))=1$,则$f_1(x)f_2(x)\mid g(x)$.
2.若 $(f(x),g(x))=1$,且$f(x)\mid g(x)h(x)$,则$f(x)\mid h(x)$.
3.设 $(f(x),g(x))=d(x), f(x)=f_1(x)d(x), g(x)=g_1(x)d(x)$,则 $(f_1(x),g_1(x))=1$.
4.设 $(f(x),g(x))=d(x)$,则 $(t(x)f(x),t(x)g(x))=t(x)d(x)$.
5.若 $(f_1(x),g(x))=1, (f_2(x),g(x))=1$,则 $(f_1(x)f_2(x),g(x))=1$.

Proof 1.不妨设$g(x)=f_1(x)s(x)=f_2(x)t(x)$

则$g(x)=f_1(x)s(x)\left(f_1(x)u(x)+f_2(x)v(x)\right)$, 其中$u(x),v(x)\in\mathbb{K}[x]$, 且使$f_1(x)u(x)+f_2(x)v(x)=1$.
于是

$$\begin{aligned}g(x)&=f_1(x)s(x)f_1(x)u(x)+f_1(x)f_2(x)s(x)v(x)\\&=f_2(x)t(x)f_1(x)u(x)+f_1(x)f_2(x)s(x)v(x)\\&=f_1(x)f_2(x)(t(x)u(x)+s(x)v(x)).\end{aligned}$$

即$f_1(x)f_2(x)\mid g(x)$.

2.设$u(x),v(x)\in\mathbb{K}[x]$, 且使$f(x)u(x)+g(x)v(x)=1$, 则$f(x)u(x)h(x)+g(x)v(x)h(x)=h(x)$.
因上式左边可被$f(x)$整除, 故$f(x)\mid h(x)$.

3.设$u(x),v(x)\in\mathbb{K}[x]$, 且使$f(x)u(x)+g(x)v(x)=d(x)$
即$f_1(x)d(x)u(x)+g_1(x)d(x)v(x)=d(x)$, 两边消去$d(x)$即得$f_1(x)u(x)+g_1(x)v(x)=1$, 因此$f_1(x),g_1(x)$互素.

4.设$u(x),v(x)\in\mathbb{K}[x]$, 且使$f(x)u(x)+g(x)v(x)=d(x)$, 则$t(x)f(x)u(x)+t(x)g(x)v(x)=t(x)d(x)$.
因此, 若$h(x)|t(x)f(x),h(x)|t(x)g(x)$, 则必有$h(x)\mid t(x)d(x)$.
又$t(x)d(x)$是$t(x)f(x),t(x)g(x)$的公因子, 因此$t(x)d(x)$是$t(x)f(x)$与$t(x)g(x)$的最大公因式.

5.设$f_1(x)u_1(x)+g(x)v_1(x)=1,\quad f_2(x)u_2(x)+g(x)v_2(x)=1$,
将上两式两边分别相乘得
$\left(f_1(x)f_2(x)\right)u_1(x)u_2(x)+g(x)\left[v_1f_2u_2+gv_1v_2+v_2f_1u_1\right]=1$.
这就是说$f_1(x)f_2(x)$和$g(x)$互素.

Corollary 6.2

1.设$f_1(x),\cdots,f_m(x),g_1(x),\cdots,g_n(x)$为多项式, 且$\left(f_i(x),g_j(x)\right)=1,1\le i\le m;1\le j\le n$
求证: $(f_1(x)f_2(x)\cdots f_m(x),g_1(x)g_2(x)\cdots g_n(x))=1$.
2.$(f(x),g(x))=1$的充要条件是对任意给定的正整数m,n, $(f(x)^m,g(x)^n)=1$.

Proof 1.利用$(f_1,g)=1$与$(f_2,g)=1\Longrightarrow(f_1f_2,g)=1$用数学归纳法即可

2.必要性由上个命题即得.
反过来, 若$d(x)\neq1$是$f(x)$和$g(x)$的公因式, 则它也是$f(x)^m$和$g(x)^n$的公因式, 因此$f(x)^m$和$g(x)^n$不可能互素.

Proposition 6.5

设$f(x),g(x)$是非零多项式, 则$f(x)g(x)\sim(f(x),g(x))[f(x),g(x)]$.

Proof 设$d(x)=(f(x),g(x))$且$f(x)=f_0(x)d(x),g(x)=g_0(x)d(x)$, 则$f_0(x),g_0(x)$互素.
下面我们就来证明: $\dfrac{fg}{(f,g)}=f_0g_0d$即为最小公倍式
一方面显然$f=f_0d\mid f_0g_0d=\dfrac{fg}{(f,g)}$同理$g|\dfrac{fg}{(f,g)}$
$\Longrightarrow\dfrac{fg}{(f,g)}$是一个公倍式, 下面来证最小性
假定$l(x)$是$f(x),g(x)$的公倍式且$l(x)=f(x)u(x)=g(x)v(x)$.故$f_0(x)d(x)u(x)=g_0(x)d(x)v(x)$.
我们只要证明$\dfrac{fg}{(f,g)}|l$即可
消去$d(x)$得$f_0(x)u(x)=g_0(x)v(x)$.因为$f_0(x),g_0(x)$互素, 所以$f_0(x)\,|v(x),g_0(x)|\,u(x)$

不妨设$u(x)=g_0(x)p(x)$.于是$l(x)=f_0(x)d(x)g_0(x)p(x)$.
即$f_0(x)d(x)g_0(x)\mid l(x)$这就是要证明的结论.

Theorem 6.2.4 (中国剩余定理)

设$\{f_i(x)\mid i=1,\cdots,n\}$是两两互素的多项式,$a_1(x),\cdots,a_n(x)$是$n$个多项式
则存在多项式$g(x),q_i(x)(i=1,\cdots,n)$,使得$g(x)=f_i(x)q_i(x)+a_i(x)$对一切$i$成立.

Proof 先证明存在多项式$g_i(x)$,使得对任意的i,有$g_i(x)=f_i(x)q_i(x)+1$且有$f_j(x)\mid g_i(x)(j\neq i)$.
一旦得证,只需令$g(x)=a_1(x)g_1(x)+\cdots+a_n(x)g_n(x)$即可.
现构造$g_1(x)$如下.因为$f_1(x)$和$f_j(x)(j\neq 1)$互素,故存在$u_j(x),v_j(x)$,使$f_1(x)u_j(x)+f_j(x)v_j(x)=1$.
令$g_1(x)=f_2(x)v_2(x)\cdots f_n(x)v_n(x)=(1-f_1(x)u_2(x))\cdots(1-f_1(x)u_n(x))$,
显然$g_1(x)$符合要求.同理可构造$g_i(x)$.

Corollary 6.3

设$d(x)$是$f_1(x),f_2(x),\cdots,f_m(x)$的最大公因式
求证：必存在多项式$g_1(x),g_2(x),\cdots,g_m(x)$,使得$f_1(x)g_1(x)+f_2(x)g_2(x)+\cdots+f_m(x)g_m(x)=d(x)$.

Proof 用数学归纳法.对$m=2$,结论已成立.
设结论对$m-1$成立.
即有：设$h(x)$是$f_1(x),f_2(x),\cdots,f_{m-1}(x)$的最大公因式,则有$g_1(x),g_2(x),\cdots,g_{m-1}(x)$
使得$f_1(x)g_1(x)+f_2(x)g_2(x)+\cdots+f_{m-1}(x)g_{m-1}(x)=h(x)$.
容易验证$d(x)$是$h(x)$和$f_m(x)$的最大公因式,故存在$u(x),v(x)$,使得$h(x)u(x)+f_m(x)v(x)=d(x)$.
将$h(x)$代入可得$f_1(x)g_1(x)u(x)+f_2(x)g_2(x)u(x)+\cdots+f_{m-1}(x)g_{m-1}(x)u(x)+f_m(x)v(x)=d(x)$,即知结论成立.

6.3 因式分解

Definition 6.4 (可约多项式与不可约多项式)

设$f(x)$是数域$\mathbb{K}$上的非常数多项式

若$f(x)$可以分解为两个次数小于$f(x)$次数的$\mathbb{K}$上多项式之积,则称$f(x)$是$\mathbb{K}$上的可约多项式.否则,称之为$\mathbb{K}$上的不可约多项式.

Proposition 6.6 (不可约多项式的性质)

1.设$f(x)$是数域$\mathbb{K}$上的不可约多项式

$\Longrightarrow$ 则对$\mathbb{K}$上任一多项式$g(x)$,或者$f(x) \mid g(x)$, 或者 $(f(x), g(x)) = 1$.

2.若$p(x)$是$\mathbb{K}$上的不可约多项式,$f(x), g(x)$是$\mathbb{K}$上的多项式且$p(x) \mid f(x)g(x)$

$\Longrightarrow$ 则或者$p(x) \mid f(x)$,或者$p(x) \mid g(x)$.

3.设$p(x)$为不可约多项式且$p(x) \mid f_1(x)f_2(x)\cdots f_m(x)$,则$p(x)$必可整除其中某个$f_i(x)$.

Proof 1设$d(x) = (f(x), g(x))$.

因为$f(x)$不可约,故$f(x)$的因式只能是非零常数多项式或$cf(x)(c \neq 0)$

从而或者$d(x) = 1$或者$d(x) = cf(x)$(首一多项式),故得结论.

2.若$p(x)$不能整除$f(x)$,则由1.知$p(x)$与$f(x)$互素,由互素性质即得$p(x) \mid g(x)$.

3.这是2.的额自然推论

Theorem 6.3.1

一次多项式无论在哪个数域下都是不可约的

Proof 通过定义即可

Theorem 6.3.2 (因式分解定理)

设$f(x)$是数域$\mathbb{K}$上的多项式且 $\deg f(x) \geq 1$, 则

(1)$f(x)$可分解为$\mathbb{K}$上的有限个不可约多项式之积

(2)若$f(x) = p_1(x)p_2(x)\cdots p_s(x) = q_1(x)q_2(x)\cdots q_t(x)$是$f(x)$的两个不可约分解

即$p_i(x), q_j(x)$都是$\mathbb{K}$上的次数大于零的不可约多项式

$\Longrightarrow$ 则$s = t$,且经过适当调换因式的次序以后,有$q_i(x) \sim p_i(x)(i = 1, 2, \cdots, s)$.

Proof (1)对多项式$f(x)$的次数用数学归纳法.

若 $\deg f(x) = 1$,结论显然成立.

设次数小于n的多项式都可以分解为$\mathbb{K}$上的不可约多项式之积

下面我们来看 $\deg f(x) = n$.

若$f(x)$不可约,结论自然成立.

若$f(x)$可约,则$f(x) = f_1(x)f_2(x)$,其中$f_1(x), f_2(x)$的次数小于n,由归纳假设它们可以分解为$\mathbb{K}$上的有限个不可约多项式之积.

所有这些多项式之积就是$f(x)$.

(2)对式中的s用数学归纳法.

若$s = 1$,则$f(x) = p_1(x)$,因此$f(x)$是不可约多项式,于是$t = 1, q_1(x) = p_1(x)$.

现假设对不可约因式个数小于s的多项式结论正确.

由式,有$p_1(x) \mid q_1(x)q_2(x)\cdots q_t(x)$,由不可约多项式性质可知,必存在某个$i$,不妨设$i=1$,使$p_1(x) \mid q_1(x)$.

但是$p_1(x), q_1(x)$都是不可约多项式,因此存在$c_1 \in \mathbb{K}$,使$q_1(x) = c_1 p_1(x)$.此即$p_1(x) \sim q_1(x)$.

将上式代入式并消去$p_1(x)$,得到$p_2(x)\cdots p_s(x) = c_1 q_2(x)\cdots q_t(x)$.

这时左边为$s-1$个不可约多项式之积,由归纳假设,$s-1=t-1$,即$s=t$.

另一方面,存在$c_i \in \mathbb{K}, c_i \neq 0, q_i(x) = c_i p_i(x)$.

Lemma 6.2

对于首一不可约多项式而言,互异 $\iff$ 互素

Proof 一方面若互素,则肯定互异,若相同,那么其最大公因式就为本身怎么会为1

另一方面,设f与g是首一不可约多项式,若不互素,那么由不可约多项式性质

$\implies f|g \implies g = fh$这迫使$h$只能为非零常数,又$f$与$g$是首一 $\implies h=1 \implies f=g$矛盾

故f与g互素

Theorem 6.3.3 (标准因式分解定理)

1.对于$f(x)$来说有如下的标准因式分解$f(x) = cp_1(x)^{e_1}p_2(x)^{e_2}\cdots p_m(x)^{e_m}$,其中$p_i(x)$是互异的首一不可约多项式,$e_i \geq 1 (i=1,2,\cdots,m)$.

2.设$f(x), g(x)$是$\mathbb{K}$上的两个多项式.在它们的标准分解式中适当添加零次项,我们不妨设它们有如下的标准分解式:

$f(x) = c_1 p_1(x)^{e_1} p_2(x)^{e_2}\cdots p_n(x)^{e_n}$

$g(x) = c_2 p_1(x)^{f_1} p_2(x)^{f_2}\cdots p_n(x)^{f_n}$,

其中$e_i \geq 0, f_i \geq 0 (i=1,2,\cdots,n)$

则$f(x)|g(x) \iff e_i \leqslant f_i$

则$f(x), g(x)$的最大公因式$(f(x), g(x)) = p_1(x)^{k_1}p_2(x)^{k_2}\cdots p_n(x)^{k_n}$,其中$k_i = \min\{e_i, f_i\}\ (i=1,2,\cdots,n)$.

则$f(x), g(x)$的最小公倍式$[f(x), g(x)] = p_1(x)^{h_1}p_2(x)^{h_2}\cdots p_n(x)^{h_n}$,其中$h_i = \max\{e_i, f_i\}\ (i=1,2,\cdots,n)$.

Proof 一方面若$e_i \leqslant f_i$显然$f(x)|g(x)$

另一方面若$f(x)|g(x)$,用反证法若$e_1 \geqslant f_1$

此时$p_1(x)^{e_1}p_2(x)^{e_2}\cdots p_n(x)^{e_n} | p_1(x)^{f_1}p_2(x)^{f_2}\cdots p_n(x)^{f_n} \implies p_1^{e_1-f_1}p_2^{e_2}\cdots p_n^{e_n} | p_2^{f_2}\cdots p_n^{f_n}$

$\implies p_1 | p_1^{e_1-f_1} | p_2^{f_2}\cdots p_n^{f_n}$由不可约多项式性质 $\implies p_1 | p_i\ (2 \leqslant i \leqslant n)$

那么$p_i = p_1 h(x)$迫使h只能为非零常数又首一使得$h(x)=1$使得$p_i = p_1$矛盾

Definition 6.5 (重因式)

在标准因式分解中,若$e_i > 1$,我们称式中的因式$p_i(x)$为$f(x)$的e_i重因式.

当$e_i = 1$时候,那么称为单因式

显然这时$p_i(x)^{e_i} \mid f(x)$,但$p_i(x)^{e_i+1}$不能整除$f(x)$.

Lemma 6.3

设$d(x) = (f(x), f'(x))$,则$\dfrac{f(x)}{d(x)}$是一个没有重因式的多项式,且这个多项式的不可约因式与$f(x)$的不可约因式相同(不计重数)

Proof 设$f(x)$有标准分解式$f = cp_1^{e_1}\cdots p_n^{e_n}$,则

$$
\begin{aligned}
f'(x) = & c e_1 p_1(x)^{e_1-1} p_2(x)^{e_2} \cdots p_s(x)^{e_s} p_1'(x) \\
& + c e_2 p_1(x)^{e_1} p_2(x)^{e_2-1} \cdots p_s(x)^{e_s} p_2'(x) \\
& + \cdots \\
& + c e_s p_1(x)^{e_1} p_2(x)^{e_2} \cdots p_s(x)^{e_s-1} p_s'(x).
\end{aligned}
$$

此时发现$p_1^{e_1-1}\cdots p_n^{e_n-1}$为$f$与$f'$的公因式

下面证其是最大公因式

设$d(x)=(f,f')=p_1^{k_1}\cdots p_m^{k_m}$

首先因为$d|f \Longrightarrow k_i \leqslant e_i$又我们还有$p_1^{e_1-1}\cdots p_n^{e_n-1}$为$f$与$f'$的公因式$\Longrightarrow e_i-1\leqslant k_i \Longrightarrow e_i-1\leqslant k_i\leqslant e_i$

断言k_i只能取e_i-1, 反证法若$k_1=e_1$

那么$d(x)=(f,f')=p_1^{e_1}\cdots p_m^{k_m} \Longrightarrow p_1^{e_1}|f' \Longrightarrow p_1|p_2(x)^{e_2}\cdots p_s(x)^{e_s}p_1'(x)$.

但是$p_1\sim p_n$是不同的故$\left(p_1,p_j\right)=1\,(2\leqslant j\leqslant n)\Longrightarrow p_1|p_1'$矛盾

$\Longrightarrow d(x)=(f,f')=p_1^{k_1}\cdots p_m^{k_m}=p_1^{e_1-1}\cdots p_n^{e_n-1}$

此时$\dfrac{f}{d}=cp_1\cdots p_m$

Theorem 6.3.4 (重因式导数判别法)

数域$\mathbb{K}$上的多项式$f(x)$没有重因式$\Longleftrightarrow (f,f')=1$

Proof 根据引理证明我们知道：$(f,f')=p_1^{e_1-1}\cdots p_m^{e_m-1}$

$f(x)$ 无重因式 $\Longleftrightarrow e_i=1 \Longleftrightarrow e_i-1=0 \Longleftrightarrow (f,f')=1$

6.4 多项式函数

Definition 6.6 (多项式的根)

设$f(x) \in \mathbb{K}[x], b \in \mathbb{K}$, 若$f(b) = 0$, 则称$b$是$f(x)$的一个根或零点.

Theorem 6.4.1 (余数定理)

设$f(x) \in \mathbb{K}[x], b \in \mathbb{K}$, 则存在$g(x) \in \mathbb{K}[x]$, 使$f(x) = (x-b)g(x) + f(b)$.
特别, b是$f(x)$的根当且仅当 $(x-b) \mid f(x)$.

Proof 由带余除法知$f(x) = (x-b)g(x) + r(x), \deg r(x) < 1$, 因此$r(x)$为常数多项式.在式中用$b$代替$x$, 即得$r(x) = f(b)$.

Definition 6.7 (重根定义)

设$f(x) \in \mathbb{K}[x], b \in \mathbb{K}$, 若存在正整数$k$, 使 $(x-b)^k \mid f(x)$, 但 $(x-b)^{k+1}$不能整除$f(x)$, 则称b是$f(x)$的一个k重根.
若$k = 1$, 则称b为单根

Lemma 6.4

设$f(x)$是数域$\mathbb{K}$上的不可约多项式且 $\deg f(x) \geq 2$, 则$f(x)$在$\mathbb{K}$中没有根.

Proof 用反证法, 设$b \in \mathbb{K}$是$f(x)$的根, 由定理知 $(x-b) \mid f(x)$, 即$f(x) = (x-b)g(x)$可分解为两个低次多项式之积, 这与$f(x)$不可约矛盾.

Theorem 6.4.2

若$f(x)$是数域$\mathbb{K}$上的n次多项式, 则$f(x)$在$\mathbb{K}$中最多只有n个根 (如果把k重根看成$f(x)$有k个根)

Proof 证明将$f(x)$作标准因式分解, 则由引理5.5.1知$f(x)$在$\mathbb{K}$中根的个数等于该分解式中一次因式的个数, 它不会超过n.

Corollary 6.4

设$f(x)$与$g(x)$是$\mathbb{K}$上的次数不超过n的两个多项式
若存在$\mathbb{K}$上$n+1$个不同的数$b_1, b_2, \cdots, b_{n+1}$, 使$f(b_i) = g(b_i), i = 1, 2, \cdots, n+1$, 则$f(x) = g(x)$.

Proof 证明作$h(x) = f(x) - g(x)$, 显然$h(x)$次数不超过n.但有$n+1$个不同的根, 因此只可能$h(x) = 0$, 即$f(x) = g(x)$.

Theorem 6.4.3

求证 : a是多项式$f(x)$的k重根的充要条件是 : $f(a) = f'(a) = \cdots = f^{(k-1)}(a) = 0, \quad f^{(k)}(a) \neq 0$.

Proof 若a是$f(x)$的k重根, 可设$f(x) = (x-a)^k g(x), g(x)$不含因式$x-a$.
通过对$f(x)$求导可发现, $x-a$可整除$f^{(j)}(x)(1 \leq j \leq k-1)$.
因此$f(a) = f'(a) = \cdots = f^{(k-1)}(a) = 0$.而$f^{(k)}(a) = k!g(a) \neq 0$, 故必要性得证.
反之, 若a是$f(x)$的m重根, 若$m > k$, 则由必要性的证明可知, 将有$f^{(k)}(a) = 0$, 这与已知矛盾.
同样, 若$m < k$, 则由$f(x) = (x-a)^m g(x)$ 其中$g(x)$ 不含有$x-a$也即a不是$g(x)$ 的根
将有$f^{(m)}(a) \neq 0$, 这也与已知矛盾, 于是只能$m = k$.

Proposition 6.7

$$
\begin{array}{ccc}
 & & \text{不可约} \\
 & & \Downarrow \\
\text{无重根} & \Leftarrow & \text{无重因式} \\
 & & \Updownarrow \\
 & & (f, f') = 1
\end{array}
$$

6.5 实数域与有理数域上不可约多项式

Lemma 6.5 (实系数多项式虚根总是共轭成对出现)

设$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$是实系数多项式, 若复数$a+b\mathrm{i}(b\neq 0)$是其根, 则$a-b\mathrm{i}$ 也是它的根.

Proof 令$z=a+b\mathrm{i}$, 其共轭复数为$\bar{z}=a-b\mathrm{i}$, 则
$f(\bar{z})=a_n\bar{z}^n+a_{n-1}\bar{z}^{n-1}+\cdots+a_1\bar{z}+a_0=\overline{a_nz^n+a_{n-1}z^{n-1}+\cdots+a_1z+a_0}=0.$
由此即得结论.

Theorem 6.5.1 (实数域上的不可约多项式结构与实数域上的因式分解定理)

1.实数域上的不可约多项式或为一次或为二次多项式 : ax^2+bx+c, 且$b^2-4ac<0$.

2.实数域上的多项式$f(x)$必可分解为有限个一次因式及不可约二次因式的乘积.

Proof 1.一次多项式显然为不可约.
当$b^2-4ac<0$时, ax^2+bx+c没有实根, 故不可约.
反过来, 任一高于二次的实系数多项式$f(x)$如有实根, 则$f(x)$可约; 如有一复根$a+b\mathrm{i}(b\neq 0)$, 则$a-b\mathrm{i}$ 也是它的根
故 $(x-(a+b\mathrm{i}))(x-(a-b\mathrm{i}))=x^2-2ax+\left(a^2+b^2\right)$ 是$f(x)$的因式, $f(x)$可约.

2.利用因式分解定理即可

Theorem 6.5.2 (整系数多项式有理根的必要条件)

设有n次整系数多项式$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$
$\Longrightarrow$ 则有理数$\frac{q}{p}$是$f(x)$的根的必要条件是$p\,|a_n, q|\,a_0$, 其中p, q是互素的整数.

Proof 将$\frac{q}{p}$代入式得$a_n\left(\frac{q}{p}\right)^n+a_{n-1}\left(\frac{q}{p}\right)^{n-1}+\cdots+a_1\frac{q}{p}+a_0=0$, 将上式两边乘以$p^n$得 :
$a_nq^n+a_{n-1}q^{n-1}p+\cdots+a_1qp^{n-1}+a_0p^n=0.$
因为$p|0$那么$p|\left(a_nq^n+a_{n-1}q^{n-1}p+\cdots+a_1qp^{n-1}+a_0q^n\right)\Longrightarrow p|a_nq^n\Longrightarrow p|a_n$
显然必须有$p\,|a_n, q|\,a_0$.

Corollary 6.5

1.首一多项式的有理系数根都为整数根

2.$f(x)$ 为一整系数多项式, 既约分数$\frac{p}{q}$为$f(x)$ 有理根 $\Longrightarrow p+q|f(-1)$ 且$p-q|f(1)$

注意2.取逆否命题就有若$p+q\nmid f(-1)$ 或者$p-q\nmid f(1)\Longrightarrow\frac{p}{q}$与$\frac{q}{p}$都不为$f(x)$ 的有理根

Proof 1.只要在上个定理中令$a_n=1$那么$\frac{p}{q}$中的$q|a_n=1\Longrightarrow q=\pm 1$故有理系数根一定为整数根
2.$f(x)$ 在$\mathbb{Q}$上可以写为$f(x)=\left(x-\frac{p}{q}\right)g(x)$
而$h(x)=qx-p=q\left(x-\frac{p}{q}\right)$是一不可约多项式且与$f(x)$ 在$\mathbb{Q}$上不互素 $\Longrightarrow h(x)\,|f(x)$
$\Longrightarrow f(x)=(qx-p)\,t(x)$
两边令$x=1$与-1得到答案

Definition 6.8 (本原多项式)

设多项式$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$是整系数多项式
若$a_n,a_{n-1},\cdots,a_1,a_0$的最大公约数等于1,则称$f(x)$为本原多项式.

Lemma 6.6 (Gauss 本原多项式引理)

两个本原多项式之积仍是本原多项式.

Proof 设两个本原多项式.
$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$,
$g(x)=b_mx^m+b_{m-1}x^{m-1}+\cdots+b_1x+b_0$
若$f(x)g(x)=c_{m+n}x^{m+n}+c_{m+n-1}x^{m+n-1}+\cdots+c_1x+c_0$不是本原多项式,则$c_0,c_1,\cdots,c_{m+n}$必有一个公共素因子$p$.
因为$f(x)$是本原多项式,p不能整除$f(x)$的所有系数
不妨设$p\mid a_0,p\mid a_1,\cdots,p\mid a_{i-1}$,但$p$不能整除$a_i$.
同理,不妨设$p\mid b_0,p\mid b_1,\cdots,p\mid b_{j-1}$,但$p$不能整除$b_j$.
注意到$c_{i+j}=\cdots+a_{i-2}b_{j+2}+a_{i-1}b_{j+1}+a_ib_j+a_{i+1}b_{j-1}+\cdots$
p可整除c_{i+j},p也能整除右式除a_ib_j以外的所有项.但p不能整除a_i和b_j,故p不能整除a_ib_j,引出矛盾.

Theorem 6.5.3

若整系数多项式$f(x)$在有理数域上可约,则它必可分解为两个次数较低的整系数多项式之积.
这个定理告诉我们,一个整系数多项式若在$\mathbb{Z}$上不可约 $\Longrightarrow$ $\mathbb{Q}$上不可约.
同样的有理系数多项式我们只需成倍数化为整系数多项式即可

Proof 假定整系数多项式$f(x)$可以分解为两个次数较低的有理系数多项式之积:
$f(x)=g(x)h(x)$
$g(x)$的各项系数为有理数,必有一个公分母记为c,于是$g(x)=\frac{1}{c}(cg(x))$,其中$cg(x)$为整系数多项式.
若把$cg(x)$中所有系数的最大公因数d提出来,则$g(x)=\frac{d}{c}\left(\frac{c}{d}g(x)\right)$,$\frac{c}{d}g(x)$是一个本原多项式.
这表明$g(x)=ag_1(x)$,a为有理数,$g_1(x)$为本原多项式.同样,$h(x)=bh_1(x)$,其中b为有理数,$h_1(x)$为本原多项式.
于是我们得到$f(x)=g(x)h(x)=abg_1(x)h_1(x)$.
由$Gauss$引理知,$g_1(x)h_1(x)$是本原多项式.
若ab不是一个整数,则$abg_1(x)h_1(x)$将不是整系数多项式,这与$f(x)$是整系数多项式相矛盾.
因此ab必须是整数,于是$f(x)$可以分解为两个次数较小的整系数多项式之积.

Corollary 6.6

已知$f(x)$是一个整系数多项式,并且本原多项式$g(x)$是$f(x)$的因式,若有$f(x)=g(x)h(x)$,$h(x)$是有理系数多项式
$\Longrightarrow h(x)$一定是整系数多项式

Proof $h(x)=ah_1(x)$此时a为有理数,$h_1(x)$为本原多项式
此时$f=agh_1$故此时$gh_1(x)$为本原多项式,故此时a为整数(可以考虑本原多项式定义)
故$f=g(x)\cdot(ah_1)=g(x)\cdot h(x)$为两个整系数多项式乘积,$h(x)$一定是整系数多项式

Corollary 6.7

1.对于实数域上的2次与3次多项式,在实系数域上可约 $\Longleftrightarrow$ 有实根
2.对于有理数域上的2次与3次多项式,在有理数域上可约 $\Longleftrightarrow$ 存在有理根

Proof 1.若有实根 $\Longrightarrow$ 可约非常容易,另一反面若可约则表示可以分解为两个次数较低的乘积
而原本的多项式仅仅只是两次三次故一定包含一个一次的多项式,则有实根
2.同理

Theorem 6.5.4

设$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$是一个整系数多项式,并且存在素数$p$使得$p\nmid a_n$; $p\mid a_{r-1},\cdots,a_1,a_0$; $p^2\nmid a_0$
则$f(x)$有一个次数大于等于r且在有理数域上不可约的因式.

Proof 设$f(x)=p_1(x)p_2(x)\cdots p_s(x)$,其中$p_1(x),p_2(x),\cdots,p_s(x)$为有理数域上不可约的整系数多项式
再设$p_i(x)$的常数项为$k_i(i=1,2,\cdots,s)$,则$k_1k_2\cdots k_s=a_0$
由$p\mid a_0$但$p^2\nmid a_0$可知$p\mid k_1k_2\cdots k_s$且$p^2\nmid k_1k_2\cdots k_s$
这意味着$k_1,k_2,\cdots,k_s$中有且仅有一个数可以被p整除,不妨设$p\mid k_1,p\nmid k_2,k_3,\cdots,k_s$
下证$\partial\left(p_1(x)\right)\geq r$
记$p_1(x)=b_mx^m+\cdots+b_1x+b_0$
记$g(x)=p_2(x)p_3(x)\cdots p_s(x)=c_lx^l+\cdots+c_1x+c_0\left(b_mc_l\neq 0\right)$.
其中$b_0=k_1,c_0=k_2k_3\cdots k_s$,即$p\mid b_0$且$p\nmid c_0$.
由于$f(x)=p_1(x)g(x)$,所以$b_mc_l=a_n$,由$p\nmid a_n$可知$p\nmid b_mc_l\Longrightarrow p\nmid b_m$且$p\nmid c_l$.
现在对$b_0,b_1,\cdots,b_m$从左往右依次检查,设第一个不被p整除的为b_t,则$m\geq t$
考虑a_t:根据$f(x)=p_1(x)g(x)$可知$a_t=b_tc_0+b_{t-1}c_1+b_{t-2}c_2+\cdots$
会发现$p\mid b_{t-1}c_1+b_{t-2}c_2+\cdots$,但$p\nmid b_tc_0$,于是$p\nmid a_t$.而已知说$p\mid a_0,a_1,\cdots,a_{r-1}$,这说明$t\geq r$,从而$m\geq t\geq r$,即$\partial\left(p_1(x)\right)\geq r$.
即$f(x)$有一个次数大于等于r的在有理数域上不可约的因式即这里的$p_1(x)$

Corollary 6.8 (艾森斯坦判别法)

设多项式$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$是整系数多项式,$a_n\neq 0,n\geq 1$
若p是一个素数.若$p\mid a_i(i=0,1,\cdots,n-1)$,但$p$不能整除$a_n$且$p^2$不能整除$a_0\Longrightarrow$则$f(x)$在有理数域上不可约.

Proof 我们只需证明$f(x)$在整数环上不可约即可.
设$f(x)$可分解为两个次数较低的整系数多项式之积:
$f(x)=\left(b_mx^m+b_{m-1}x^{m-1}+\cdots+b_0\right)\left(c_tx^t+c_{t-1}x^{t-1}+\cdots+c_0\right)$,其中$m+t=n$.
显然$a_0=b_0c_0,a_n=b_mc_t$.
由假设$p\mid a_0$,故$p\mid b_0$或$p\mid c_0$.
又p^2不能整除a_0,故p不能同时整除b_0及c_0.
不妨设$p\mid b_0$但p不能整除c_0.又由假设,p不能整除$a_n=b_mc_t$,故p既不能整除b_m又不能整除c_t.
因此不妨设$p\mid b_0,p\mid b_1,\cdots,p\mid b_{j-1}$但$p$不能整除$b_j$,其中$0<j\leq m<n$.
而$a_j=b_jc_0+b_{j-1}c_1+\cdots+b_0c_j$,根据假定,$p\mid a_j$,又$p$可整除上式右端除$b_jc_0$外的其余项,而不能整除$b_jc_0$这一项,引出矛盾.

Corollary 6.9

设整系数多项式$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_0$没有有理根,又有素数$p$满足$p\nmid a_n$; $p\mid a_{n-2},\cdots,a_0$; $p^2\nmid a_0$.
则$f(x)$在有理数域上不可约.

Proof 由定理知道存在一个次数大于等于$n-1$的$f(x)$的因式且该因式在有理数域上不可约
若该因式的此时为$n-1$那么就有$f(x)=g(x)h(x)$其中$g(x)$次数$n-1$那么$h(x)$次数为1
那么$f(x)$在有理数域上就存在有理根了,矛盾故$g(x)$的次数只能为n而不是$n-1$
$\Longrightarrow f(x)=g(x)$即$f(x)$在有理数域上不可约

Proposition 6.8

$f(x) \in P[x]$，那么$\forall a \in P$此时$f(x)$与$f(x+a)$在$P[x]$中的可约性是相同的

Example 6.1

$x^n - p$即为整数环(有理系数域)上的一个任意多次不可约多项式

Problem 6.1 若p为素数

证明：$f(x) = x^{p-1} + x^{p-2} + \cdots + x + 1$在有理数域上不可约.

Proof 作变量代换$x = y + 1$，得$f(x) = \dfrac{x^p - 1}{x - 1} = \dfrac{(y+1)^p - 1}{y} = y^{p-1} + C_p^1 y^{p-2} + C_p^2 y^{p-3} + \cdots + C_p^{p-1}$.

注意$p \mid C_p^i (1 \le i \le p-1)$，$p$不能整除首项系数1，$p^2$不能整除$C_p^{p-1} = p$，因此上述关于$y$的多项式在有理数域上不可约.

显然$f(x)$在有理数域上也不可约.

6.6 多元多项式

Lemma 6.7

若$f\left(x_1,x_2,\cdots,x_n\right)$及$g\left(x_1,x_2,\cdots,x_n\right)$都是$\mathbb{K}$上的$n$元多项式且非零

则按字典排列法排列后乘积的首项等于f的首项与g的首项之积.

Proof 设$ax_1^{i_1}x_2^{i_2}\cdots x_n^{i_n}$和$bx_1^{j_1}x_2^{j_2}\cdots x_n^{j_n}$分别是$f$和$g$的首项(按字典排列法),它们的乘积为$abx_1^{i_1+j_1}x_2^{i_2+j_2}\cdots x_n^{i_n+j_n}$.

其他任意两个单项式$cx_1^{k_1}x_2^{k_2}\cdots x_n^{k_n}$和$dx_1^{r_1}x_2^{r_2}\cdots x_n^{r_n}$之积为$cdx_1^{k_1+r_1}x_2^{k_2+r_2}\cdots x_n^{k_n+r_n}$

设$i_1=k_1,\cdots,i_{t-1}=k_{t-1},i_t>k_t;j_1=r_1,\cdots,j_{s-1}=r_{s-1},j_s>r_s$.

不妨设$t\le s$,显然$i_1+j_1=k_1+r_1,\cdots,i_{t-1}+j_{t-1}=k_{t-1}+r_{t-1},i_t+j_t>k_t+r_t$.

因此$abx_1^{i_1+j_1}x_2^{i_2+j_2}\cdots x_n^{i_n+j_n}$先于$cdx_1^{k_1+r_1}x_2^{k_2+r_2}\cdots x_n^{k_n+r_n}$.

同理可证明:$abx_1^{i_1+j_1}x_2^{i_2+j_2}\cdots x_n^{i_n+j_n}$先于$adx_1^{i_1+r_1}x_2^{i_2+r_2}\cdots x_n^{i_n+r_n}$和$cbx_1^{k_1+j_1}x_2^{k_2+j_2}\cdots x_n^{k_n+j_n}$.

因此它确是fg的首项.

Corollary 6.10

若$f\left(x_1,x_2,\cdots,x_n\right)\neq0,g\left(x_1,x_2,\cdots,x_n\right)\neq0$,则$f\left(x_1,x_2,\cdots,x_n\right)g\left(x_1,x_2,\cdots,x_n\right)\neq0$.

若$h\left(x_1,x_2,\cdots,x_n\right)\neq0$,且$f\left(x_1,x_2,\cdots,x_n\right)h\left(x_1,x_2,\cdots,x_n\right)=g\left(x_1,x_2,\cdots,x_n\right)h\left(x_1,x_2,\cdots,x_n\right)$,则$f\left(x_1,x_2,\cdots,x_n\right)=g\left(x_1,x_2,\cdots,x_n\right)$.

Lemma 6.8

1.两个次数相同的齐次多项式之和若不为零,则必仍是同次齐次多项式.

2.任意两个齐次多项式之积仍为齐次多项式.

3.任一n元多项式均可表示为若干个齐次多项式之和,这只需要将各次数相等的项放在一起即可.(且分量齐次多项式唯一)

Proposition 6.9

1.设$f\left(x_1,x_2,\cdots,x_n\right)$是$\mathbb{K}$上的$n$元非零多项式

$\Longrightarrow$ 则必存在$\mathbb{K}$中元$a_1,a_2,\cdots,a_n$,使$f\left(a_1,a_2,\cdots,a_n\right)\neq0$.

2.数域$\mathbb{K}$上的两个n元多项式$f\left(x_1,x_2,\cdots,x_n\right)$与$g\left(x_1,x_2,\cdots,x_n\right)$相等的 $\Longleftrightarrow$

对一切$a_1,a_2,\cdots,a_n\in\mathbb{K}$,均有$f\left(a_1,a_2,\cdots,a_n\right)=g\left(a_1,a_2,\cdots,a_n\right)$.

Proof 1.对未定元个数n用数学归纳法.

当$n=1$时,多项式$f(x)$只有有限个零点,故总有$a\in\mathbb{K}$使$f(a)\neq0$.

现设对有$n-1$个未定元的多项式结论成立.将$f\left(x_1,x_2,\cdots,x_n\right)$写成未定元$x_n$的多项式:

$f\left(x_1,x_2,\cdots,x_n\right)=b_0+b_1x_n+\cdots+b_mx_n^m$,其中$b_i=b_i\left(x_1,x_2,\cdots,x_{n-1}\right)$是$n-1$元多项式.

因为$f\left(x_1,x_2,\cdots,x_n\right)\neq0$,故可设$b_m\neq0$.由归纳假设,存在$a_1,\cdots,a_{n-1}\in\mathbb{K}$,使$b_m\left(a_1,\cdots,a_{n-1}\right)\neq0$.

因而

$f\left(a_1,\cdots,a_{n-1},x_n\right)=b_0\left(a_1,\cdots,a_{n-1}\right)+b_1\left(a_1,\cdots,a_{n-1}\right)x_n+\cdots+b_m\left(a_1,\cdots,a_{n-1}\right)x_n^m$是一个非零的以$x_n$为未定元的一元多项式

故存在$a_n\in\mathbb{K}$,使$f\left(a_1,a_2,\cdots,a_n\right)\neq0$.

2.只需证明充分性.作$h\left(x_1,x_2,\cdots,x_n\right)=f\left(x_1,x_2,\cdots,x_n\right)-g\left(x_1,x_2,\cdots,x_n\right)$.若$h\left(x_1,x_2,\cdots,x_n\right)\neq0$

则必有$a_1,a_2,\cdots,a_n\in\mathbb{K}$,使$h\left(a_1,a_2,\cdots,a_n\right)\neq0$.与假设矛盾.

Definition 6.9 (对称多项式)

设$f(x_1,x_2,\cdots,x_n)$是数域$\mathbb{K}$上的n元多项式

若对任意的$i\neq j(1\leq i,j\leq n)$,均有$f(x_1,\cdots,x_i,\cdots,x_j,\cdots,x_n)=f(x_1,\cdots,x_j,\cdots,x_i,\cdots,x_n)$

则称$f(x_1,x_2,\cdots,x_n)$是数域$\mathbb{K}$上的n元对称多项式.

Proposition 6.10

1.对称多项式之和仍是对称多项式

2.对称多项式之积也是对称多项式.

3.因此对称多项式的多项式还是对称多项式.

Definition 6.10

在对称多项式中,有一类基本的多项式,称为初等对称多项式.它们是这样定义的:

$\sigma_1=x_1+x_2+\cdots+x_n=\sum_{i=1}^n x_i,$

$\sigma_2=x_1x_2+x_1x_3+\cdots+x_{n-1}x_n=\sum_{1\leq i<j\leq n}x_ix_j,$

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

$\sigma_n=x_1x_2\cdots x_n.$

Theorem 6.6.1

设$f(x_1,x_2,\cdots,x_n)$是数域$\mathbb{K}$上的对称多项式

$\Longrightarrow$ 则必存在$\mathbb{K}$上唯一的一个多项式$g(y_1,y_2,\cdots,y_n)$,使$f(x_1,x_2,\cdots,x_n)=g(\sigma_1,\sigma_2,\cdots,\sigma_n)$.

Proof 存在性:设$f(x_1,x_2,\cdots,x_n)$按字典排列法的首项为$ax_1^{i_1}x_2^{i_2}\cdots x_n^{i_n},a\neq 0$.则必有$i_1\geq i_2\geq\cdots\geq i_n$.

事实上$f(x_1,x_2,\cdots,x_n)$是对称多项式,若$i_k<i_{k+1}$,则将x_k与x_{k+1}对换得到$ax_1^{i_1}\cdots x_k^{i_{k+1}}x_{k+1}^{i_k}\cdots x_n^{i_n}$

这一项也是$f(x_1,x_2,\cdots,x_n)$中的项,但它应在$ax_1^{i_1}x_2^{i_2}\cdots x_n^{i_n}$之前,此与首项的假定相矛盾.

作多项式$g_1(x_1,x_2,\cdots,x_n)=a\sigma_1^{i_1-i_2}\sigma_2^{i_2-i_3}\cdots\sigma_{n-1}^{i_{n-1}-i_n}\sigma_n^{i_n}$.

显然g_1是$x_1,x_2,\cdots,x_n$的对称多项式,且g_1的首项为$\underbrace{ax_1^{i_1-i_2}(x_1x_2)^{i_2-i_3}\cdots(x_1\cdots x_n)^{i_n}=ax_1^{i_1}x_2^{i_2}\cdots x_n^{i_n}}_{\text{利用了首项的乘积为乘积的首项}}$.

因此g_1与f的首项相同,于是$f_1=f-g_1$是一个首项后于f的对称多项式.

对f_1重复上述做法,不断做下去,便得到一列对称多项式:

$f_0=f,f_1=f_0-g_1,f_2=f_1-g_2,\cdots$,每个$f_i$的首项都后于$f_{i-1}$的首项.

设$bx_1^{k_1}x_2^{k_2}\cdots x_n^{k_n}$是某个$f_i(i\geq 1)$的首项,则因它后于$f$的首项,故有$i_1\geq k_1\geq k_2\geq\cdots\geq k_n\geq 0$.

这样的$k_1,k_2,\cdots,k_n$只有有限个,故多项式f_i不能无限地构造下去,即存在某个正整数s,使$f_s=0$.

于是$f=g_1+f_1=g_1+g_2+f_2=\cdots=g_1+g_2+\cdots+g_s$.由于每个$g_i$都可表示为$\sigma_1,\sigma_2,\cdots,\sigma_n$的多项式

故f也可表示为$\sigma_1,\sigma_2,\cdots,\sigma_n$的多项式.

唯一性:

设$g(y_1,\cdots,y_n)$及$h(y_1,\cdots,y_n)$都是$\mathbb{K}$上的n元多项式,且$f(x_1,\cdots,x_n)=g(\sigma_1,\cdots,\sigma_n)=h(\sigma_1,\cdots,\sigma_n)$.

令$\varphi(y_1,\cdots,y_n)=g(y_1,\cdots,y_n)-h(y_1,\cdots,y_n)$.

由假定,有$\varphi(\sigma_1,\cdots,\sigma_n)=0$.我们要证明多项式$\varphi(y_1,\cdots,y_n)=0$.

假定$\varphi(y_1,\cdots,y_n)\neq 0$,不妨设$\varphi(y_1,y_2,\cdots,y_n)=\cdots+cy_1^{k_1}y_2^{k_2}\cdots y_n^{k_n}+dy_1^{j_1}y_2^{j_2}\cdots y_n^{j_n}+\cdots$

其中$c,d,\cdots$均不为零且假定各单项式彼此不是同类项.

在$\varphi(\sigma_1,\sigma_2,\cdots,\sigma_n)$中

$c\sigma_1^{k_1}\sigma_2^{k_2}\cdots\sigma_n^{k_n}=cx_1^{k_1}(x_1x_2)^{k_2}\cdots(x_1x_2\cdots x_n)^{k_n}+\cdots\quad=cx_1^{k_1+k_2+\cdots+k_n}x_2^{k_2+k_3+\cdots+k_n}\cdots x_n^{k_n}+\cdots$

$d\sigma_1^{j_1}\sigma_2^{j_2}\cdots\sigma_n^{j_n}=dx_1^{j_1}(x_1x_2)^{j_2}\cdots(x_1x_2\cdots x_n)^{j_n}+\cdots\quad=dx_1^{j_1+j_2+\cdots+j_n}x_2^{j_2+j_3+\cdots+j_n}\cdots x_n^{j_n}+\cdots$

因此$c\sigma_1^{k_1}\sigma_2^{k_2}\cdots\sigma_n^{k_n}$与$d\sigma_1^{j_1}\sigma_2^{j_2}\cdots\sigma_n^{j_n}$等化成$x_1,x_2,\cdots,x_n$的多项式后其首项都不是同类项

从而$\varphi(\sigma_1,\sigma_2,\cdots,\sigma_n)$的首项一定是这些首项中排在最前的那一个.特别,$\varphi(\sigma_1,\sigma_2,\cdots,\sigma_n)\neq 0$,引出矛盾.

笔记 如何将一个对称多项式转化为初等对称多项式的多项式

1.我们先来看对称齐次多项式

将对称多项式$f(x_1,x_2,x_3)=\left(x_1^2+x_2^2\right)\left(x_1^2+x_3^2\right)\left(x_2^2+x_3^2\right)$表示为初等对称多项式的多项式.

注意到f是一个齐次多项式,次数等于6.又f的首项是$x_1^4x_2^2$,它的指数组为$(4,2,0)$.

从定理的证明中可看出我们在不断消去首项的过程中,齐次$(=6)$这个性质一直被保留

且还会有一个上界限制$i_1\geqslant k_1\geqslant\cdots\geqslant k_n\geqslant 0$

f_i首项的指数组(i,j,k)只可能是$(4,1,1),(3,3,0),(3,2,1),(2,2,2)$

相对应的f_i的首项为

$\sigma_1^{4-1}\sigma_2^{1-1}\sigma_3^1=\sigma_1^3\sigma_3$

$\sigma_1^{3-3}\sigma_2^{3-0}\sigma_3^0=\sigma_2^3$

$\sigma_1^{3-2}\sigma_2^{2-1}\sigma_3^1=\sigma_1\sigma_2\sigma_3$

$\sigma_1^{2-2}\sigma_2^{2-2}\sigma_3^2=\sigma_3^2$

因此可设$f=\sigma_1^2\sigma_2^2+a\sigma_1^3\sigma_3+b\sigma_2^3+c\sigma_1\sigma_2\sigma_3+d\sigma_3^2$.

取x_1,x_2,x_3的一些特殊值便得到关于a,b,c,d的线性方程组.

不难解得$a=-2,b=-2,c=4,d=-1$.因此$f=\sigma_1^2\sigma_2^2-2\sigma_1^3\sigma_3-2\sigma_2^3+4\sigma_1\sigma_2\sigma_3-\sigma_3^2$.

2.若一开始不是齐次的,我们则可以先进行齐次分解

Lemma 6.9

令$s_k=x_1^k+x_2^k+\cdots+x_n^k(k\geq 1)$; $s_0=n$.

设$f(x)=(x-x_1)(x-x_2)\cdots(x-x_n)=x^n-\sigma_1x^{n-1}+\sigma_2x^{n-2}+\cdots+(-1)^n\sigma_n$

则$x^{k+1}f'(x)=\left(s_0x^k+s_1x^{k-1}+\cdots+s_k\right)f(x)+g(x)$,其中$g(x)$作为$x$的多项式次数小于$n$.

Proof 容易看出$f'(x)=\sum\limits_{i=1}^n\dfrac{f(x)}{x-x_i}$.

因此

$$
\begin{aligned}
x^{k+1}f'(x)&=\sum_{i=1}^n\frac{x^{k+1}}{x-x_i}f(x)\\
&=\sum_{i=1}^n\frac{x^{k+1}-x_i^{k+1}+x_i^{k+1}}{x-x_i}f(x)\\
&=f(x)\sum_{i=1}^n\frac{x^{k+1}-x_i^{k+1}}{x-x_i}+g(x),
\end{aligned}
$$

其中$g(x)=\sum\limits_{i=1}^n\dfrac{x_i^{k+1}f(x)}{x-x_i}$作为$x$的多项式其次数小于$n$.

又

$$
\begin{aligned}
\sum_{i=1}^n\frac{x^{k+1}-x_i^{k+1}}{x-x_i}&=\sum_{i=1}^n\left(x^k+x_ix^{k-1}+\cdots+x_i^k\right)\\
&=nx^k+(x_1+\cdots+x_n)x^{k-1}+\cdots+\left(x_1^k+\cdots+x_n^k\right)\\
&=s_0x^k+s_1x^{k-1}+\cdots+s_k
\end{aligned}
$$

于是$x^{k+1}f'(x)=\left(s_0x^k+s_1x^{k-1}+\cdots+s_k\right)f(x)+g(x)$.

Theorem 6.6.2 (Newton 公式)

$s_k = x_1^k + \cdots + x_n^k$且$s_0 = n$

若$k \le n-1$, 则$s_k - s_{k-1}\sigma_1 + s_{k-2}\sigma_2 - \cdots + (-1)^{k-1}s_1\sigma_{k-1} + (-1)^k k\sigma_k = 0$

若$k \ge n$, 则$s_k - s_{k-1}\sigma_1 + s_{k-2}\sigma_2 - \cdots + (-1)^n s_{k-n}\sigma_n = 0.$

Proof 对$f(x) = x^n - \sigma_1 x^{n-1} + \sigma_2 x^{n-2} + \cdots + (-1)^n\sigma_n$求导并乘以$x^{k+1}$得到：

$x^{k+1}f'(x) = nx^{n+k} - (n-1)\sigma_1 x^{n+k-1} + (n-2)\sigma_2 x^{n+k-2} + \cdots + (-1)^{n-1}\sigma_{n-1}x^{k+1}.$

上式中x^i的项为$(-1)^{n+k-1}(i-k)\,\sigma_{n+k-i}x^i$

由引理, 有

$$\begin{aligned} x^{k+1}f'(x) &= \left(s_0x^k + s_1x^{k-1} + \cdots + s_k\right) f(x) + g(x) \\ &= \left(s_0x^k + s_1x^{k-1} + \cdots + s_k\right)\left(x^n - \sigma_1x^{n-1} + \sigma_2x^{n-2} + \cdots + (-1)^n\sigma_n\right) + g(x) \end{aligned}$$

比较x^n的系数即知

当$k \le n-1$时, 有$s_k - s_{k-1}\sigma_1 + s_{k-2}\sigma_2 - \cdots + (-1)^{k-1}s_1\sigma_{k-1} + (-1)^k k\sigma_k = 0$;

若$k \ge n$, 则$s_k - s_{k-1}\sigma_1 + s_{k-2}\sigma_2 - \cdots + (-1)^n s_{k-n}\sigma_n = 0$

Theorem 6.6.3 (Vieta 定理)

数域P上的n次多项式$f(x) = a_nx^n + \cdots + a_1x + a_0$的$n$个复数根为$x_1, x_2, \cdots, x_n$, 则

$$\begin{cases} \sigma_1 = x_1 + x_2 + \cdots + x_n = -\frac{a_{n-1}}{a_n}; \\ \sigma_2 = x_1x_2 + x_1x_3 + \cdots + x_1x_n + x_2x_3 + \cdots + x_{n-1}x_n = \dfrac{a_{n-2}}{a_n}; \\ \cdots\cdots \\ \sigma_i = \sum\limits_{j_1j_2\cdots j_i} x_{j_1}x_{j_2}\cdots x_{j_i} = (-1)^i\dfrac{a_{n-i}}{a_n} \; (\text{所有可能的}i\text{ 个不同的}x_i\text{ 的乘积之和}) \\ \cdots\cdots \\ \sigma_n = x_1x_2\cdots x_n = (-1)^n\dfrac{a_0}{a_n}. \end{cases}$$

第 7 章 相似标准型

7.1 特征值与特征向量

Definition 7.1 (特征值和特征向量)

设$\varphi \in \mathscr{L}(V^n)$；若存在$\lambda \in \mathbb{K}; 0 \neq e \in V^n \ st.\varphi(e) = \lambda e$

则称λ为φ的特征值。e为φ对应于λ特征值的特征向量。

令$V_\lambda = \{v \in V|\varphi(v) = \lambda v\} = \{\lambda\text{特征向量}\} \cup \{0\}$.易知$V_\lambda$为$V$子空间，更是$\varphi$不变子空间,称为特征值$\lambda$的特征子空间

代数版本：任取V的一组基，φ表示矩阵为A；e的坐标向量$\alpha \in K^n$; 则$\varphi(e)$坐标向量为$A\alpha$

有$\varphi(e) = \lambda e \to A\alpha = \lambda\alpha \quad V_\lambda : (\lambda I_n - A)x = 0$解空间 $|\lambda I_n - A|$ 为矩阵A的特征多项式

Lemma 7.1

相似矩阵具有相同的特征多项式

Proof 设$B = P^{-1}AP$ (其中P非异)

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

笔记 反之不对例如$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ $B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$

Definition 7.2 (线性变换的特征值和特征向量)

线性变换$\varphi \in \mathscr{L}(V^n)$.任取一组基，表示矩阵为$A$

则φ的特征多项式 $\triangleq |\lambda I_n - A|$ 即线性变换的特征多项式不依赖于基的选取

Corollary 7.1

根据韦达定理不难得到 $\lambda_1 + \lambda_2 + \cdots + \lambda_n = tr(A) \quad \lambda_1\lambda_2\cdots\lambda_n = |A|$

不难有 $A_{n\times n}$为非异阵 $\Leftrightarrow$ A特征值$\lambda_1\ \lambda_n$全不等于0

Theorem 7.1.1

任一复数域上的方阵必定相似于一个上三角阵若是在一般数域 K 上考虑，若增加假设其特征值全部在 K 上同理也会有如此定理

Proof 设A是n阶复方阵,现对n用数学归纳法.当$n = 1$时结论显然.假定对$n-1$阶矩阵结论成立,现对n阶矩阵A来证明.

由代数学基本定理。设λ是A的一个特征值,则存在非零向量α_1,使$A\alpha_1 = \lambda\alpha_1$.

将α_1作为V的一个基向量,并扩展为V的一组基$\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$则$\varphi$在这组基下的表示矩阵形如：$\begin{pmatrix} \lambda & * \\ 0 & A_1 \end{pmatrix}$

令$P = (\alpha_1, \alpha_2 \cdots \alpha_n)$；其中$\alpha_i \in \mathbb{C}^n$； $AP = A(\alpha_1, \alpha_2 \cdots \alpha_n) = (A\alpha_1, A\alpha_2 \cdots A\alpha_n) = (\alpha_1, \alpha_2 \cdots \alpha_n)\begin{pmatrix} \lambda & * \\ 0 & A_1 \end{pmatrix} = P\begin{pmatrix} \lambda & * \\ 0 & A_1 \end{pmatrix}$

$\Rightarrow P^{-1}AP = \begin{pmatrix} \lambda & * \\ 0 & A_1 \end{pmatrix}$.因为$A_1$是一个$n-1$阶矩阵,所以由归纳假设可知,存在非异的$n-1$阶矩阵$Q$,使$Q^{-1}A_1Q$是一个上三角阵.

令$R=\begin{pmatrix}1 & O\\ O & Q\end{pmatrix}\Rightarrow R^{-1}P^{-1}APR=\begin{pmatrix}1 & O\\ O & Q\end{pmatrix}^{-1}\begin{pmatrix}\lambda & *\\ 0 & A_1\end{pmatrix}\begin{pmatrix}1 & O\\ O & Q\end{pmatrix}=\begin{pmatrix}\lambda_1 & *\\ O & Q^{-1}A_1Q\end{pmatrix}$
这是一个上三角阵且与A相似

Lemma 7.2 (相似关系的保持性质)

$f(x)=a_mx^m+\cdots+a_1x+a_0$

$f(A)=a_mA^m+\cdots+a_1A+a_0I_n$

(1) $f\left(P^{-1}AP\right)=P^{-1}f(A)P$

(2) $\left(P^{-1}AP\right)^{-1}=P^{-1}(A)^{-1}P$

(3) $\left(P^{-1}AP\right)^{*}=P^{*}A^{*}\left(P^{*}\right)^{-1}$

Theorem 7.1.2 (矩阵之间的特殊关系对特征值的影响)

设矩阵A是n阶方阵,λ是它的一个特征值.又$f(x)$是一个多项式.

若$\lambda_1,\lambda_2,\cdots,\lambda_n$是$A$的全部特征值,则$f(\lambda_1),f(\lambda_2),\cdots,f(\lambda_n)$是$f(A)$的全部特征值.

设n阶矩阵A适合一个多项式$g(x)$,即$g(A)=0\Rightarrow$ A的特征值λ也必适合$g(\lambda)=0$.

设n阶矩阵A是可逆阵,若A的特征值为$\lambda_1,\lambda_2,\cdots,\lambda_n,\Rightarrow$ 则A^{-1}的特征值为$\lambda_1^{-1},\lambda_2^{-1},\cdots,\lambda_n^{-1}$.

设n阶矩阵$\boldsymbol{A}$的全体特征值为$\lambda_1,\lambda_2,\cdots,\lambda_n$,求证:$\boldsymbol{A}^*$的全体特征值为$\prod\limits_{i\neq 1}\lambda_i,\prod\limits_{i\neq 2}\lambda_i,\cdots,\prod\limits_{i\neq n}\lambda_i$.

Proof 证明因为任意一个n阶矩阵均(复)相似于上三角阵,可设$P^{-1}AP=\begin{pmatrix}\lambda_1 & * & \cdots & *\\ 0 & \lambda_2 & \cdots & *\\ \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & \cdots & \lambda_n\end{pmatrix}$

因为上角阵的和、数乘及乘方仍是上三角阵,经计算不难得到$P^{-1}f(A)P=f\left(P^{-1}AP\right)=\begin{pmatrix}f(\lambda_1) & * & \cdots & *\\ 0 & f(\lambda_2) & \cdots & *\\ \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & \cdots & f(\lambda_n)\end{pmatrix}$.

因此$f(A)$的全部特征值为:$f(\lambda_1),f(\lambda_2),\cdots,f(\lambda_n)$.证毕.

设A的特征值为$\lambda_1\sim\lambda_n$,则$g(A)$的特征值为$g(\lambda_1)\sim g(\lambda_n)$;而$g(A)=0$其特征值只为0

$\Rightarrow g(\lambda_1)\sim g(\lambda_n)=0$

可设$P^{-1}AP=\begin{pmatrix}\lambda_1 & * & \cdots & *\\ 0 & \lambda_2 & \cdots & *\\ \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & \cdots & \lambda_n\end{pmatrix}\Rightarrow$ 两边求逆:$P^{-1}A^{-1}P=\begin{pmatrix}\lambda_1^{-1} & * & \cdots & *\\ 0 & \lambda_2^{-1} & \cdots & *\\ \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & \cdots & \lambda_n^{-1}\end{pmatrix}$

再根据相似矩阵具有相同特征多项式我们知道$\lambda_1^{-1}\sim\lambda_n^{-1}$就是$A^{-1}$的特征值

Proof 因为任一n阶矩阵均复相似于上三角矩阵,故可设$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\begin{pmatrix}\lambda_1 & * & \cdots & *\\ 0 & \lambda_2 & \cdots & *\\ \vdots & \vdots & & \vdots\\ 0 & 0 & \cdots & \lambda_n\end{pmatrix}$注意到上三角矩阵的伴随矩阵仍是上三角矩阵

经计算可得$\boldsymbol{P}^{-1}\boldsymbol{A}^*\boldsymbol{P}=\boldsymbol{P}^*\boldsymbol{A}^*\left(\boldsymbol{P}^{-1}\right)^*=\left(\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}\right)^*=\begin{pmatrix}\prod\limits_{i\neq1}\lambda_i & * & \cdots & * \\ 0 & \prod\limits_{i\neq2}\lambda_i & \cdots & * \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \prod\limits_{i\neq n}\lambda_i\end{pmatrix}$因此$\boldsymbol{A}^*$的全部特征值为$\prod\limits_{i\neq1}\lambda_i,\prod\limits_{i\neq2}\lambda_i,\cdots,\prod\limits_{i\neq n}\lambda_i$.

7.2 对角化

Definition 7.3

设$\varphi \in \mathscr{L}(V^n)$,若$\varphi$能够在一组基下的表示矩阵为对角阵那么我们称$\varphi$是可对角化

Theorem 7.2.1 (矩阵可对角化定理 1(特征向量版本))

$\varphi \in \mathscr{L}(V^n)$ 可对角化 $\Leftrightarrow$ φ有n个线性无关的特征向量

Proof 必要性：设φ可对角化，即φ在某组基下的矩阵为对角阵，设基为$\{e_1 \sim e_n\}$,化为了$\operatorname{diag}\{\lambda_1 \cdots \lambda_n\}$

那么$(\varphi(e_1), \varphi(e_2) \cdots \varphi(e_n)) = (e_1, e_2, \cdots e_n)\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}$

$\Rightarrow \varphi(e_1) = \lambda_1 e_1 \cdots ; \varphi(e_n) = \lambda_n e_n$

这就得到了特征值$\lambda_1 \sim \lambda_n$,特征向量$e_1 \sim e_n$;其中$e_1 \sim e_n$是基必定线性无关有n个

充分性：设φ有n个线性无关的特征向量记为$e_1 \sim e_n$,对应的特征值记为$\lambda_1 \sim \lambda_n$

有$\varphi(e_i) = \lambda_i e_i\ (1 \leqslant i \leqslant n)$ 此时因为$dimV = n; \Rightarrow e_1 \sim e_n$为$V$的一组基

那么$(\varphi(e_1), \varphi(e_2) \cdots \varphi(e_n)) = (e_1, e_2, \cdots e_n)\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}$

$\Rightarrow \varphi$在$e_1 \sim e_n$的表示矩阵为对角阵因而可以对角化

Theorem 7.2.2 (不同特征值对应子空间直和定理)

设$\varphi \in \mathscr{L}(V^n)$.$\lambda_1 \sim \lambda_k$是不同特征值，$V_1 \sim V_m$为对应的特征子空间

$\Rightarrow V_1 + V_2 + \cdots + V_m = V_1 \oplus V_2 \cdots \oplus V_m$

Proof

对k用数学归纳法.若$k=1$,结论显然.现设对$k-1$个特征值$\lambda_1, \lambda_2, \cdots, \lambda_{k-1}$,它们相应的特征子空间$V_1, V_2, \cdots, V_{k-1}$之和是直和.我们要证明$V_1, V_2, \cdots, V_{k-1}, V_k$之和为直和.这只需证明：$V_k \cap (V_1 + V_2 + \cdots + V_{k-1}) = 0$即可. 设$v \in V_k \cap (V_1 + V_2 + \cdots + V_{k-1})$,

$$\text{则} \boldsymbol{v} = v_1 + v_2 + \cdots + v_{k-1}, \text{其中} \boldsymbol{v}_i \in V_i (i = 1, 2, \cdots, k-1). \tag{1}$$

在 (1) 式两边作用φ,得

$$\varphi(v) = \varphi(v_1) + \varphi(v_2) + \cdots + \varphi(v_{k-1}).$$

但$\boldsymbol{v}_1, v_2, \cdots, v_{k-1}$都是$\varphi$的特征向量或零向量因此

$$\lambda_k \boldsymbol{v} = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_{k-1} v_{k-1}. \tag{2}$$

在 (1) 式两边乘以λ_k减去 (2) 式得$0 = (\lambda_k - \lambda_1)\boldsymbol{v}_1 + (\lambda_k - \lambda_2)\boldsymbol{v}_2 + \cdots + (\lambda_k - \lambda_{k-1})\boldsymbol{v}_{k-1}$.

由归纳假设,$V_1 + V_2 + \cdots + V_{k-1}$是直和因此$(\lambda_k - \lambda_i) v_i = 0$,而$\lambda_k - \lambda_i \neq 0$,从而$\boldsymbol{v}_i = 0 (i = 1, 2, \cdots, k-1)$

从而$v = 0$证毕.

Corollary 7.2

(1) 不同特征值的特征向量是线性无关的

(2) φ若有n个不同特征值那么其一定可以对角化

Proof 令$\lambda_1 \sim \lambda_k$为φ的不同特征值。$v_1 \sim v_k$为对应的特征向量
令$c_1v_1 + \cdots + c_kv_k = 0$因为$V_1 \sim V_k$是直和那么根据定理必定有0向量分块表示唯一
即$c_iv_i = 0$但是v_i是特征向量不为$0 \Rightarrow c_i = 0 \Rightarrow$ 线性无关

Theorem 7.2.3 (矩阵可对角化定理 2(特征子空间版本))

设$\varphi \in \mathscr{L}(V^n)$;设$\lambda_1 \sim \lambda_k$为其全部不同的特征值.$V_1 \sim V_k$为特征子空间
$\Rightarrow \varphi$可对角化 $\Leftrightarrow V = V_1 \oplus V_2 \oplus \cdots \oplus V_k$

Proof 充分性:假设$V = V_1 \oplus V_2 \oplus \cdots \oplus V_k$
取V_1基$\left\{e_{11}, \cdots e_{1n_1}\right\} \cdots V_i$基$\left\{e_{i1} \cdots e_{in_i}\right\}$;那么根据直和定理知道
这些基拼成了V的一足基。那么这些基向量一共n个且线性无关，从而可对角化

必要性:设φ可对角化那么有n个线性无关的特征向量$\{e_1 \sim e_n\}$又因为$dimV = n$
$\{e_1 \sim e_n\}$为V的一组基。设前面t_1个关于特征值λ_1的特征向量。接下去t_2个是关于特征值λ_2的特征向量$\cdots$
那么对于$\forall \alpha \in V$; 设$\alpha = k_1e_1 + \cdots + k_ne_n$则$\alpha$可以写为$V_1 \sim V_k$的和。又因为是直和则
$V = V_1 \oplus V_2 \oplus \cdots \oplus V_k$

Definition 7.4 (几何重数, 代数重数, 特征向量系)

设φ是有限维线性空间V上的线性变换,λ_0是φ的一个特征值,V_0是属于λ_0的特征子空间
则称 $\dim V_0$为λ_0的度数或几何重数.λ_0作为φ的特征多项式根的重数称为λ_0的重数或代数重数.
若φ的任一特征值的重数等于度数,则称φ有完全特征向量系.

Lemma 7.3

设$\varphi \in \mathscr{L}(V^n)$;$\lambda_0$为其特征值，几何重数 $\leqslant$ 代数重数

注 λ_0几何重数 $= dimV_{\lambda_0}$而$V_{\lambda_0} = \{v \in \mathbb{C}^n | (\lambda_0 I - A)x = 0\}$ 解空间
所以$dimV_{\lambda_0} = n - rank(\lambda_0 I - A)$

Proof 设$t = dimV_0$几何重数；m为λ_0代数重数
取V_0的一组基$\{e_1 \sim e_t\}$;根据基扩张定理，那么$\{e_1 \sim e_t\} \longrightarrow \{e_1 \sim e_t, e_{t+1} \cdots e_n\}$
我们要计算一个代数重数实际上要根据其特征多项式来看。就得看一个线性变换在给定一组基下的表示矩阵
于是我们有针对$\{e_1 \sim e_t, e_{t+1} \cdots e_n\}$那么有

$$(\varphi(e_1) \cdots \varphi(e_t)\varphi(e_{t+1}) \cdots \varphi(e_n)) = (e_1 \cdots e_t, e_{t+1}, \cdots e_n)\begin{bmatrix} \lambda_0 I_n & P \\ O & Q \end{bmatrix}, \text{其中} \begin{bmatrix} \lambda_0 I_n & P \\ O & Q \end{bmatrix} = A$$

$$\Rightarrow |\lambda I_V - \varphi| = |\lambda I_n - A| = \begin{vmatrix} (\lambda - \lambda_0) I_t & P \\ O & \lambda I_{n-t} - Q \end{vmatrix} = (\lambda - \lambda_0)^t |\lambda I_{n-t} - Q|$$

所以代数重数 $\geqslant t =$ 几何重数

Theorem 7.2.4 (矩阵可对角化定理 3(几何与代数重数版本))

$\varphi \in \mathscr{L}(V^n)$ 可对角化 $\Leftrightarrow$ φ有完全特征向量系

Proof 只要证φ有完全特征向量系 $\Leftrightarrow V = V_1 \oplus V_2 \oplus \cdots \oplus V_k$

其中$\lambda_1 \sim \lambda_k$为全体不同特征值，$V_1 \sim V_k$特征子空间

充分性：设$t_i = dimV_i$;m_i为代数重数，那么显然有$t_i \leqslant m_i\ (1 \leqslant i \leqslant k)$ 且$m_1 + m_2 + \cdots + m_k = n$

$n = dimV = dim\,(V_1 \oplus V_2 \oplus \cdots \oplus V_k) = dimV_1 + dimV_2 + \cdots + dimV_k = t_1 + t_2 + \cdots + t_k \leqslant m_1 + \cdots + m_k = n$

则$t_1 + t_2 + \cdots + t_k \leqslant m_1 + \cdots + m_k$只能取等号所以每个都得$t_i = m_i$

必要性：若有完全特征向量系

$dim\,(V_1 \oplus V_2 \oplus \cdots \oplus V_k) = t_1 + \cdots + t_k = m_1 + \cdots + m_k = n$

那么$V = V_1 \oplus V_2 \oplus \cdots \oplus V_k$

7.3 极小多项式和 Hamilton-Cayley 定理

我们已经知道,数域K上的全体$n\times n$矩阵组成了K上的线性空间,其维数等于n^2.因此下列n^2+1个矩阵必线性相关:
$A^{n^2}, A^{n^2-1}, \cdots, A, I$.也就是说,存在K中不全为零的数$c_i\ (i=0,1,2,\cdots,c_{n^2})$,使得$c_{n^2}A^{n^2}+c_{n^2-1}A^{n^{2\cdots 1}}+\cdots+c_1A+c_0I=0$.
这表明矩阵A适合数域K上的一个多项式$g(x)=c_{n^2}x^{n^2}+\cdots+c_1x+c_0$
我们继续考虑集合$S=\left\{f(x)\in K[x]\,|f(x)\neq 0\text{且}f(A)=O\right\}$则$S\neq\oslash$因为$S$至少有$g(x)$
;我们考虑$\min_{f(x)\in S}\partial(f(x))=k$一定存在
那么一定$\exists h(x)\in S\ st.\deg h(x)=k$.那么首一化得到$m(x)$;
$\Rightarrow m(x)\in K[x]\,; m(x)\in S$即$m(A)=O; m(x)\neq 0; m(x)$次数是$S$中最小的

Definition 7.5 (极小多项式)

假设A为数域$\mathbb{K}$上的n阶方阵。$m(x)\in K[x]$为首一多项式且$m(A)=O$
若$m(x)$是A适合的非零多项式中的次数最小的某一个。那么称$m(x)$为A的极小多项式 根据上文极小多项式一定存在

下面我们考虑极小多项式的唯一性

Lemma 7.4

设$m(x)$是A的极小多项式,若还有另一个多项式$f(x)\in K[x]\,; st.f(A)=O$
$\Rightarrow m(x)\,\big|f(x)$

Proof 由带余除法$f(x)=m(x)q(x)+r(x)\quad\partial(r(x))<\partial(m(x))$
有$f(A)=m(A)q(A)+r(A)\Rightarrow r(A)=O$
若$r(A)\neq 0$;那么$r(x)$为A适合的一个非零多项式且$r(x)$次数$<m(x)$这与$m(x)$次数极小矛盾

Theorem 7.3.1

设A为n阶方阵,那么其极小多项式一定存在且唯一

Proof 假设$m(x), h(x)$都为A的极小多项式。
那么有$m(x)\,|h(x)\quad h(x)\,|m(x)\Rightarrow m(x)=c_1h(x)\,(c_1\neq 0)\Rightarrow$又极小多项式都是首一的$\Rightarrow m(x)=h(x)$

Proposition 7.1

$A=cI_n, A$的极小多项式为$m(x)=x-c$
由题得到极小多项式非零,次数一定大于等于1;且适合$m(x)$那么$x-c$就为极小多项式

Lemma 7.5 (极小多项式是相似关系下的不变量)

相似矩阵具有相同的极小多项式

Proof 设$B=P^{-1}AP$; A的极小$m(x)\quad B$的极小$g(x)$
$g(A)=g\left(PBP^{-1}\right)=Pg(B)P^{-1}=O$那么$A$适合$g(x)\Rightarrow m(x)\,|g(x)$
同理$g(x)\,|m(x)\Rightarrow g(x)=c_1m(x)\,(c_1\neq 0)\Rightarrow$极小为首一$\Rightarrow g(x)=m(x)$

Lemma 7.6 (分块矩阵的极小多项式)

设A是一个分块矩阵$A = \begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_m \end{pmatrix}$求证：$A$的极小多项式等于诸$A_i$的极小多项式之最小公倍式.

Proof 设A_i的极小多项式为$f_i(x)$, A的极小多项式为$f(x)$.诸$f_i(x)$的最小公倍式是$g(x)$

由题$f_i(x)\,|g(x)$ 且$f(A_i) = 0 \Rightarrow$ 则$g(A_i) = 0$, 故$g(A) = \begin{pmatrix} g(A_1) & & & \\ & g(A_2) & & \\ & & \ddots & \\ & & & g(A_m) \end{pmatrix} = 0$

因此A适合于$g(x) \Rightarrow f(x)\,|g(x)$

又因为$f(A) = \begin{pmatrix} f(A_1) & & & \\ & f(A_2) & & \\ & & \ddots & \\ & & & f(A_m) \end{pmatrix} = 0$, 因此对每个$i$有$f(A_i) = 0$

因此A_i适合于$f(x) \Rightarrow$ 有$f_i(x) \mid f(x)$.

而$g(x)$是诸$f_i(x)$的最小公倍式, 故$g(x) \mid f(x)$.

综上所述, $f(x) = c_1 g(x)\ (c_1 \neq 0) \Rightarrow$ 极小多项式是首一的 $\Rightarrow f(x) = g(x)$

Theorem 7.3.2

设A的特征值$\lambda_1 \sim \lambda_k$。若A可对角化。$m(x)$ 为A的极小多项式

$\Rightarrow m(x) = (x - \lambda_1)(x - \lambda_2)\cdots(x - \lambda_k)$

笔记 之后会证明其逆命题也成立即：若一个矩阵的极小多项式没有重根 $\Rightarrow$ 可对角化

Proof 由题得不防适当选择$P\ st.P^{-1}AP = \begin{pmatrix} \lambda_1 I & & & \\ & \lambda_2 I & & \\ & & \ddots & \\ & & & \lambda_k I \end{pmatrix} = B$其中$\lambda_i I$阶数为$\lambda_i$的重数

由题A与B相似。那么A与B极小多项式相同所以$m(x) = [m(\lambda_1 I), \cdots m(\lambda_k I)] = (x - \lambda_1)\cdots(x - \lambda_k)$

Theorem 7.3.3 (特征值一定为极小多项式的根)

设λ_0为A的任一特征值 $\Rightarrow (x - \lambda_0)\,|m(x)$　　$m(x)$ 为A的极小多项式

Proof 由题$m(A) = O \Rightarrow m(\lambda_0) = 0 \Rightarrow (x - \lambda_0)\,|m(x)$

Lemma 7.7

设$A=\begin{pmatrix}\lambda_1 & a_{12} & \cdots & a_{1n}\\ & \lambda_2 & \cdots & a_{2n}\\ & & \ddots & \vdots\\ & & & \lambda_n\end{pmatrix}$其特征值显然为$\lambda_1\sim\lambda_n$

则 $(A-\lambda_1 I_n)(A-\lambda_2 I_n)\cdots(A-\lambda_n I_n)=O$ 即A适合于其特征多项式$f(x)=(x-\lambda_1)(x-\lambda_2)\cdots(x-\lambda_n)$

Proof 设e_i为标准列向量则

$Ae_1=\lambda_1 e_1$; $Ae_2=a_{12}e_1+\lambda_2 e_2$; $Ae_i=a_{1i}e_1+a_{2i}e_2+\cdots+\lambda_i e_i$; $Ae_n=a_{1n}e_1+a_{2n}e_2+\cdots+\lambda_n e_n$

那么为证 $(A-\lambda_1 I_n)(A-\lambda_2 I_n)\cdots(A-\lambda_n I_n)=O$只需证明 $(A-\lambda_1 I_n)(A-\lambda_2 I_n)\cdots(A-\lambda_n I_n)e_i=O\ (1\leqslant i\leqslant n)$

下面证一个更强的结论 $(A-\lambda_1 I_n)(A-\lambda_2 I_n)\cdots(A-\lambda_i I_n)e_i=O\ (\forall 1\leqslant i\leqslant n)$

对i进行归纳，当$i=1$时即 $(A-\lambda_1 I_n)e_1=O$显然成立

假设当 $<i$时成立。那么当 $=i$时

$(A-\lambda_1 I_n)(A-\lambda_2 I_n)\cdots(A-\lambda_i I_n)e_i=(A-\lambda_1 I_n)(A-\lambda_2 I_n)\cdots\left(a_{1n}e_1+\cdots+a_{i-1,i}e_{i-1}\right)$

那么根据归纳假设知道上述 $=O$

Theorem 7.3.4 (Hamilton-Cayley 定理)

设A为n阶方阵，$f(\lambda)$ 为 $|\lambda I_n-A|$ 的特征多项式

$\Rightarrow f(A)=O$

设$\varphi\in\mathscr{L}(V^n)$; 特征多项式$f(\lambda)=|\lambda I_V-\varphi|$ 则$f(\varphi)=O$线性变换

Proof 由题干知道$\exists P$可逆$st.P^{-1}AP=B=\begin{pmatrix}\lambda_1 & a_{12} & \cdots & a_{1n}\\ & \lambda_2 & \cdots & a_{2n}\\ & & \ddots & \vdots\\ & & & \lambda_n\end{pmatrix}$

而相似关系有相同的特征多项式 $\Rightarrow f(\lambda)=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n)$

由上面引理知道$f(B)=O$那么此时$f(A)=f\left(PBP^{-1}\right)=Pf(B)P^{-1}=O$

Corollary 7.3

A为n阶方阵，设特征多项式为$f(\lambda)$; 极小多项式为$m(\lambda)$

(1) $m(\lambda)\,|\,f(\lambda)$，特别的 $\deg m(\lambda)\leqslant n$

(2) $f(\lambda)$ 与$m(\lambda)$ 有相同的根(不计重数)

(3) $f(\lambda)\,|m(\lambda)^n$

Proof 由$Hamilton-Cayley$定理知道A适合于其特征多项式

那么$m(\lambda)\,|f(\lambda)\Rightarrow m(\lambda)$ 次数 $\leqslant f(\lambda)$ 次数

由$m(\lambda)\,|f(\lambda)$ 知道$m(\lambda)$ 根一定是$f(\lambda)$ 的根。且由定理4.3.3知道特征值一定是极小多项式的根

即$f(\lambda)$ 的根一定是$m(\lambda)$ 的根

$f(\lambda)=(\lambda-\lambda_1)^{m_1}(\lambda-\lambda_2)^{m_2}\cdots(\lambda-\lambda_k)^{m_k}$

$m(\lambda)=(\lambda-\lambda_1)^{r_1}(\lambda-\lambda_2)^{r_2}\cdots(\lambda-\lambda_k)^{r_k}$ 其中$r_1\sim r_k\in\mathbb{Z}$

且$m_1 + \cdots + m_k = n \Rightarrow m_i \leqslant n \leqslant n \cdot r_i \Rightarrow f(\lambda) | m(\lambda)^n$

Example 7.1 A有n个不同特征值$\lambda_1 \sim \lambda_n$

$f(\lambda) = (\lambda - \lambda_1) \cdots (\lambda - \lambda_n) = m(\lambda)$

$A = cI_n \quad f(\lambda) = (\lambda - c)^n \quad m(\lambda) = \lambda - c \Rightarrow f(\lambda) = m(\lambda)^n$

7.4 λ 矩阵与矩阵多项式

Definition 7.6 (λ 矩阵定义)

$$A(\lambda)=\begin{pmatrix} a_{11}(\lambda) & a_{12}(\lambda) & \cdots & a_{1n}(\lambda) \\ a_{21}(\lambda) & a_{22}(\lambda) & \cdots & a_{2n}(\lambda) \\ \vdots & \vdots & & \vdots \\ a_{m1}(\lambda) & a_{m2}(\lambda) & \cdots & a_{mn}(\lambda) \end{pmatrix}$$ 其中$a_{ij}(\lambda)$是以λ为未定元的数域$\mathbb{K}$上的多项式

我们称之为多项式矩阵, 或$\lambda-$矩阵.

$\lambda-$矩阵的加法、数乘及乘法与数域上的矩阵运算一样, 只需在运算过程中将数的运算代之以多项式运算即可.

Definition 7.7 (λ 矩阵的初等变换)

对$\lambda-$矩阵$\boldsymbol{A}(\lambda)$施行的下列3种变换称为$\lambda-$矩阵的初等行变换:

(1)将$\boldsymbol{A}(\lambda)$的两行对换;

(2)将$\boldsymbol{A}(\lambda)$的第i行乘以常数c, c是数域$\mathbb{K}$中的非零数;

(3)将$\boldsymbol{A}(\lambda)$的第i行乘以$\mathbb{K}$上的多项式$f(\lambda)$后加到第j行上去.

同理我们可以定义3种$\lambda-$矩阵的初等列变换.

Definition 7.8 (λ 矩阵的相抵或等价)

若$\boldsymbol{A}(\lambda)$, $\boldsymbol{B}(\lambda)$都是$\lambda-$矩阵且$\boldsymbol{A}(\lambda)$经过初等变换后可变为$\boldsymbol{B}(\lambda)$, 则称$\lambda-$矩阵$\boldsymbol{A}(\lambda)$与$\boldsymbol{B}(\lambda)$相抵.

与数字矩阵一样, $\lambda-$矩阵的相抵关系也是一种等价关系, 即

(1)$\boldsymbol{A}(\lambda)$与自身相抵;

(2)若$\boldsymbol{A}(\lambda)$与$\boldsymbol{B}(\lambda)$相抵, 则$\boldsymbol{B}(\lambda)$与$\boldsymbol{A}(\lambda)$相抵;

(3)若$\boldsymbol{A}(\lambda)$与$\boldsymbol{B}(\lambda)$相抵, $\boldsymbol{B}(\lambda)$与$\boldsymbol{C}(\lambda)$相抵, 则$\boldsymbol{A}(\lambda)$与$\boldsymbol{C}(\lambda)$相抵.

Definition 7.9 (初等$\lambda-$矩阵)

初等$\lambda-$矩阵:

(1)将n阶单位阵的第i行与第j行对换, 记为$\boldsymbol{P}_{ij}$;

(2)将n阶单位阵的第i行乘以非零常数c, 记为$\boldsymbol{P}_i(c)$;

(3)将n阶单位阵的第i行乘以多项式$f(\lambda)$后加到第j行上去后得到的矩阵, 记为$\boldsymbol{T}_{ij}(f(\lambda))$.

注意初等 λ 矩阵和普通第一类第二类初等数值矩阵并无不同

笔记 类似的我们会有对 λ 矩阵实施初等变换相当于用初等 λ 矩阵去相乘且有若干个有限个初等 λ 矩阵相乘仍然为初等 λ 矩阵

Example 7.2 下列$\lambda-$矩阵的变换不是$\lambda-$矩阵的初等变换: $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \to \begin{pmatrix} \lambda & \lambda \\ 0 & 1 \end{pmatrix}$.

这是因为前面一个矩阵的第一行乘以λ不是$\lambda-$矩阵的初等变换.

同理下面的变换需第一行乘以λ^{-1}, 因此也不是$\lambda-$矩阵的初等变换: $\begin{pmatrix} \lambda & 0 \\ 0 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

Definition 7.10 (可逆 λ 矩阵)

若$\boldsymbol{A}(\lambda),\boldsymbol{B}(\lambda)$都是$n$阶$\lambda-$矩阵,且$\boldsymbol{A}(\lambda)\boldsymbol{B}(\lambda)=\boldsymbol{B}(\lambda)\boldsymbol{A}(\lambda)=\boldsymbol{I}_n$

则称$\boldsymbol{B}(\lambda)$是$\boldsymbol{A}(\lambda)$的逆λ矩阵,这时称$\boldsymbol{A}(\lambda)$为可逆λ矩阵,有时在不引起混淆的情形下,简称之为可逆阵.

笔记 (1) 不难验证,初等λ矩阵都是可逆阵且逆阵为同类的初等λ矩阵

(2) 有限个可逆λ矩阵的乘积仍然是可逆λ矩阵

Definition 7.11 (矩阵多项式)

设$\boldsymbol{M}(\lambda)$是一个n阶$\lambda-$矩阵,则$\boldsymbol{M}(\lambda)$可以化为如下形状:$\boldsymbol{M}(\lambda)=\boldsymbol{M}_m\lambda^m+\boldsymbol{M}_{m-1}\lambda^{m-1}+\cdots+\boldsymbol{M}_0$,

定义 $\deg M(\lambda)=m\ (若M_m\neq O)$

笔记 (1) 两个矩阵多项式相等 $:=\deg(f(\lambda))=\deg(g(\lambda))$ 且$M_i=N_i$ (即系数矩阵相同)

(2) 可定义两个矩阵多项式加法,数乘,但是注意两个多项式乘法却会出问题,因为系数矩阵的乘积可能会为零矩阵

Lemma 7.8

$\deg(M(\lambda)\cdot N(\lambda))\leqslant\deg(M(\lambda))+\deg(N(\lambda))$

若$M(\lambda)$或$N(\lambda)$的首项系数矩阵为可逆阵即M_n或者N_n可逆那么等号成立

Proof $M(\lambda)$与$N(\lambda)$相乘首先为$M_mN_n\lambda^{m+n}$这时M_mN_n两个非零阵，且有一个为可逆阵，那么相乘必定不会为零阵

Theorem 7.4.1 (λ 矩阵带余除法)

设$M(\lambda)$与$N(\lambda)$是两个n阶$\lambda-$矩阵且都不等于零.又设B为n阶数字矩阵

则必存在$\lambda-$矩阵$\boldsymbol{Q}(\lambda)$及$\boldsymbol{S}(\lambda)$和数字矩阵$\boldsymbol{R}$及$\boldsymbol{T}$,使下式成立:$\begin{cases}\boldsymbol{M}(\lambda)=(\lambda\boldsymbol{I}-\boldsymbol{B})\boldsymbol{Q}(\lambda)+\boldsymbol{R},\\ \boldsymbol{N}(\lambda)=\boldsymbol{S}(\lambda)(\lambda\boldsymbol{I}-\boldsymbol{B})+\boldsymbol{T}.\end{cases}$

Proof 将$M(\lambda)$写为$\boldsymbol{M}(\lambda)=\boldsymbol{M}_m\lambda^m+\boldsymbol{M}_{m-1}\lambda^{m-1}+\cdots+\boldsymbol{M}_0$,其中$\boldsymbol{M}_m\neq\boldsymbol{O}$.

可对m用归纳法,若$m=0$,则已适合要求(取$\boldsymbol{Q}(\lambda)=\boldsymbol{O}$).

现设对小于m次的矩阵多项式,成立.

令$\boldsymbol{Q}_1(\lambda)=\boldsymbol{M}_m\lambda^{m-1}$,则$\boldsymbol{M}(\lambda)-(\lambda\boldsymbol{I}-\boldsymbol{B})\boldsymbol{Q}_1(\lambda)=(\boldsymbol{B}\boldsymbol{M}_m+\boldsymbol{M}_{m-1})\lambda^{m-1}+\cdots+\boldsymbol{M}_0$.

上式是一个次数小于m的矩阵多项式,由归纳假设得$\boldsymbol{M}(\lambda)-(\lambda\boldsymbol{I}-\boldsymbol{B})\boldsymbol{Q}_1(\lambda)=(\lambda\boldsymbol{I}-\boldsymbol{B})\boldsymbol{Q}_2(\lambda)+\boldsymbol{R}$.

于是$\boldsymbol{M}(\lambda)=(\lambda\boldsymbol{I}-\boldsymbol{B})\left[\boldsymbol{Q}_1(\lambda)+\boldsymbol{Q}_2(\lambda)\right]+\boldsymbol{R}$.令$\boldsymbol{Q}(\lambda)=\boldsymbol{Q}_1(\lambda)+\boldsymbol{Q}_2(\lambda)$即得式.

同理可证另一边带余除法.

Theorem 7.4.2 (λ 矩阵的相抵与原矩阵的相似互为充要)

设A,B是数域$\mathbb{K}$上的矩阵,则$\boldsymbol{A}$与$\boldsymbol{B}$相似的 $\Leftrightarrow$ λ矩阵:$\lambda\boldsymbol{I}-\boldsymbol{A}$与$\lambda\boldsymbol{I}-\boldsymbol{B}$相抵.

Proof 反过来,若$\lambda\boldsymbol{I}-\boldsymbol{A}$与$\lambda\boldsymbol{I}-\boldsymbol{B}$相抵,即存在$\boldsymbol{M}(\lambda)$及$\boldsymbol{N}(\lambda)$,使

$\boldsymbol{M}(\lambda)(\lambda\boldsymbol{I}-\boldsymbol{A})\boldsymbol{N}(\lambda)=\lambda\boldsymbol{I}-\boldsymbol{B}$(1)

其中$\boldsymbol{M}(\lambda)$与$\boldsymbol{N}(\lambda)$都是有限个初等矩阵之积的复合,因而都是可逆阵.因此可将(1)式写为

$\boldsymbol{M}(\lambda)(\lambda\boldsymbol{I}-\boldsymbol{A})=(\lambda\boldsymbol{I}-\boldsymbol{B})\boldsymbol{N}(\lambda)^{-1}$(2)

由带余除法,可设$\boldsymbol{M}(\lambda)=(\lambda\boldsymbol{I}-\boldsymbol{B})\boldsymbol{Q}(\lambda)+\boldsymbol{R}$,代入(2)式经整理得

$\boldsymbol{R}(\lambda\boldsymbol{I}-\boldsymbol{A})=(\lambda\boldsymbol{I}-\boldsymbol{B})\left[\boldsymbol{N}(\lambda)^{-1}-\boldsymbol{Q}(\lambda)(\lambda\boldsymbol{I}-\boldsymbol{A})\right]$.

上式左边是次数小于等于1的矩阵多项式,因此上式右边中括号内的矩阵多项式的次数必须小于等于零,即必是一个常数矩阵,设为$\boldsymbol{P}$.

即$\boldsymbol{P} = \boldsymbol{N}(\lambda)^{-1} - \boldsymbol{Q}(\lambda)(\lambda\boldsymbol{I} - \boldsymbol{A})$

于是$\boldsymbol{R}(\lambda\boldsymbol{I} - \boldsymbol{A}) = (\lambda\boldsymbol{I} - \boldsymbol{B})\boldsymbol{P}$.(3)

(3) 式又可整理为 $(\boldsymbol{R} - \boldsymbol{P})\lambda = \boldsymbol{RA} - \boldsymbol{BP}$.

再次比较次数得$\boldsymbol{R} = \boldsymbol{P}$, 那么 $(\boldsymbol{R} - \boldsymbol{P})\lambda = \boldsymbol{O} = \boldsymbol{RA} - \boldsymbol{BP}$.

$\Rightarrow \boldsymbol{RA} = \boldsymbol{BP}$.(4)

现只需证明$\boldsymbol{P}$是一个非异阵即可.由假设$\boldsymbol{P} = \boldsymbol{N}(\lambda)^{-1} - \boldsymbol{Q}(\lambda)(\lambda\boldsymbol{I} - \boldsymbol{A})$, 将上式两边右乘$\boldsymbol{N}(\lambda)$并移项得

$\boldsymbol{PN}(\lambda) + \boldsymbol{Q}(\lambda)(\lambda\boldsymbol{I} - \boldsymbol{A})\boldsymbol{N}(\lambda) = \boldsymbol{I}$.(5)

但由 (2) 可知 $(\lambda\boldsymbol{I} - \boldsymbol{A})\boldsymbol{N}(\lambda) = \boldsymbol{M}(\lambda)^{-1}(\lambda\boldsymbol{I} - \boldsymbol{B})$, 带入 (5) 因此

$\boldsymbol{PN}(\lambda) + \boldsymbol{Q}(\lambda)\boldsymbol{M}(\lambda)^{-1}(\lambda\boldsymbol{I} - \boldsymbol{B}) = \boldsymbol{I}$.(6)

再由带余除法可设$\boldsymbol{N}(\lambda) = \boldsymbol{S}(\lambda)(\lambda\boldsymbol{I} - \boldsymbol{B}) + \boldsymbol{T}$, 代入 (6) 式并整理得

$\left[\boldsymbol{PS}(\lambda) + \boldsymbol{Q}(\lambda)\boldsymbol{M}(\lambda)^{-1}\right](\lambda\boldsymbol{I} - \boldsymbol{B}) = \boldsymbol{I} - \boldsymbol{PT}$.

上式右边是次数小于等于零的矩阵多项式, 因此上式左边中括号内的矩阵多项式必须为零, 从而

$\left[\boldsymbol{PS}(\lambda) + \boldsymbol{Q}(\lambda)\boldsymbol{M}(\lambda)^{-1}\right](\lambda\boldsymbol{I} - \boldsymbol{B}) = \boldsymbol{I} - \boldsymbol{PT} = \boldsymbol{O}$

$\Rightarrow \boldsymbol{PT} = \boldsymbol{I}$, 即$\boldsymbol{P}$是非异阵.(7)

由 (4) (7) 即可

7.5 矩阵的法式

在上一节中,我们把矩阵的相似归结为$\lambda-$矩阵的相抵.现在我们来求$\lambda-$矩阵的相抵标准型.
我们自然地希望任一$\lambda-$矩阵相抵于一个对角$\lambda-$矩阵.

Lemma 7.9

设$\boldsymbol{A}(\lambda)=\left(a_{ij}(\lambda)\right)_{m\times n}$是任一非零$\lambda-$矩阵
则$\boldsymbol{A}(\lambda)$必相抵于这样的一个$\lambda-$矩阵$\boldsymbol{B}(\lambda)=\left(b_{ij}(\lambda)\right)_{m\times n}$,其中$b_{11}(\lambda)\neq 0$且$b_{11}(\lambda)$可整除$\boldsymbol{B}(\lambda)$中的任一元素$b_{ij}(\lambda)$.

Proof 设$k=\min\left\{\deg a_{ij}(\lambda)\mid a_{ij}(\lambda)\neq 0, 1\leq i\leq m; 1\leq j\leq n\right\}$,我们对$k$用数学归纳法.
首先,经行对换及列对换可将$\boldsymbol{A}(\lambda)$的第$(1,1)$元素变成次数最低的非零多项式,因此不妨设$a_{11}(\lambda)\neq 0$且$\deg a_{11}(\lambda)=k$.
若$k=0$,则$a_{11}(\lambda)$是一个非零常数,结论显然成立.
假设对非零元素次数的最小值小于k的任一$\lambda-$矩阵,引理的结论成立,现考虑非零元素次数的最小值等于k的$\lambda-$矩阵$\boldsymbol{A}(\lambda)$.
若$a_{11}(\lambda)$可整除所有的$a_{ij}(\lambda)$,则结论已成立.
若否,设在第一列中有元素$a_{i1}(\lambda)$不能被$a_{11}(\lambda)$整除,作带余除法:$a_{i1}(\lambda)=a_{11}(\lambda)q(\lambda)+r(\lambda)$.
用$-q(\lambda)$乘以第一行加到第i行上,第$(i,1)$元素就变为$r(\lambda)$.注意到$r(\lambda)\neq 0$且$\deg r(\lambda)<\deg a_{11}(\lambda)=k$,由归纳假设即知结论成立.
同样的方法可施于第一行.因此我们不妨设$a_{11}(\lambda)$可整除第一行及第一列.
这时,设$a_{21}(\lambda)=a_{11}(\lambda)g(\lambda)$.将第一行乘以$-g(\lambda)$加到第二行上,则第$(2,1)$元素变为零.
用同样的方法可消去第一行、第一列除$a_{11}(\lambda)$以外的所有元素,于是$\boldsymbol{A}(\lambda)$经初等变换后变成下列形状:

$$\begin{pmatrix} a_{11}(\lambda) & 0 & \cdots & 0 \\ 0 & a'_{22}(\lambda) & \cdots & a'_{2n}(\lambda) \\ \vdots & \vdots & & \vdots \\ 0 & a'_{m2}(\lambda) & \cdots & a'_{mn}(\lambda) \end{pmatrix}.$$

这时,若$a_{11}(\lambda)$可整除所有其他元素,则结论已成立.
若否,比如$a_{11}(\lambda)$不能整除$a'_{ij}(\lambda)$,则将第i行加到第一行上去,这时在第一行又出现了一元素$a'_{ij}(\lambda)$,它不能被$a_{11}(\lambda)$整除.
重复上面的做法,通过归纳假设即可得到结论.

Theorem 7.5.1

设$\boldsymbol{A}(\lambda)$是一个n阶$\lambda-$矩阵
则$\boldsymbol{A}(\lambda)$相抵于$\mathrm{diag}\left\{d_1(\lambda), d_2(\lambda), \cdots, d_r(\lambda); 0, \cdots, 0\right\}$其中$d_i(\lambda)$是非零首一多项式且$d_i(\lambda)\mid d_{i+1}(\lambda)(i=1,2,\cdots,r-1)$.
上述得到的矩阵成为λ矩阵的法式或相抵标准型

Proof 对n用数学归纳法,当$n=1$时结论显然
现设$\boldsymbol{A}(\lambda)$是n阶$\lambda-$矩阵.由引理可知$\boldsymbol{A}(\lambda)$相抵于n阶$\lambda-$矩阵$\boldsymbol{B}(\lambda)=\left(b_{ij}(\lambda)\right)$,其中$b_{11}(\lambda)\mid b_{ij}(\lambda)$对一切$i,j$成立.
因此,将$\boldsymbol{B}(\lambda)$的第一行乘以某个λ的多项式加到第二行上去便可消去第二行第一列元素$b_{21}(\lambda)$.
同理可依次消去第一列除$b_{11}(\lambda)$外的所有元素.用类似方法消去第一行其余元素,这样便得到了一个矩阵:

$$\begin{pmatrix} b_{11}(\lambda) & 0 & \cdots & 0 \\ 0 & b'_{22}(\lambda) & \cdots & b'_{2n}(\lambda) \\ \vdots & \vdots & & \vdots \\ 0 & b'_{n2}(\lambda) & \cdots & b'_{nn}(\lambda) \end{pmatrix}.$$

不难看出,这时$b_{11}(\lambda)$仍可整除所有的$b'_{ij}(\lambda)$.
若$b_{i1}(\lambda)=b_{11}(\lambda)q(\lambda)+r(\lambda)$我们做这样的操作第一行乘以$-q(\lambda)$加到第$i$行那么就有$b'_{ij}(\lambda)=-q(\lambda)b_{1j}(\lambda)+b_{ij}(\lambda)$
由引理可知$b_{11}(\lambda)|b_{1j}(\lambda)\quad b_{11}(\lambda)|b_{ij}(\lambda)\Rightarrow b_{11}(\lambda)|b'_{ij}(\lambda)$

设c为$b_{11}(\lambda)$的首项系数,记$d_1(\lambda)=c^{-1}b_{11}(\lambda)$,设$\overline{\boldsymbol{B}}(\lambda)$为上面的矩阵中右下方的$n-1$阶$\lambda-$矩阵

则由归纳假设可知存在$\boldsymbol{P}(\lambda)$及$\boldsymbol{Q}(\lambda)$,使$\boldsymbol{P}(\lambda)\overline{\boldsymbol{B}}(\lambda)\boldsymbol{Q}(\lambda)=\operatorname{diag}\{d_2(\lambda),\cdots,d_r(\lambda);0,\cdots,0\}$,且$d_i(\lambda)\mid d_{i+1}(\lambda)(i=2,\cdots,r-1)$
其中$\boldsymbol{P}(\lambda)$与$\boldsymbol{Q}(\lambda)$可写成为有限个初等$\lambda-$矩阵之积.于是

$$\begin{pmatrix}c^{-1}&0\\0&\boldsymbol{P}(\lambda)\end{pmatrix}\begin{pmatrix}b_{11}(\lambda)&0\\0&\overline{\boldsymbol{B}}(\lambda)\end{pmatrix}\begin{pmatrix}1&0\\0&\boldsymbol{Q}(\lambda)\end{pmatrix}=\operatorname{diag}\{d_1(\lambda),d_2(\lambda),\cdots,d_r(\lambda);0,\cdots,0\}$$

且$\begin{pmatrix}1&0\\0&\boldsymbol{P}(\lambda)\end{pmatrix},\begin{pmatrix}1&0\\0&\boldsymbol{Q}(\lambda)\end{pmatrix}$可写为若干个$n$阶初等$\lambda-$矩阵之积.
于是只需证明$d_1(\lambda)\mid d_2(\lambda)$即可.但这点很容易看出,事实上由于$\overline{\boldsymbol{B}}(\lambda)$中的任一元素均可被$d_1(\lambda)$整除
因此$\boldsymbol{P}(\lambda)\overline{\boldsymbol{B}}(\lambda)\boldsymbol{Q}(\lambda)$中的任一元素也可被$d_1(\lambda)$整除,这就证明了定理.

笔记 (1):一般的同理可以得到$A_{m\times n}(\lambda)$相抵与$\begin{bmatrix}d_1(\lambda)&&&\\&d_2(\lambda)&&O\\&&d_3(\lambda)&\\O&&&O\end{bmatrix}$

(2):若方阵$A(\lambda)$秩等于n,但是未必行列式$\neq 0$,例子$\begin{bmatrix}1&0\\0&\lambda\end{bmatrix}$

Lemma 7.10

$A(\lambda),B(\lambda)$为n阶λ矩阵

(1) $|A(\lambda)B(\lambda)|=|A(\lambda)||B(\lambda)|$

(2) $A(\lambda)A(\lambda)^*=A(\lambda)^*A(\lambda)=|A(\lambda)|I_n$

Proof (1) $f(\lambda)=|A(\lambda)B(\lambda)|-|A(\lambda)||B(\lambda)|$是$\lambda$多项式那么$\forall a\in\mathbb{K},f(a)=0$
$\Rightarrow f(\lambda)=0$即可
(2) 令$(f_{ij}(\lambda))_{n\times n}=A(\lambda)A(\lambda)^*-|A(\lambda)|I_n$故$f_{ij}(\lambda)\in\mathbb{K}[\lambda]$
$\forall a\in\mathbb{K}$,都有$f_{ij}(a)=0\Rightarrow f_{ij}(\lambda)=0$即可

Theorem 7.5.2

以下说法等价

(1) $A(\lambda)$为可逆λ矩阵

(2) $|A(\lambda)|$是非零常数

(3) $A(\lambda)$相抵标准型为I_n

(4) $A(\lambda)$只通过初等行(列变换)变为I_n

(5) $A(\lambda)$是初等λ矩阵的乘积

Proof (1) $\Rightarrow$ (2)
$\exists B(\lambda)\quad st.A(\lambda)B(\lambda)=B(\lambda)A(\lambda)=I_n\Rightarrow 1=|A(\lambda)B(\lambda)|=|A(\lambda)||B(\lambda)|$
则有$|A(\lambda)|$为非零常数

(2) $\Rightarrow$ (3)
$\exists P(\lambda),Q(\lambda)$是初等$\lambda$矩阵乘积使得$P(\lambda)A(\lambda)Q(\lambda)=\operatorname{diag}\{d_1(\lambda)\cdots d_r(\lambda),0\cdots 0\}$
其中$d_i(\lambda)$为非零首一多项式且$d_1(\lambda)|d_2(\lambda)|\cdots|d_r(\lambda)$
其中$P(\lambda),Q(\lambda)$为初等可逆矩阵那么行列式为非零常数由(1)可得
此时$|P(\lambda)A(\lambda)Q(\lambda)|=|P(\lambda)||A(\lambda)||Q(\lambda)|=|d_1(\lambda)|\cdots|d_r(\lambda)|\times 0\cdots 0$
上式左边$\neq 0$，故右边不会有0. $\Rightarrow r=n$,且有$0\neq c=d_1(\lambda)\cdots d_n(\lambda)$而右边是$\lambda$多项式且$d_1(\lambda)|d_2(\lambda)|\cdots|d_r$
$\Rightarrow d_i(\lambda)$只能为非零常数又首一则为1,相抵于I_n

(3) $\Rightarrow$ (4)

$\exists P(\lambda), Q(\lambda)$ 是初等λ矩阵乘积使得$P(\lambda)A(\lambda)Q(\lambda) = I_n$

$\Rightarrow P(\lambda)A(\lambda) = Q^{-1}(\lambda) \Rightarrow Q(\lambda)P(\lambda)A(\lambda) = I_n$即可

(4) $\Rightarrow$ (5)

$P_r(\lambda)\cdots P_1(\lambda)A(\lambda) = I_n \Rightarrow A(\lambda) = P_r^{-1}(\lambda)\cdots P_1^{-1}(\lambda)$ 即可

(5) $\Rightarrow$ (1)

$A(\lambda)$ 是初等λ矩阵乘积从而是可逆λ矩阵乘积从而是可逆阵

Corollary 7.4 (特征矩阵的法式形状)

设$\boldsymbol{A}$是数域$\mathbb{K}$上的n阶矩阵

则$\boldsymbol{A}$的特征矩阵$\lambda \boldsymbol{I}_n - \boldsymbol{A}$必相抵于$\mathrm{diag}\{1, \cdots, 1, d_1(\lambda), \cdots, d_m(\lambda)\}$ 其中$d_i(\lambda) \mid d_{i+1}(\lambda)(i = 1, \cdots, m-1)$.

Proof

由上述定理, 存在$\boldsymbol{P}(\lambda), \boldsymbol{Q}(\lambda)$, 使$\boldsymbol{P}(\lambda)(\lambda \boldsymbol{I}_n - \boldsymbol{A})\boldsymbol{Q}(\lambda) = \mathrm{diag}\{d_1(\lambda), d_2(\lambda), \cdots, d_r(\lambda); 0, \cdots, 0\}$, 其中$\boldsymbol{P}(\lambda), \boldsymbol{Q}(\lambda)$为若干个初等$\lambda$矩阵之积

根据$\lambda-$矩阵初等变换的定义以及行列式的性质可得, 上式左边的行列式等于$c\,|\lambda \boldsymbol{I}_n - \boldsymbol{A}|$, 其中$c$是一个非零常数

从而上式右边的行列式不为零, 故$r = n$.把$d_i(\lambda)$中的常数多项式写出来 (因是首一多项式, 故为常数1), 就可得结论.

笔记 若上述 $\deg d_i(\lambda) \geqslant 1 \Rightarrow A = cI_n$

$|\lambda I_n - A| = d_1(\lambda)\cdots d_n(\lambda)$ 故 $\deg n = \deg \sum_{i=1}^{n} d_i(\lambda) \geqslant n$从而$d_i(\lambda)$ 都为首一一次多项式

又先后整除 $\Rightarrow d_i(\lambda) = \lambda - c \Rightarrow \lambda I_n - A$相抵于$\mathrm{diag}\{\lambda - c \cdots\cdots \lambda - c\} = \lambda I_n - cI_n$

$\Rightarrow A$相似于$cI_n \Rightarrow A = cI_n$

7.6 不变因子

Definition 7.12 (法式的行列式因子)

设$\boldsymbol{A}(\lambda)$是n阶$\lambda-$矩阵, k是小于等于n的某个正整数.
如果$\boldsymbol{A}(\lambda)$的所有k阶子式的最大公因子 (它是首一多项式) 不等于零, 则称这个多项式为$\boldsymbol{A}(\lambda)$的k阶行列式因子, 记为$D_k(\lambda)$.
如果$\boldsymbol{A}(\lambda)$的所有k阶子式都等于零, 则规定$\boldsymbol{A}(\lambda)$的k阶行列式因子为零.

Proposition 7.2 (对角 λ 矩阵的行列式因子)

求下列矩阵的行列式因子$\boldsymbol{A}(\lambda)=\begin{pmatrix} d_1(\lambda) & & & & & \\ & \ddots & & & & \\ & & d_r(\lambda) & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}$, 其中$d_i(\lambda)$为非零首一多项式且$d_i(\lambda)\mid d_{i+1}(\lambda)(i=1,2,\cdots,r-1)$.

Proof $\boldsymbol{A}(\lambda)$的非零行列式因子为$D_1(\lambda)=d_1(\lambda), D_2(\lambda)=d_1(\lambda)d_2(\lambda),\cdots,D_r(\lambda)=d_1(\lambda)d_2(\lambda)\cdots d_r(\lambda)$

Lemma 7.11 (行列式因子的整除性)

设$D_1(\lambda),D_2(\lambda),\cdots,D_r(\lambda)$是$\boldsymbol{A}(\lambda)$的非零行列式因子, 则$D_i(\lambda)\mid D_{i+1}(\lambda), i=1,2,\cdots,r-1$.
(这是上个命题的推广, 由上个命题我们很轻易的可以知道法式是满足该引理结论的)

Proof 设A_{i+1}是$\boldsymbol{A}(\lambda)$的任一$i+1$阶子式, 即在$\boldsymbol{A}(\lambda)$中任意取出$i+1$行及$i+1$列组成的行列式.
将这个行列式按某一行展开, 则它的每一展开项都是一个多项式与一个i阶子式的乘积.
由于$D_i(\lambda)$是所有i阶子式的公因子, 因此$D_i(\lambda)\mid A_{i+1}$.而$D_{i+1}(\lambda)$是所有$i+1$阶子式的最大公因子
因此$D_i(\lambda)\mid D_{i+1}(\lambda)$对一切$i=1,2,\cdots,r-1$成立.

Definition 7.13 (不变因子定义)

设$D_1(\lambda),D_2(\lambda),\cdots,D_r(\lambda)$是$\lambda-$矩阵$\boldsymbol{A}(\lambda)$的非零行列式因子
则$g_1(\lambda)=D_1(\lambda), g_2(\lambda)=D_2(\lambda)/D_1(\lambda),\cdots,g_r(\lambda)=D_r(\lambda)/D_{r-1}(\lambda)$称为$\boldsymbol{A}(\lambda)$的不变因子.

Proposition 7.3 (法式的不变因子)

法式矩阵的不变因子为$d_1(\lambda),d_2(\lambda),\cdots,d_r(\lambda)$.且有$d_1(\lambda)\,|d_2(\lambda)\,|\cdots|d_r(\lambda)$

Theorem 7.6.1 (行列式因子与不变因子在相抵下保持)

相抵的λ矩阵有相同的行列式因子, 从而有相同的不变因子, 也有相同的法式.

Proof 设$A(\lambda)$与$B(\lambda)$相抵故存在可逆λ矩阵$P(\lambda),Q(\lambda)$使得$B(\lambda)=P(\lambda)A(\lambda)Q(\lambda)$
设$A(\lambda)$与$B(\lambda)$的行列式因子为$D_k(\lambda),E_k(\lambda)\ (1\leqslant k\leqslant n)$
要证 : $D_k(\lambda)=E_k(\lambda)$

故$B(\lambda)\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{pmatrix} \xlongequal{Cauchy-Binet} \sum P(\lambda)\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ r_1 & r_2 & \cdots & r_k \end{pmatrix} A(\lambda)\begin{pmatrix} r_1 & r_2 & \cdots & r_k \\ s_1 & s_2 & \cdots & s_k \end{pmatrix} Q(\lambda)\begin{pmatrix} s_1 & s_2 & \cdots & s_k \\ j_1 & j_2 & \cdots & j_k \end{pmatrix}$

1.若$D_k(\lambda)=0$由上式得到$E_k(\lambda)=0$

2.若$D_k(\lambda)\neq 0$, 那么$D_k(\lambda)\,|B(\lambda)\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{pmatrix}$即$D_k(\lambda)$是$B(\lambda)$所有$k$阶子式的公因式故$D_k(\lambda)\,|E_k(\lambda)$

但是在相抵关系下$P^{-1}(\lambda)B(\lambda)Q^{-1}(\lambda)=A(\lambda)$那么同理再做一遍即可

总之

$1^\circ D_k(\lambda)=0\Leftrightarrow E_k(\lambda)=0$

2°若$D_k(\lambda)\neq 0$且$E_k(\lambda)\neq 0$时, $D_k(\lambda)\,|E_k(\lambda).E_k(\lambda)\,|D_k(\lambda)$

$\Rightarrow\exists 0\neq c$ 使得$D_k(\lambda)=cE_k(\lambda)$又由首一性故$D_k(\lambda)=E_k(\lambda)$从而行列式因子相同, 不变因子也相同

Corollary 7.5 (法式的不变因子 (行列式因子))

设n阶λ矩阵$\boldsymbol{A}(\lambda)$的法式为$\Lambda=\operatorname{diag}\{d_1(\lambda),d_2(\lambda),\cdots,d_r(\lambda);0,\cdots,0\}$其中$d_i(\lambda)$是非零首一多项式且$d_i(\lambda)\mid d_{i+1}(\lambda)$

(1) $\boldsymbol{A}(\lambda)$的不变因子为$d_1(\lambda),d_2(\lambda),\cdots,d_r(\lambda)$.

(2) 特别, 法式和不变因子之间相互唯一确定.

Proof 由相抵关系知道$A(\lambda)$和$\boldsymbol{\Lambda}=\operatorname{diag}\{d_1(\lambda),d_2(\lambda),\cdots,d_r(\lambda);0,\cdots,0\}$具有相同的行列式因子和不变因子
而前文已经计算过对角λ矩阵的行列式因子和不变因子故可得

Theorem 7.6.2 (λ 矩阵相抵关系下，下的全系不变量)

设$\boldsymbol{A}(\lambda),\boldsymbol{B}(\lambda)$为$n$阶$\lambda-$矩阵, 则$\boldsymbol{A}(\lambda)$与$\boldsymbol{B}(\lambda)$相抵$\Leftrightarrow$它们有相同的法式 (不变因子)

Proof 若$\boldsymbol{A}(\lambda)$与$\boldsymbol{B}(\lambda)$有相同的法式, 显然它们相抵.
若$\boldsymbol{A}(\lambda)$与$\boldsymbol{B}(\lambda)$相抵, 则法式$\Lambda_A=\Lambda_B$相抵故由定理知$\boldsymbol{A}(\lambda)$与$\boldsymbol{B}(\lambda)$有相同的不变因子, 从而有相同的法式.

Corollary 7.6

法式或相抵标准型不依赖于λ矩阵的初等变换选取

Proof $A\sim\Lambda_1\quad A\sim\Lambda_2$那么由传递性$\Lambda_1\sim\Lambda_2$那么具有相同的法式 (不变因子) 则$\Lambda_1=\Lambda_2$

Theorem 7.6.3 (相似矩阵的全系不变量)

设A与B相似$\Leftrightarrow\lambda I-A$与$\lambda I-B$具有相同的行列式因子组或不变因子组 (法式)

把$\lambda I-A$的行列式因子组或不变因子组称为A的行列式因子组或不变因子组

Proof 由前一节A与B相似$\Leftrightarrow\lambda I-A$与$\lambda I-B$相抵$\Leftrightarrow\lambda I-A$与$\lambda I-B$具有相同的行列式因子组或不变因子组 (法式)

Corollary 7.7

1.设$A,B\in\mathrm{M}_n(\mathbb{F})$且$\mathbb{F}\subseteq\mathbb{K}$则$A$与$B$在$\mathbb{F}$上相似$\Leftrightarrow A$与$B$在$\mathbb{K}$上相似

2.不变因子组 (行列式因子组) 在基域扩张下不改变

Proof 设$A,B\in\mathrm{M}_n(\mathbb{F})$且$\mathbb{F}\subseteq\mathbb{K}$则$A$与$B$在$\mathbb{F}$上相似$\Leftrightarrow A$与$B$在$\mathbb{K}$上相似
显然必要性显然, 下证充分性

由定理A与B在$\mathbb{K}$上相似 $\Leftrightarrow \lambda I - A$与$\lambda I - B$具有相同的不变因子组 (法式)
实际上$\lambda I - A$与$\lambda I - B$也是$\mathbb{F}$上的λ矩阵.但由于法式或相抵标准型结果不依赖初等变换选取
故只仅仅选取$\mathbb{F}$上的变换就可以得到$\lambda I - A$与$\lambda I - B$法式不防设为$\operatorname{diag}\{1, 1 \cdots, 1, d_1(\lambda) \cdots d_k(\lambda)\}$
所以存在$\mathbb{F}$上的可逆λ阵使得$P(\lambda)(\lambda I - A) Q(\lambda) = \operatorname{diag}\{1, 1 \cdots, 1, d_1(\lambda) \cdots d_k(\lambda)\} = M(\lambda)(\lambda I - B) N(\lambda)$
$\Rightarrow \lambda I - B = M^{-1}(\lambda) P(\lambda)(\lambda I - A) Q(\lambda) N^{-1}(\lambda)$
那么由于$M^{-1}(\lambda) P(\lambda) Q(\lambda) N^{-1}(\lambda)$ 都是$\mathbb{F}$上的矩阵故$\lambda I - B$与$\lambda I - A$作为$\mathbb{F}$上的矩阵相抵
故A，B作为$\mathbb{F}$上的矩阵相似

Theorem 7.6.4 (特征矩阵的不变因子)

矩阵$\boldsymbol{A}$的特征矩阵$\lambda \boldsymbol{I} - \boldsymbol{A}$的法式为$\operatorname{diag}\{1, \cdots, 1, d_1(\lambda), \cdots, d_k(\lambda)\}$ 其中$d_i(\lambda)$为首一非常数多项式且$d_i(\lambda) \mid d_{i+1}(\lambda)(i = 1, 2, \cdots, k-1)$
则$\boldsymbol{A}$的不变因子就是$1, \cdots, 1, d_1(\lambda), \cdots, d_k(\lambda)$.

再由λ矩阵相抵下行列式因子不改变所以两边取n阶行列式因子
故 $|\lambda I - A| = d_1(\lambda) \cdots d_k(\lambda)$ 即特征多项式为非常数不变因子乘积

7.7 有理标准型

笔记 利用矩阵的不变因子,现在可以来构造所谓的有理标准型了.
我们的想法是寻找一个比较简单的矩阵,使它与给定的矩阵有相同的不变因子.
由前面两节我们已经知道,矩阵$\boldsymbol{A}$的特征矩阵$\lambda \boldsymbol{I}-\boldsymbol{A}$的法式为$\operatorname{diag}\{1,\cdots,1,d_1(\lambda),\cdots,d_k(\lambda)\}$
其中$d_i(\lambda)$为首一非常数多项式且$d_i(\lambda)\mid d_{i+1}(\lambda)(i=1,2,\cdots,k-1)$,则$\boldsymbol{A}$的不变因子就是$1,\cdots,1,d_1(\lambda),\cdots,d_k(\lambda)$.

Lemma 7.12

设r阶矩阵$\boldsymbol{F}=\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_r & -a_{r-1} & -a_{r-2} & \cdots & -a_1 \end{pmatrix}$

(1)$\boldsymbol{F}$的行列式因子为$\underbrace{1,\cdots,1}_{r-1\text{个}},f(\lambda)\quad f(\lambda)=\lambda^r+a_1\lambda^{r-1}+\cdots+a_r$,

(2) $\boldsymbol{F}$的不变因子组:$\underbrace{1,\cdots,1}_{r-1\text{个}},f(\lambda)\quad f(\lambda)=\lambda^r+a_1\lambda^{r-1}+\cdots+a_r$,式给出;

(3)$\boldsymbol{F}$的极小多项式,特征多项式等于$f(\lambda)$.

(4) F的法式即为$\operatorname{diag}(1,1\cdots f(\lambda))$

并且记$F=F(f(\lambda))$表示由$f(\lambda)$生成的

Proof (1)$\boldsymbol{F}$的r阶行列式因子就是它的特征多项式,由行列式得特征多项式为$|\lambda \boldsymbol{I}-\boldsymbol{F}|=\lambda^r+a_1\lambda^{r-1}+\cdots+a_r$.
对任一$k<r,\lambda \boldsymbol{I}-\boldsymbol{F}$总有一个$k$阶子式其值等于$(-1)^k$,故$D_k(\lambda)=1$.
(2)
设$\boldsymbol{F}$的极小多项式为$m(\lambda)$只要证$m(\lambda)=f(\lambda)$
由$Hamilton-Cayley$定理知道$m(\lambda)\,|f(\lambda)\Rightarrow \deg m(\lambda)\leqslant r$
若$\deg m(\lambda)=r+$首一性,显然得到
若$\deg m(\lambda)<r$那么设$m(\lambda)=c_{r-1}\lambda^{r-1}+\cdots+c_1\lambda+c_0$ (其中$c_{r-1}\cdots c_1,c_0$不全为0)
设$\boldsymbol{e}_i(i=1,2,\cdots,r)$是$r$维标准单位行向量,则不难算出:$e_1\boldsymbol{F}=e_2,e_1\boldsymbol{F}^2=e_3,\cdots,e_1\boldsymbol{F}^i=e_{i+1}.(\forall 1\leqslant i\leqslant r-1)$
$0=m(F)=c_{r-1}F^{r-1}+\cdots+c_1F+c_0I_r\Rightarrow e_10=0=e_1\left(c_{r-1}F^{r-1}+\cdots+c_1F+c_0I_r\right)$
$\Rightarrow 0=c_{r-1}e_r+c_{r-2}e_{r-1}+\cdots+c_1e_2+c_0e_1=(c_0,c_1\cdots c_{r-1})\Rightarrow c_0=c_1=\cdots=c_{r-1}=0$矛盾
证毕

Lemma 7.13

设$\lambda-$矩阵$\boldsymbol{A}(\lambda)$相抵于对角$\lambda-$矩阵$\operatorname{diag}\{d_1(\lambda),d_2(\lambda),\cdots,d_n(\lambda)\}$
设$\lambda-$矩阵$\boldsymbol{B}(\lambda)$相抵于对角$\lambda-$矩阵$\operatorname{diag}\{d_1'(\lambda),d_2'(\lambda),\cdots,d_n'(\lambda)\}$
且$d_1'(\lambda),d_2'(\lambda),\cdots,d_n'(\lambda)$是$d_1(\lambda),d_2(\lambda),\cdots,d_n(\lambda)$的一个置换(即若不计次序,这两组多项式完全相同)
则$\boldsymbol{A}(\lambda)$相抵于$\boldsymbol{B}(\lambda)$.

Proof 利用行列对换很轻易的知道如果知识置换两个非常轻松，那么多个置换只需利用两两即可

Theorem 7.7.1 (有理标准型)

设$\boldsymbol{A}$是数域$\mathbb{K}$上的n阶方阵,$\boldsymbol{A}$的不变因子组为$\underbrace{1,\cdots,1}_{n-k个},d_1(\lambda),\cdots,d_k(\lambda)$,其中 $\deg d_i(\lambda)=m_i$

则$\boldsymbol{A}$相似于下列分块对角阵:$\boldsymbol{F}=\begin{pmatrix}\boldsymbol{F}_1&&&\\&\boldsymbol{F}_2&&\\&&\ddots&\\&&&\boldsymbol{F}_k\end{pmatrix}$

其中$\boldsymbol{F}_i$的阶等于m_i,且$\boldsymbol{F}_i$是形如引理中的矩阵,$\boldsymbol{F}_i$的最后一行由$d_i(\lambda)$系数(除最高次项)的负值组成.

称矩阵$\boldsymbol{A}$的有理标准型或$Frobenius$(弗罗本纽斯)标准型,每个$\boldsymbol{F}_i$称为$Frobenius$块.

Proof

注意到$\lambda\boldsymbol{I}-\boldsymbol{A}$的第$n$个行列式因子就是$\boldsymbol{A}$的特征多项式$|\lambda\boldsymbol{I}-\boldsymbol{A}|$,再由行列式因子的相抵不变性可知:$|\lambda\boldsymbol{I}-\boldsymbol{A}|=d_1(\lambda)d_2(\lambda)\cdots d_k(\lambda)$.

而$|\lambda\boldsymbol{I}-\boldsymbol{A}|$是一个$n$次多项式,因此$m_1+m_2+\cdots+m_k=n$.其中$m_i$为$d_i(\lambda)$多项式最高次数

由引理矩阵$\boldsymbol{F}_i$的不变因子为$\underbrace{1,\cdots,1}_{m_i-1个},d_i(\lambda)$,其中共有$m_i-1$个1.即$\lambda\boldsymbol{I}-\boldsymbol{F}\left(d_{i(\lambda)}\right)$相抵于$\mathrm{diag}_i\left\{\underbrace{1,1\cdots}_{m_i-1个}d_i(\lambda)\right\}$

$$|\lambda\boldsymbol{I}-\boldsymbol{F}|=\begin{pmatrix}\lambda\boldsymbol{I}-\boldsymbol{F}(d_1(\lambda))&&&\\&\lambda\boldsymbol{I}-\boldsymbol{F}(d_2(\lambda))&&\\&&\ddots&\\&&&\lambda\boldsymbol{I}-\boldsymbol{F}(d_k(\lambda))\end{pmatrix}$$

通过行列变换即可得到$\begin{pmatrix}\mathrm{diag}_1&&&\\&\mathrm{diag}_2&&\\&&\ddots&\\&&&\mathrm{diag}_k\end{pmatrix}$

再通过行列对换可以得到$\mathrm{diag}\{1,1\cdots1,d_1(\lambda),d_2(\lambda)\cdots d_k(\lambda)\}$而且1的个数一共有$m_1+\cdots+m_k-k=n-k$个

所以$\lambda\boldsymbol{I}-\boldsymbol{A}$与$\lambda\boldsymbol{I}-\boldsymbol{F}$相抵故由定理可知$\boldsymbol{A}$与$\boldsymbol{F}$相似

笔记 (1) 由有理标准型 $\Rightarrow$ 不变因子

1.确定有理标准型阶数n,确定$Frobenius$块个数k

2.不变因子最前面有$n-k$个1

3.写出每个$Frobenius$块的多项式可见本节开头$lemma$

4.总共写出不变因子组

(2) 由不变因子 $\Rightarrow$ 有理标准型

1.仅依靠k个非常数不变因子$d_1(\lambda)\cdots d_k(\lambda)$通过每个$d_i(\lambda)$最高次数$n_i$做和即$n_1+\cdots+n_k=n$即为$F$矩阵的阶数

2.用$n-k$即1的个数这是蕴含的信息

3.每个块是n_i阶,写出每个$Frobenius$块可见开头$lemma$

Theorem 7.7.2

设数域$\mathbb{K}$上的n阶矩阵$\boldsymbol{A}$的不变因子为$1,\cdots,1,d_1(\lambda),\cdots,d_k(\lambda)$,其中$d_i(\lambda)\mid d_{i+1}(\lambda)(i=1,\cdots,k-1)$

则$\boldsymbol{A}$的极小多项式$m(\lambda)=d_k(\lambda)$.

Proof 设$\boldsymbol{A}$的有理标准型为$\boldsymbol{F}=\begin{pmatrix}\boldsymbol{F}_1 & & & \\ & \boldsymbol{F}_2 & & \\ & & \ddots & \\ & & & \boldsymbol{F}_k\end{pmatrix}$.因为相似矩阵有相同的极小多项式,故只需证明$\boldsymbol{F}$的极小多项式是$d_k(\lambda)$即可.
但$\boldsymbol{F}$是分块对角阵,由前文知$\boldsymbol{F}$的极小多项式是诸$\boldsymbol{F}_i$极小多项式的最小公倍式.又由引理知$\boldsymbol{F}_i$的极小多项式为$d_i(\lambda)$.
因为$d_i(\lambda)\mid d_{i+1}(\lambda)$,故诸$d_i(\lambda)$的最小公倍式等于$d_k(\lambda)$.

Example 7.3 下面两个4阶矩阵$\boldsymbol{A}=\begin{pmatrix}0&0&0&0\\0&0&0&0\\0&0&0&1\\0&0&0&0\end{pmatrix}$, $\boldsymbol{B}=\begin{pmatrix}0&1&0&0\\0&0&0&0\\0&0&0&1\\0&0&0&0\end{pmatrix}$的不变因子分别为$\boldsymbol{A}:1,\lambda,\lambda,\lambda^2$和$\boldsymbol{B}:1,1,\lambda^2,\lambda^2$.
它们的特征多项式和极小多项式分别相等,但它们不相似.

Theorem 7.7.3

$\boldsymbol{A}_{n\times n}$的不变因子组为$1,1\cdots d_1(\lambda)\cdots d_k(\lambda)$

1.法式为$\operatorname{diag}(1,1\cdots d_1(\lambda)\cdots d_k(\lambda))$ (法式和不变因子互相唯一确定)

2.$|\lambda I-A|$特征多项式为$d_1(\lambda)\times\cdots\times d_k(\lambda)$

3.A的极小多项式为$d_k(\lambda)$

4.$deg(d_1(\lambda))+\cdots+deg(d_k(\lambda))=n$

7.8 初等因子

由前面两节我们已经知道,矩阵$\boldsymbol{A}$的特征矩阵$\lambda\boldsymbol{I}-\boldsymbol{A}$的法式为$\mathrm{diag}\{1,\cdots,1,d_1(\lambda),\cdots,d_k(\lambda)\}$
其中$d_i(\lambda)$为首一非常数多项式且$d_i(\lambda)\mid d_{i+1}(\lambda)(i=1,2,\cdots,k-1)$,则$\boldsymbol{A}$的不变因子就是$1,\cdots,1,d_1(\lambda),\cdots,d_k(\lambda)$.
利用矩阵的不变因子,我们可以求一个矩阵的有理标准型.有理标准型对任何数域$\mathbb{K}$都可以求出来,它有着诸多的用途.
但是有理标准型也有一些缺点,主要是它有时不够简单,即有时每个$Frobenius$块太大,用起来不太方便.
有理标准型中$Frobenius$块太大的原因是不变因子$d_i(\lambda)$的次数可能比较高.如果我们用因式分解的方法分解每个$d_i(\lambda)$
这就有可能造出更细的标准型来.为此,我们先引进初等因子的概念.

Definition 7.14 (准素因子)

设$f(\lambda)\in\mathbb{K}(\lambda)$, $p(\lambda)$为不可约多项式,若存在$e\in\mathbb{Z}^+$使得$p(\lambda)^e\mid f(\lambda)$但是$p(\lambda)^{e+1}\nmid f(\lambda)$
则称$p(\lambda)^e$为$f(\lambda)$的准素因子

Definition 7.15 (初等因子组)

考虑A的不变因子$d_1(\lambda)\cdots d_k(\lambda)$

$$\begin{cases} d_1(\lambda)=p_1(\lambda)^{e_{11}}\,p_2(\lambda)^{e_{12}}\cdots p_t(\lambda)^{e_{1t}} \\ d_2(\lambda)=p_1(\lambda)^{e_{21}}\,p_2(\lambda)^{e_{22}}\cdots p_t(\lambda)^{e_{2t}} \\ \cdots\cdots \\ d_k(\lambda)=p_1(\lambda)^{e_{k1}}\,p_2(\lambda)^{e_{k2}}\cdots p_t(\lambda)^{e_{kt}} \end{cases}$$

规定$e_{ij}\geqslant 0$目的是保持各个$d_i(\lambda)$的分解具有类似的形式
由于$d_i(\lambda)\,|d_{i+1}(\lambda)$因此$e_{1j}\leqslant e_{2j}\leqslant\cdots\leqslant e_{kj}\,(j=1,2,\cdots t)$
若上式中的某个$e_{ij}>0$我们则称$p_j(\lambda)^{e_{ij}}$为A的一个初等因子.全体初等因子称为初等因子组
亦即A的不变因子的准素因子全体

Theorem 7.8.1

数域$\mathbb{K}$上的两个矩阵A与$\boldsymbol{B}$相似$\Leftrightarrow$它们有相同的初等因子组
即矩阵的初等因子组是矩阵相似关系的全系不变量.

Proof 显然不变因子组决定了初等因子组,反之
给定A的初等因子组,涉及的不可约多项式设为$p_1(\lambda)\cdots p_t(\lambda)$
由于每个不可约多项式的可能出现的个数不一样(因为次数可能不一样).但是总可以找到一个最大的
其余的利用零次幂补充不妨假设补充完之后针对每个不可约多项式按照降幂排列为如下

$$\begin{array}{l} p_1(\lambda)^{e_{k1}},p_1(\lambda)^{e_{k-1,1}},\cdots,p_1(\lambda)^{e_{11}} \\ p_2(\lambda)^{e_{k2}},p_2(\lambda)^{e_{k-1,2}},\cdots,p_2(\lambda)^{e_{12}} \\ \qquad\cdots\cdots\cdots\cdots \\ p_t(\lambda)^{e_{kt}},p_t(\lambda)^{e_{k-1,t}},\cdots,p_t(\lambda)^{e_{1t}} \end{array}$$

由于初等因子定义当中的整除性与次数逐渐升高的性质
那么我们令

$$
\begin{aligned}
d_k(\lambda) =& p_1(\lambda)^{e_{k1}} p_2(\lambda)^{e_{k2}} \cdots p_t(\lambda)^{e_{kt}} \\
d_{k-1}(\lambda) =& p_1(\lambda)^{e_{k-1,1}} p_2(\lambda)^{e_{k-1,2}} \cdots p_t(\lambda)^{e_{k-1,t}} \\
& \cdots\cdots\cdots\cdots \\
d_1(\lambda) =& p_1(\lambda)^{e_{11}} p_2(\lambda)^{e_{12}} \cdots p_t(\lambda)^{e_{1t}}
\end{aligned}
$$

简单可知初等因子组与不变因子组之间是一一对应的

Example 7.4 设9阶矩阵 (这个信息由前文可以自行推理得到) $\boldsymbol{A}$的不变因子组为$1, \cdots, 1, (\lambda-1)\left(\lambda^2+1\right), (\lambda-1)^2\left(\lambda^2+1\right)\left(\lambda^2-2\right)$
试分别在有理数域、实数域和复数域上求$\boldsymbol{A}$的初等因子组.

$\boldsymbol{A}$在有理数域上的初等因子组为$\lambda-1, (\lambda-1)^2, \lambda^2+1, \lambda^2+1, \lambda^2-2$.
$\boldsymbol{A}$在实数域上的初等因子组为$\lambda-1, (\lambda-1)^2, \lambda^2+1, \lambda^2+1, \lambda+\sqrt{2}, \lambda-\sqrt{2}$.
$\boldsymbol{A}$在复数域上的初等因子组为$\lambda-1, (\lambda-1)^2, \lambda+\mathrm{i}, \lambda+\mathrm{i}, \lambda-\mathrm{i}, \lambda-\mathrm{i}, \lambda+\sqrt{2}, \lambda-\sqrt{2}$.

Example 7.5 设$\boldsymbol{A}$是一个矩阵, 它的初等因子组为$\lambda-1, \lambda-1, (\lambda-1)^2, (\lambda+1)^2, (\lambda+1)^3, \lambda-2$.
求$\boldsymbol{A}$的不变因子组.
由定义可知初等因子是不变因子中非常数的准素因子
故仍然可知矩阵的阶数为所有多项式最高次数之和为10阶
将上述多项式分类按降幂排列：

$$
\begin{array}{lll}
(\lambda-1)^2 & \lambda-1 & \lambda-1 \\
(\lambda+1)^3 & (\lambda+1)^2 & \color{red}{1} \\
\lambda-2 & \color{red}{1} & \color{red}{1}
\end{array}
$$

红色代表补充的0次幂

于是$\begin{cases} d_3(\lambda)=(\lambda-1)^2(\lambda+1)^3(\lambda-2) \\ d_2(\lambda)=(\lambda-1)(\lambda+1)^2 \\ d_1(\lambda)=\lambda-1. \end{cases}$

从而1的个数也知道了为$10-3=7$个
从而$\boldsymbol{A}$的不变因子组为$1, \cdots, 1, \lambda-1, (\lambda-1)(\lambda+1)^2, (\lambda-1)^2(\lambda+1)^3(\lambda-2)$, 其中有7个1.

7.9 Jordan 标准型

借用有理标准型的思想, 我们希望对于某个初等因子 $d_k(\lambda)^{e_k}$ 找到较为简单的矩阵其初等因子就只有一个就为 $d_k(\lambda)^{e_k}$。

Lemma 7.14

r阶矩阵$\boldsymbol{J}_r(\lambda_0)=\begin{pmatrix}\lambda_0 & 1 & & & \\ & \lambda_0 & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda_0\end{pmatrix}_{r\times r}$ 考虑其$\lambda\boldsymbol{I}-\boldsymbol{J}_r(\lambda_0)$

1.行列式因子组为$\underbrace{1,1\cdots1}_{r-1\text{个}},\quad(\lambda-\lambda_0)^r$.

2.不变因子组为$\underbrace{1,1\cdots1}_{r-1\text{个}},\quad(\lambda-\lambda_0)^r$.

3.初等因子组为 $(\lambda-\lambda_0)^r$.

4.即$\lambda\boldsymbol{I}-\boldsymbol{J}_r(\lambda_0)$ 相抵于$\mathrm{diag}\{1,1,\cdots(\lambda-\lambda_0)^r\}$

Proof 显然$\boldsymbol{J}$的特征多项式为 $(\lambda-\lambda_0)^r$.对任一小于r的正整数k, $\lambda\boldsymbol{I}-\boldsymbol{J}$总有一个$k$阶子式, 其值等于 $(-1)^k$
因此$\boldsymbol{J}$的行列式因子为$1,\cdots,1,(\lambda-\lambda_0)^r$ 式也是$\boldsymbol{J}$的不变因子组, 故$\boldsymbol{J}$的初等因子组只有一个多项式 $(\lambda-\lambda_0)^r$.

我们知道如果去求一个$\lambda I-A$的初等因子往往需要求出其法式得到不变因子然后做分解进而得到初等因子
我们现在退一步若$\lambda I-A$能够相抵于一个对角矩阵也可以较为简单的求出初等因子

Lemma 7.15

设特征矩阵$\lambda\boldsymbol{I}-\boldsymbol{A}$经过初等变换化为下列对角阵 : $\begin{pmatrix}f_1(\lambda) & & & \\ & f_2(\lambda) & & \\ & & \ddots & \\ & & & f_n(\lambda)\end{pmatrix}$

其中$f_i(\lambda)(i=1,\cdots,n)$为非零首一多项式.
则矩阵$\boldsymbol{A}$的初等因子组等于所有$f_i(\lambda)$的准素因子的集合.

Proof 1.若$f_i(\lambda),f_j(\lambda)(i\neq j)$的最大公因式和最小公倍式分别为$g(\lambda),h(\lambda)$
则$\mathrm{diag}\left\{f_1(\lambda),\cdots,f_i(\lambda),\cdots,f_j(\lambda),\cdots,f_n(\lambda)\right\}$ 经过初等变换可以变为$\mathrm{diag}\{f_1(\lambda),\cdots,g(\lambda),\cdots,h(\lambda),\cdots,f_n(\lambda)\}$
且这两个对角阵具有相同的准素因子组.
不失一般性, 令$i=1,j=2$.因为 $(f_1(\lambda),f_2(\lambda))=g(\lambda)$, 所以存在$u(\lambda),v(\lambda)$, 使$f_1(\lambda)u(\lambda)+f_2(\lambda)v(\lambda)=g(\lambda)$.
又令$f_1(\lambda)=g(\lambda)q(\lambda)$, 则$h(\lambda)=f_2(\lambda)q(\lambda)$.对式作下列初等变换 :

$$\begin{pmatrix}f_1(\lambda) & & & \\ & f_2(\lambda) & & \\ & & \ddots & \\ & & & f_n(\lambda)\end{pmatrix}\xrightarrow{\text{第一行}\times u(\lambda)+\text{第二行}}\begin{pmatrix}f_1(\lambda) & 0 & & \\ f_1(\lambda)u(\lambda) & f_2(\lambda) & & \\ & & \ddots & \\ & & & f_n(\lambda)\end{pmatrix}\xrightarrow{\text{第二列}\times(v\lambda)+\text{第一列}}$$

$$\begin{pmatrix}f_1(\lambda) & & & \\ g(\lambda) & f_2(\lambda) & & \\ & & \ddots & \\ & & & f_n(\lambda)\end{pmatrix}\xrightarrow{\text{第二列行}\times(-q(\lambda))+\text{第一行}}\begin{pmatrix}0 & -h(\lambda) & & \\ g(\lambda) & f_2(\lambda) & & \\ & & \ddots & \\ & & & f_n(\lambda)\end{pmatrix}\xrightarrow{\text{通过}g(\lambda)\text{消去}h(\lambda)}\begin{pmatrix}g(\lambda) & & & \\ & h(\lambda) & & \\ & & \ddots & \\ & & & f_n(\lambda)\end{pmatrix}.$$

现来考察$g(\lambda)$与$h(\lambda)$的准素因子.将$f_1(\lambda),f_2(\lambda)$作标准因式分解, 其分解式不妨设为

$f_1(\lambda)=(\lambda-\lambda_1)^{c_1}(\lambda-\lambda_2)^{c_2}\cdots(\lambda-\lambda_t)^{c_t}$
$f_2(\lambda)=(\lambda-\lambda_1)^{d_1}(\lambda-\lambda_2)^{d_2}\cdots(\lambda-\lambda_t)^{d_t}$
其中c_i,d_i为非负整数$(\geqslant 0)$.令$e_i=\max\{c_i,d_i\}$, $k_i=\min\{c_i,d_i\}$,则
$g(\lambda)=(\lambda-\lambda_1)^{k_1}(\lambda-\lambda_2)^{k_2}\cdots(\lambda-\lambda_t)^{k_t}$,
$h(\lambda)=(\lambda-\lambda_1)^{e_1}(\lambda-\lambda_2)^{e_2}\cdots(\lambda-\lambda_t)^{e_t}$.
不难看出$g(\lambda),h(\lambda)$的准素因子组与$f_1(\lambda),f_2(\lambda)$的准素因子组相同.

2.证明原式所示矩阵的法式可通过上述变换得到.先将第$(1,1)$位置的元素依次和第$(2,2)$位置,$\cdots$,第(n,n)位置的元素进行上述变换
此时第$(1,1)$元素的所有一次因式的幂都是最小的;
再将第$(2,2)$位置的元素依次和第$(3,3)$位置,$\cdots$,第(n,n)位置的元素进行上述变换;
最后将第$(n-1,n-1)$位置的元素和第(n,n)位置的元素进行上述变换.

那么最后我们可以得到原来$\mathrm{diag}\{f_1(\lambda)\cdots f_n(\lambda)\}$相抵于$\mathrm{diag}\{d_1(\lambda)\cdots d_n(\lambda)\}$
且$d_1(\lambda)\,|d_2(\lambda)|\cdots|d_n(\lambda)$即法式.注意到在每一次变换的过程中,准素因子组都保持不变
故$\mathrm{diag}\{f_1(\lambda)\cdots f_n(\lambda)\}$与法式$\mathrm{diag}\{d_1(\lambda)\cdots d_n(\lambda)\}$有相同准素因子组且法式$\mathrm{diag}\{d_1(\lambda)\cdots d_n(\lambda)\}$的初等因子组即其准素因子组
这就证明了结论.

Example 7.6 设$\lambda I-\boldsymbol{A}$经过初等变换后化为下列对角阵:$\begin{pmatrix}1&&&&\\&(\lambda-1)^2(\lambda+2)&&&\\&&(\lambda+2)&&\\&&&1&\\&&&&(\lambda-1)\end{pmatrix}$,求$\boldsymbol{A}$的初等因子组.

解$\boldsymbol{A}$的初等因子组为$\lambda-1,(\lambda-1)^2,\lambda+2,\lambda+2$.

Lemma 7.16

设J是分块对角阵:$\begin{pmatrix}\boldsymbol{J}_1&&&\\&\boldsymbol{J}_2&&\\&&\ddots&\\&&&\boldsymbol{J}_k\end{pmatrix}$,其中每个$J_i$都是形如引理中的矩阵,$\boldsymbol{J}_i$的初等因子组为$(\lambda-\lambda_i)^{r_i}$
则J的初等因子组为$(\lambda-\lambda_1)^{r_1},(\lambda-\lambda_2)^{r_2},\cdots,(\lambda-\lambda_k)^{r_k}$.

Proof $\lambda\boldsymbol{I}-\boldsymbol{A}=\begin{pmatrix}\lambda\boldsymbol{I}-\boldsymbol{J}_{r_1}(\lambda_1)&&&\\&\lambda\boldsymbol{I}-\boldsymbol{J}_{r_2}(\lambda_2)&&\\&&\ddots&\\&&&\lambda\boldsymbol{I}-\boldsymbol{J}_{r_k}(\lambda_k)\end{pmatrix}$
由引理1可知$\lambda\boldsymbol{I}-\boldsymbol{J}_{r_1}(\lambda_1)$相抵于$\mathrm{diag}\{1,1\cdots(\lambda-\lambda_1)^{r_1}\}$诸如此类
$\lambda\boldsymbol{I}-\boldsymbol{A}$是一个分块对角$\lambda-$矩阵.由于对分块对角阵中某一块施行初等变换时其余各块保持不变
因此$\lambda\boldsymbol{I}-\boldsymbol{J}$相抵于下列分块对角阵:$\boldsymbol{H}=\begin{pmatrix}\boldsymbol{H}_1&&&\\&\boldsymbol{H}_2&&\\&&\ddots&\\&&&\boldsymbol{H}_k\end{pmatrix}$,其中$\boldsymbol{H}_i=\mathrm{diag}\{1,\cdots,1,(\lambda-\lambda_i)^{r_i}\}$.
再由引理2只需要拿到$\boldsymbol{H}$的准素因子组即可得结论.

Theorem 7.9.1 (Jordan 标准型)

1.设$A \in M_n(\mathbb{C})$.的初等因子组$(\lambda-\lambda_1)^{r_1}(\lambda-\lambda_2)^{r_2}\cdots(\lambda-\lambda_k)^{r_k}$
则A相似于$\boldsymbol{J}=\operatorname{diag}\left\{\boldsymbol{J}_{r_1}(\lambda_1),\boldsymbol{J}_{r_2}(\lambda_2)\cdots\boldsymbol{J}_{r_k}(\lambda_k)\right\}$,
A称为$Jordan$标准型,每一个$\boldsymbol{J}_{r_i}(\lambda_i)$称为$Jordan$块

2.特别的A的特征多项式为不变因子乘积即为初等因子乘积即$(\lambda-\lambda_1)^{r_1}(\lambda-\lambda_2)^{r_2}\cdots(\lambda-\lambda_k)^{r_k}$
那么$\lambda_1\sim\lambda_k$为特征值

注 1.$Jordan$标准型中,$Jordan$块换了次序不改变初等因子组因而从相似角度不变还是相似标准型
2.一个$Jordan$标准块由一个初等因子唯一决定
从而在不考虑块顺序下,$Jordan$标准型由A的初等因子唯一决定即由A唯一决定
3.设$\varphi\in\mathscr{L}\left(V_{\mathbb{C}}^n\right)$是复数域上线性空间$V$上的线性变换,则必存在$V$的一组基,使得$\varphi$在这组基下的表示矩阵为式所示的$Jordan$标准型.

Theorem 7.9.2 (矩阵可对角化定理 4,5)

代数版本
设$\boldsymbol{A}$是n阶复矩阵,则下列结论等价:
(1)$\boldsymbol{A}$可对角化;
(2)$\boldsymbol{A}$的极小多项式无重根;
(3)$\boldsymbol{A}$的初等因子都是一次多项式.(等价$Jordan$块都是一阶的)

几何版本
设φ是复线性空间V上的线性变换,则φ可对角化当且仅当φ的极小多项式无重根,当且仅当φ的初等因子都是一次多项式.

Proof (1) $\Rightarrow$ (2):由定理6.3.2的结论即得.
(2) $\Rightarrow$ (3):设$\boldsymbol{A}$的极小多项式$m(\lambda)$无重根.
由于$m(\lambda)$是$\boldsymbol{A}$的最后一个不变因子,故$\boldsymbol{A}$的所有不变因子都无重根(由于递推整除性)从而$\boldsymbol{A}$的初等因子都是一次多项式.
(3) $\Rightarrow$ (1):设$\boldsymbol{A}$的初等因子组为$\lambda-\lambda_1,\lambda-\lambda_2,\cdots,\lambda-\lambda_n$(初等因子组相乘即为特征多项式肯定有$n$个根)
再由定理可得,$\boldsymbol{A}$相似于对角阵$\operatorname{diag}\{\lambda_1,\lambda_2,\cdots,\lambda_n\}$,即$\boldsymbol{A}$可对角化.(这里$\lambda_1\cdots\lambda_n$可能相同)

Corollary 7.8 (线性变换可对角化则限制也可对角化)

1.设φ是复线性空间V上的线性变换,V_0是φ的不变子空间.如果φ可对角化,则φ在V_0上的限制也可对角化.
2.若$M=\begin{pmatrix} A^m & C \\ O & B^n \end{pmatrix}$可对角化,则$A,B$均可对角化

Proof 因为φ的极小多项式$g(\lambda)$无重根,而φ在V_0上限制的极小多项式是$g(\lambda)$的因子,所以也没有重根.则φ在V_0上的限制也可对角化.

设$M(\lambda)$的极小多项式为$m(\lambda)$故$O=m(M)=\begin{pmatrix} m(A) & * \\ O & m(B) \end{pmatrix}$(由上三角矩阵经过多项式计算得到仍然为上三角矩阵)
那么$m(A),m(B)=O$那么$m_A(\lambda)$作为A的极小多项式就有$m_A(\lambda)\,|m(\lambda)$还有$m_B(\lambda)\,|m(\lambda)$
由定理可以知道M可对角化得到$m(\lambda)$无重根所以$m_A(\lambda),m_B(\lambda)$也无重根
再由定理可知可对角化

Corollary 7.9

设φ是复线性空间V上的线性变换,且$V = V_1 \oplus V_2 \oplus \cdots \oplus V_k$,其中每个$V_i$都是$\varphi$的不变子空间
则φ可对角化的 $\Leftrightarrow$ φ在每个V_i上的限制都可对角化.

$$A = \begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_k \end{pmatrix}$$若A可对角化那么A_i也可对角化

Proof 必要性由上个即得,下证充分性.
若φ在每个V_i上的限制都可对角化,由定义存在V_i的一组基,使得 $\varphi|_{V_i}$ 在这组基下的表示矩阵是对角阵.
再由定理知所有V_i的一组基可以拼成V的一组基,因此φ在这组基下的表示阵是对角阵,即φ可对角化.(由对角化定义)

这是代数版本的叙述显而易见或者利用最小多项式是分块矩阵最小多项式的最小公倍式$m_A(\lambda) = \left[m_{A_1}(\lambda) \cdots m_{A_k}(\lambda)\right]$
利用没有重根也可

Corollary 7.10

设A是数域$\mathbb{K}$上的矩阵,如果$\boldsymbol{A}$的特征值全在$\mathbb{K}$中,则$\boldsymbol{A}$在$\mathbb{K}$上相似于其$Jordan$标准型.

Proof 由于$\boldsymbol{A}$的特征值全在$\mathbb{K}$中,故$\boldsymbol{A}$的$Jordan$标准型$\boldsymbol{J}$实际上是$\mathbb{K}$上的矩阵.
因为$\boldsymbol{A}$在复数域上相似于$\boldsymbol{J}$,由相似关系在基域下不改变知$\boldsymbol{A}$在$\mathbb{K}$上也相似于$\boldsymbol{J}$.

Example 7.7 设复数域上的四维线性空间V上的线性变换φ在一组基$\{e_1, e_2, e_3, e_4\}$下的表示矩阵为$\boldsymbol{A} = \begin{pmatrix} 3 & 1 & 0 & 0 \\ -4 & -1 & 0 & 0 \\ 6 & 1 & 2 & 1 \\ -14 & -5 & -1 & 0 \end{pmatrix}$
求V的一组基,使φ在这组基下的表示矩阵为$Jordan$标准型,并求出从原来的基到新基的过渡矩阵。

由题干可得存在P使得$P^{-1}AP = J$
先确定J,就要确定初等因子就要知道不变因子就要知道法式
用初等变换把$\lambda \boldsymbol{I} - \boldsymbol{A}$化为对角$\lambda-$矩阵并求出它的初等因子组为$(\lambda-1)^2, (\lambda-1)^2$.
因此,$\boldsymbol{A}$的$Jordan$标准型为$\boldsymbol{J} = \begin{pmatrix} 1 & 1 & & \\ 0 & 1 & & \\ & & 1 & 1 \\ & & 0 & 1 \end{pmatrix}$.
设矩阵$\boldsymbol{P}$是从$\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3, \boldsymbol{e}_4\}$到新基的过渡矩阵,则$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{J}$,此即$\boldsymbol{A}\boldsymbol{P} = \boldsymbol{P}\boldsymbol{J}$.
设$\boldsymbol{P} = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3, \boldsymbol{\alpha}_4)$,其中$\boldsymbol{\alpha}_i$是四维列向量,代入式得$(\boldsymbol{A}\boldsymbol{\alpha}_1, \boldsymbol{A}\boldsymbol{\alpha}_2, \boldsymbol{A}\boldsymbol{\alpha}_3, \boldsymbol{A}\boldsymbol{\alpha}_4) = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3, \boldsymbol{\alpha}_4)\begin{pmatrix} 1 & 1 & & \\ 0 & 1 & & \\ & & 1 & 1 \\ & & 0 & 1 \end{pmatrix}$

化成方程组为
$$\begin{aligned} \boldsymbol{A}\boldsymbol{\alpha}_1 &= \boldsymbol{\alpha}_1, \\ \boldsymbol{A}\boldsymbol{\alpha}_2 &= \boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_2, \\ \boldsymbol{A}\boldsymbol{\alpha}_3 &= \boldsymbol{\alpha}_3, \\ \boldsymbol{A}\boldsymbol{\alpha}_4 &= \boldsymbol{\alpha}_3 + \boldsymbol{\alpha}_4. \end{aligned}$$
即α_1与α_3是关于A矩阵特征值1的两个线性无关的特征向量 (因为P可逆肯定无关)

求出α_1, α_3回代求解 $(A - E)\,\alpha_2 = \alpha_1$方程组即可

7.10 Jordan 标准型的运用

设V是n维复线性空间,φ是V上的线性变换.设φ的初等因子组为$(\lambda-\lambda_1)^{r_1},(\lambda-\lambda_2)^{r_2},\cdots,(\lambda-\lambda_k)^{r_k}$

定理告诉我们,存在V的一组基$\{e_{11},e_{12},\cdots,e_{1r_1};e_{21},e_{22},\cdots,e_{2r_2};\cdots;e_{k1},e_{k2},\cdots,e_{kr_k}\}$

使得φ在这组基下的表示矩阵为$\boldsymbol{J}=\begin{pmatrix}\boldsymbol{J}_1 & & & \\ & \boldsymbol{J}_2 & & \\ & & \ddots & \\ & & & \boldsymbol{J}_k\end{pmatrix}$.

$\forall 1\leqslant i\leqslant k$考虑某个$\boldsymbol{J}_{r_i}(\lambda_i)$我们有$\left[\varphi(e_{i1}),\varphi(e_{i2})\cdots\cdots\varphi(e_{ir_i})\right]=(e_{i1},e_{i2}\cdots e_{ir_i})\begin{bmatrix}\lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1\\ & & & \lambda_i\end{bmatrix}$

$$\Rightarrow\begin{cases}\varphi(e_{i1})=\lambda_i e_{i1}\\ \varphi(e_{i2})=e_{i1}+\lambda_i e_{i2}\\ \cdots\cdots\cdots\\ \varphi(e_{ir_i})=\boldsymbol{e}_{ir_i-1}+\lambda_i e_{ir_i}\end{cases}$$

令$V_i=L(e_{i1},e_{i2}\cdots e_{ir_i})$ [显然$(e_{i1},e_{i2}\cdots e_{ir_i})$为$V_i$的一组基]

1.显然$\varphi(V_i)\subseteq V_i\Rightarrow V_i$为$\varphi-$不变子空间

2.显然V_i的基可以拼成V的基,故$V=V_1\oplus V_2\oplus\cdots\oplus V_k$

下面研究各个特征值的几何与代数重数

显然φ的特征多项式$f(\lambda)=(\lambda-\lambda_1)^{r_1}(\lambda-\lambda_2)^{r_2}\cdots(\lambda-\lambda_k)^{r_k}\ (r_1+\cdots+r_s=n)$

不妨设前面$\lambda_1=\lambda_2=\cdots\lambda_s.\lambda_1\neq\lambda_j\ (\forall s<j\leqslant k)$假设前面$s$个特征值都是$\lambda_1$

所以$f(\lambda)=(\lambda-\lambda_1)^{r_1+\cdots r_s}(\lambda-\lambda_{s+1})^{r_{s+1}}\cdots(\lambda-\lambda_k)^{r_k}$

$\Rightarrow\lambda_1$的代数重数$=r_1+r_2+\cdots+r_s=$属于λ_1 $Jordan$块的阶数之和

λ_1的几何重数$=\dim(V_{\lambda_1})=\dim\mathrm{Ker}(\varphi-\lambda_1 I_V)=n-\dim\mathrm{Im}(\varphi-\lambda_1 I_V)=n-rank(\varphi-\lambda_1 I_V)$

下面来计算$\varphi-\lambda_1 I_V$的秩即可根据其表示矩阵秩即可我们选取$Jordan$标准型来计算

而$Jordan$标准型为:$\{J_{r_1}(\lambda_1),J_{r_2}(\lambda_2)\cdots J_{r_k}(\lambda_k)\}$那么只要计算每个分块的秩即可

我们对$J_{r_i}(\lambda_i)-\lambda_1 I_{r_i}=\begin{bmatrix}\lambda_i-\lambda_1 & 1 & & \\ & \lambda_i-\lambda_1 & \ddots & \\ & & \ddots & 1\\ & & & \lambda_i-\lambda_1\end{bmatrix}$研究$\Rightarrow rank\left(J_{r_i}(\lambda_i)-\lambda_1 I_{r_i}\right)=\begin{cases}r_i-1 & \lambda_i=\lambda_1\text{即}1\leqslant i\leqslant s\\ r_i & \lambda_i\neq\lambda_1\text{即}s<i\leqslant k\end{cases}$

那么$rank(\varphi-\lambda_1 I_V)=rank(J-\lambda_1 I_n)=\sum rank\left(J_{r_i}(\lambda_i)-\lambda_1 I_{r_i}\right)=(r_1-1)+\cdots+(r_s-1)+r_{s+1}+\cdots+r_k=n-s$

$\Rightarrow\lambda_1$的几何重数$=s=$属于特征值$Jordan$块的个数

Theorem 7.10.1

线性变换φ的特征值λ_1的几何重数等于φ的$Jordan$标准型中属于特征值λ_1的$Jordan$块的个数

λ_1的代数重数等于所有属于特征值λ_1的$Jordan$块的阶数之和.

进一步,我们来考虑某个λ_i的特征子空间(λ_1为例)的一组基.在上文中我们已经知道$\dim V_{\lambda_1}=$几何重数$=s$

由前文知$\varphi(e_{11})=\lambda_1 e_{11}\cdots\varphi(e_{k1})=\lambda_1 e_{k1}\cdots\varphi(e_{s1})=\lambda_1 e_{s1}$

且由题目可知基$\{e_{11}\cdots e_{1r_1};e_{21}\cdots e_{2r_2};\cdots\cdots e_{k1}\cdots e_{kr_k}\}$是全空间的基

故$e_{11}\cdots e_{k1}\cdots e_{s1}\ (1\leqslant k\leqslant s)$线性无关故这些向量就为$V_{\lambda_1}$ (λ_1特征子空间的一组基)

令$\psi=\varphi-\lambda_i I_V$那么由上文就有$\begin{cases}\psi\left(e_{i1}\right)=0\\ \psi\left(e_{i_2}\right)=e_{i1}\\ \cdots\cdots\\ \psi\left(e_{ir_i}\right)=e_{ir_i-1}\end{cases}$

即$e_{ir_i}\xrightarrow{\psi}e_{ir_i-1}\cdots\cdots\xrightarrow{\psi}e_{i1}\xrightarrow{\psi}0$

V_i是关于$\psi=\varphi-\lambda_i I_V$的循环子空间$(\forall 1\leqslant i\leqslant k)$

Definition 7.16

设$\psi\in\mathscr{L}(V)$;V_0是r维的$\psi-$不变子空间，若存在$0\neq\alpha\in V_0$.使得$\left\{\alpha,\psi(\alpha)\cdots\psi^{r-1}(\alpha)\right\}$为$V_0$的一组基

则称V_0为$\psi-$循环子空间(前提是不变子空间);α称为循环向量

Theorem 7.10.2

设$R(\lambda_1)=V_1\oplus V_2\oplus\cdots\oplus V_s$则$R(\lambda_1)=\operatorname{Ker}(\varphi-\lambda_1 I_V)^n=\{v\in V\ |(\varphi-\lambda_1 I_V)^n=0\}$

Proof 考虑$R(\lambda_1)\subseteq\operatorname{Ker}(\varphi-\lambda_1 I_V)^n$

任取$v\in R(\lambda_1)$;$v=v_1+v_2+\cdots+v_s$; $v_i\in V_i$

对于某个V_i基 : $e_{ir_i}\xrightarrow{\varphi-\lambda_i I_V}e_{ir_i-1}\rightarrow\cdots\xrightarrow{\varphi-\lambda_i I_V}e_{i1}\xrightarrow{\varphi-\lambda_i I_V}0$

那么对于V_i下的基最多经过其实r_i次即可变为0，那么V_i下的任一一个普通向量通过基表示最多经过r_i次也可变为0

那么对于v分解成了$v_1\sim v_s$后，最多其实经过$\max\{r_1\cdots r_s\}$就可变为0, 也可以是$r_1+\cdots+r_s$(即λ_1的代数重数)

自然经过n次更是0了， $\Rightarrow(\varphi-\lambda_1 I_V)^n(v)=0$

再证$\operatorname{Ker}(\varphi-\lambda_1 I_V)^n\subseteq R(\lambda_1)$

设$v\in V$在基下坐标向量为$x=\left(x_{11},x_{12}\cdots x_{1r_1};x_{21}\cdots\cdots\right)'$

对于$(J-\lambda_1 I_n)^n x=0$的解

而由前文知道$\left(J_{r_i}(\lambda_i)-\lambda_1 I_{r_1}\right)^n=\begin{pmatrix}\lambda_i-\lambda_1 & 1 & & \\ & \lambda_i-\lambda_1 & \ddots & \\ & & \ddots & 1\\ & & & \lambda_i-\lambda_1\end{pmatrix}^n=\begin{cases}O & 1\leqslant i\leqslant s\\ \text{非异} & s<i\leqslant k\end{cases}$

那么对于$(J-\lambda_1 I_n)^n x=0$的解自然变成了分块去求解

$\Rightarrow x_{i1}=x_{i2}=\cdots=x_{ir_i}=0\,(\forall s<i\leqslant k)$且$\left(x_{11},x_{12}\cdots x_{1r_1};\cdots\cdots x_{s1},x_{s2}\cdots x_{sr_s}\right)$为任意解

所以$\operatorname{Ker}(\varphi-\lambda_1 I_V)^n=V_1\oplus V_2\cdots\oplus V_s$

Definition 7.17

$R(\lambda_1)=\operatorname{Ker}(\varphi-\lambda_1 I)^n=V_1\oplus V_2\cdots\oplus V_s$(所有关于特征值$\lambda_1$的循环子空间的直和)

$R(\lambda_1)$称为关于特征值λ_1的根子空间

且n可以换为$\max\{r_1\cdots r_s\}$或者$r_1+\cdots r_s$

$\dim(R(\lambda_1))=\dim V_1+\cdots\dim V_s=r_1+\cdots+r_s=\lambda_0$的代数重数

Corollary 7.11

上述定义的$n=1$时即为特征子空间.即特征子空间$\subseteq$根子空间$\left(V_{\lambda_1}\subseteq R(\lambda_1)\right)$

Corollary 7.12

$\varphi \in \mathscr{L}(V_{\mathbb{C}})$ 可对角化 $\Longleftrightarrow$ 对任一特征值λ_0有$R(\lambda_0) = V_{\lambda_0}$ (即根子空间与特征子空间相等)

Proof $\forall \lambda_0$; $\underbrace{R(\lambda_0) = V_{\lambda_0} \Longleftrightarrow \dim(R(\lambda_0)) = \dim(V_{\lambda_0})}_{\text{因为显然有} V_{\lambda_0} \subseteq R(\lambda_0)} \Longleftrightarrow \lambda_0$的代数重数 $= \lambda_0$几何重数 $\Longleftrightarrow$ 可对角化

Theorem 7.10.3

设$\varphi \in \mathscr{L}\left(V_{\mathbb{C}}^n\right)$

(1)若φ的初等因子组为 $(\lambda-\lambda_1)^{r_1}, (\lambda-\lambda_2)^{r_2}, \cdots, (\lambda-\lambda_k)^{r_k}$

则V可分解为k个不变子空间的直和 : $V = V_1 \oplus V_2 \oplus \cdots \oplus V_k$, 其中$V_i$的维数等于$r_i$且是$\boldsymbol{\varphi} - \lambda_i \boldsymbol{I}$的循环子空间;

(2)若$\lambda_1, \cdots, \lambda_s$是$\boldsymbol{\varphi}$的全体不同特征值

则V可分解为s个不变子空间的直和 : $V = R(\lambda_1) \oplus R(\lambda_2) \oplus \cdots \oplus R(\lambda_s)$, 其中$R(\lambda_i)$ 是λ_i的根子空间

$R(\lambda_i)$ 的维数等于λ_i的重数, 且每个$R(\lambda_i)$ 又可分解为 (1) 式中若干个V_j的直和.

Proposition 7.4

证明 : 复数域上的方阵$\boldsymbol{A}$必可分解为两个对称阵的乘积.

Proof 设$\boldsymbol{P}$是非异阵且使$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{J}$为$\boldsymbol{A}$的$Jordan$标准型.于是$\boldsymbol{A} = \boldsymbol{P}\boldsymbol{J}\boldsymbol{P}^{-1} = \operatorname{diag}\left\{\boldsymbol{J}_{r_1}(\lambda_1); \boldsymbol{J}_{r_2}(\lambda_2) \cdots\cdots \boldsymbol{J}_{r_k}(\lambda_k)\right\}$

$$\boldsymbol{J}_{r_i}(\lambda_i) = \begin{bmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{bmatrix} = \begin{bmatrix} & & 1 & \lambda_i \\ & & \lambda_i & \\ 1 & & & \\ \lambda_i & & & \end{bmatrix} \begin{bmatrix} & & 1 \\ & 1 & \\ 1 & & \end{bmatrix} = S_i T_i$$

令$S = \operatorname{diag}\{S_1 \cdots S_k\}$ $\quad T = \operatorname{diag}\{T_1 \cdots T_k\} \Rightarrow J = ST$

$A = PJP^{-1} = PSTP^{-1} = (PSP')\left((P')^{-1} TP^{-1}\right)$

Proposition 7.5

$A \in M_n(\mathbb{C})$ 求出A^k

Proof 在非异阵P使得$P^{-1}AP = J = \operatorname{diag}\left\{J_{r_1}(\lambda_1) \cdots J_{r_k}(\lambda_k)\right\}$

那么$A^k = P\operatorname{diag}\left\{J_{r_1}(\lambda_1)^k \cdots J_{r_k}(\lambda_k)^k\right\} P^{-1}$

而对于某个$J_n(\lambda_0) = \begin{bmatrix} \lambda_0 & 1 & & \\ & \lambda_0 & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_0 \end{bmatrix} = \lambda_0 I_n + N \left(N = \begin{bmatrix} & 1 & & \\ & & \ddots & \\ & & & 1 \\ & & & \end{bmatrix} \right)$

而针对幂零阵我们已经做过计算研究了.

所以个$J_n(\lambda_0)^k = (\lambda_0 I_n + N)^k$ 不防先假设$k \geqslant N-1$利用二项式展开

$$J_n(\lambda_0)^k = \begin{bmatrix} \lambda_0^k & C_k^1 \lambda_0^{k-1} & \cdots & C_k^{n-1} \lambda_0^{k-n+1} \\ & \lambda_0^k & \ddots & \vdots \\ & & \ddots & C_k^1 \lambda_0^{k-1} \\ & & & \lambda_0^k \end{bmatrix}$$

Lemma 7.17

设$\boldsymbol{A},\boldsymbol{B}$是两个$n$阶可对角化复矩阵且$\boldsymbol{AB}=\boldsymbol{BA}$,则它们可同时对角化,即存在可逆阵$P$,使$P^{-1}AP$和$P^{-1}BP$都是对角阵.

Proof 我们现将引理换为几何语言：

$\varphi\psi\in\mathscr{L}\left(V_{\mathbb{C}}^{n}\right).\varphi,\psi$可对角化且$\varphi\psi=\psi\varphi$则$\varphi,\psi$可同时对角化即$\exists$一组公共基使得$\varphi,\psi$表示阵均为对角阵

对空间维数n进行归纳,$n=1$显然,假设维数$<n$时.下证维数$=n$情形

设φ的全体不同特征值为$\lambda_1\cdots\lambda_s$

若$s=1\Rightarrow\varphi=\lambda_1I_V,\psi$可对角化$\exists\{e_1\cdots e_n\}$使得$\psi$表示阵为对角阵

那我们就选取公共基为$\{e_1\cdots e_n\}$即可

若$s\geqslant 2$,设特征子空间为$V_1\cdots V_s$

φ可对角化$\Rightarrow V=V_1\oplus\cdots\oplus V_s\ (V_i<n)$

下面断言有结论$\varphi\psi=\psi\varphi\Rightarrow V_i$都是$\psi$不变子空间

任取$v_i\in V_i$,即$\varphi(v_i)=\lambda_iv_i$那么$\varphi(\psi(v_i))=\psi(\varphi(v_i))=\psi(\lambda_iv_i)=\lambda_i\psi(v_i)$

$\Rightarrow\psi(v_i)\in V_i\Rightarrow V_i$都是$\psi$不变子空间

做限制$\varphi|_{V_i};\psi|_{V_i}$

第一：φ,ψ可对角化，做限制也可对角化上节内容

第二：限制也可交换顺序

第三：维数比原来的小

有归纳假设$\exists V_i$的基使得$\varphi|_{V_i};\psi|_{V_i}$在该基下表示阵为对角阵

$V=V_1\oplus\cdots\oplus V_s\ (V_i<n)$故$V_i$基拼为全空间一组基故也是对角阵

Theorem 7.10.4 ($Jordan-Chevalley$分解)

设$\boldsymbol{A}$是n阶复矩阵,则$\boldsymbol{A}$可分解为$\boldsymbol{A}=\boldsymbol{B}+\boldsymbol{C}$,其中$\boldsymbol{B},\boldsymbol{C}$适合下面条件：

(1)$\boldsymbol{B}$是一个可对角化矩阵;

(2)$\boldsymbol{C}$是一个幂零阵;

(3)$\boldsymbol{BC}=\boldsymbol{CB}$;

(4)$\boldsymbol{B},\boldsymbol{C}$均可表示为$\boldsymbol{A}$的多项式.

不仅如此,上述满足条件(1)$\sim$(3)的分解是唯一的.

Proof 先对$\boldsymbol{A}$的$Jordan$标准型J证明结论.设$\boldsymbol{A}$的全体不同特征值为$\lambda_1,\lambda_2,\cdots,\lambda_s$且$\boldsymbol{J}=\begin{pmatrix}\boldsymbol{J}_1&&&\\&\boldsymbol{J}_2&&\\&&\ddots&\\&&&\boldsymbol{J}_s\end{pmatrix}$

其中J_i是属于特征值λ_i的根子空间对应的块,其阶设为$m_i=r_1+\cdots+r_k$ (即λ_i的代数重数)

$$\boldsymbol{J}_i=\begin{bmatrix}\boldsymbol{J}_{r_1}(\lambda_i)&&\\&\ddots&\\&&\boldsymbol{J}_{r_k}(\lambda_i)\end{bmatrix}_{m_i}=\lambda_i\boldsymbol{I}+\boldsymbol{N}_i\ (\text{诸多小幂零阵的拼起来})$$

显然对每个i均有$\boldsymbol{J}_i=\boldsymbol{M}_i+\boldsymbol{N}_i$,其中$\boldsymbol{M}_i=\lambda_i\boldsymbol{I}$是对角阵,$\boldsymbol{N}_i$是幂零阵 (因为是$m_i$阶所以至少$N_i^{m_i}=O$) 且 $\underbrace{\boldsymbol{M}_i\boldsymbol{N}_i=\boldsymbol{N}_i\boldsymbol{M}_i}_{\text{因为}\lambda_i\boldsymbol{I}\text{总是可交换的}}$.

$$\text{令}\boldsymbol{M}=\begin{pmatrix}\boldsymbol{M}_1&&&\\&\boldsymbol{M}_2&&\\&&\ddots&\\&&&\boldsymbol{M}_s\end{pmatrix},\boldsymbol{N}=\begin{pmatrix}\boldsymbol{N}_1&&&\\&\boldsymbol{N}_2&&\\&&\ddots&\\&&&\boldsymbol{N}_s\end{pmatrix}$$

则每个$\boldsymbol{M}_i$都是对角阵固然$\boldsymbol{M}$也为对角阵,$\boldsymbol{N}$也是幂零阵且$\boldsymbol{MN}=\boldsymbol{NM}$,且$\boldsymbol{J}=\boldsymbol{M}+\boldsymbol{N}$其中$\boldsymbol{M}$是对角阵,$\boldsymbol{N}$是幂零阵.

上述一行说明了满足了题干条件的 (1) ~ (3)

我们知道每个$\boldsymbol{J}_i$的特征多项式：$(\lambda-\lambda_i)^{m_i}$；有$Camley-Hamilton$定理得到 $(\boldsymbol{J}_i-\lambda_i\boldsymbol{I})^{m_i}=O$

因为$\lambda_1\sim\lambda_s$是互不相同的特征值故 $(\lambda-\lambda_1)^{m_1}\cdots(\lambda-\lambda_s)^{m_s}$ 都两两互素

那么$\exists g(x)$ 使得$g(\lambda)=(\lambda-\lambda_i)^{m_i}q_i(\lambda)+\lambda_i\ (\forall 1\leqslant i\leqslant s)$

$\Rightarrow g(\boldsymbol{J}_i)=\lambda_i\boldsymbol{I}=\boldsymbol{M}_i$

所以$g(\boldsymbol{J})=\operatorname{diag}\{g(\boldsymbol{J}_1),g(\boldsymbol{J}_2)\cdots\cdots g(\boldsymbol{J}_s)\}=\operatorname{diag}\{\boldsymbol{M}_1,\boldsymbol{M}_2\cdots\cdots\boldsymbol{M}_s\}=\boldsymbol{M}$

$\boldsymbol{N}=\boldsymbol{J}-\boldsymbol{M}=\boldsymbol{J}-g(\boldsymbol{J})$

故对$Jordan$标准型是满足定理的

下面考虑一般矩阵

$\boldsymbol{A}=\boldsymbol{PJP}^{-1}=\boldsymbol{P}(\boldsymbol{M}+\boldsymbol{N})\boldsymbol{P}^{-1}=\boldsymbol{PMP}^{-1}+\boldsymbol{PNP}^{-1}$

$\boldsymbol{PMP}^{-1}$即与一个对角阵相似的矩阵这就是可对角化的定义。$\boldsymbol{PNP}^{-1}$仍然是幂零阵直接用定义n次方即可看出

就令$\boldsymbol{B}=\boldsymbol{PMP}^{-1}\quad \boldsymbol{C}=\boldsymbol{PNP}^{-1}$

检查交换性：$\boldsymbol{BC}=\left(\boldsymbol{PMP}^{-1}\right)\left(\boldsymbol{PNP}^{-1}\right)=\boldsymbol{PNMP}^{-1}=\left(\boldsymbol{PNP}^{-1}\right)\left(\boldsymbol{PMP}^{-1}\right)$

下面检查多项式：仍然考察前面定义过的多项式$g(\lambda)$　此时$g(\boldsymbol{A})=g\left(\boldsymbol{PJP}^{-1}\right)=\boldsymbol{P}g(\boldsymbol{J})\boldsymbol{P}^{-1}=\boldsymbol{PMP}^{-1}=\boldsymbol{B}$

$\boldsymbol{C}=\boldsymbol{A}-\boldsymbol{B}=\boldsymbol{A}-g(\boldsymbol{A})$

下面考虑唯一性

设$\boldsymbol{A}=\boldsymbol{B}_1+\boldsymbol{C}_1=\boldsymbol{B}+\boldsymbol{C}$那么有$\boldsymbol{B}_1-\boldsymbol{B}=\boldsymbol{C}-\boldsymbol{C}_1$

由可交换条件知道$\boldsymbol{B}_1\boldsymbol{C}_1=\boldsymbol{C}_1\boldsymbol{B}_1\Rightarrow\boldsymbol{AB}_1=(\boldsymbol{B}_1+\boldsymbol{C}_1)\boldsymbol{B}_1=\boldsymbol{B}_1^2+\boldsymbol{C}_1\boldsymbol{B}_1=\boldsymbol{B}_1^2+\boldsymbol{B}_1\boldsymbol{C}_1=\boldsymbol{B}_1\boldsymbol{A}$

因为$\boldsymbol{B}_1$与$\boldsymbol{A}$可交换那么和A的多项式也可交换那么因为$\boldsymbol{B}=g(\boldsymbol{A})\Rightarrow\boldsymbol{B}$与$\boldsymbol{B}_1$可交换.同理$\boldsymbol{C}$与$\boldsymbol{C}_1$也可交换

又$\boldsymbol{B}_1-\boldsymbol{B}=\boldsymbol{C}-\boldsymbol{C}_1$显然$\boldsymbol{C}$与$\boldsymbol{C}_1$因为幂零那么$\boldsymbol{C}-\boldsymbol{C}_1$也是幂零由二项式展开即可

$\Rightarrow\boldsymbol{B}-\boldsymbol{B}_1$也是幂零的

由引理知道$\boldsymbol{B}_1\boldsymbol{B}$可交换且分别可对角化那么就可以同时对角化

那么存在可逆阵Q,使得$Q^{-1}BQ$和$Q^{-1}B_1Q$都是对角阵.

注意到$\left(\boldsymbol{Q}^{-1}\boldsymbol{BQ}-\boldsymbol{Q}^{-1}\boldsymbol{B}_1\boldsymbol{Q}\right)^{r+t}=\left(\boldsymbol{Q}^{-1}(\boldsymbol{B}-\boldsymbol{B}_1)\boldsymbol{Q}\right)^{r+t}=\boldsymbol{Q}^{-1}(\boldsymbol{B}-\boldsymbol{B}_1)^{r+t}\boldsymbol{Q}=\boldsymbol{O}$

两个对角阵之差仍是一个对角阵,这个差的幂要等于零矩阵

这两个矩阵必相等,由此即得$B=B_1$,于是$C=C_1$.

7.11 重点结论与知识

Theorem 7.11.1 (分块对角矩阵与全空间的直和分解下的性质)

设φ是线性空间V上的线性变换,V有一个直和分解 : $V = V_1 \oplus V_2 \oplus \cdots \oplus V_m$, 其中$V_i$都是$\varphi-$不变子空间.

(1)设φ限制在V_i上的特征多项式为$f_i(\lambda)$, 求证 : φ的特征多项式$f(\lambda) = f_1(\lambda) f_2(\lambda) \cdots f_m(\lambda)$

(2)设λ_0是φ的特征值, $V_0 = \{\boldsymbol{v} \in V \mid \boldsymbol{\varphi}(\boldsymbol{v}) = \lambda_0 \boldsymbol{v}\}$ 为特征子空间

$V_{i,0} = V_i \cap V_0 = \{\boldsymbol{v} \in V_i \mid \boldsymbol{\varphi}(\boldsymbol{v}) = \lambda_0 \boldsymbol{v}\}$, 求证 : $V_0 = V_{1,0} \oplus V_{2,0} \oplus \cdots \oplus V_{m,0}$

(3) 说明 : 分块对角矩阵的代数重数是每个分块代数重数之和, 几何重数是每个分块的几何重数之和

Proof (1)取V_i的一组基, 将它们拼成V的一组基.

记$\boldsymbol{A}_i$是φ在V_i上的限制在V_i所取基下的表示矩阵, 则φ在V的这组基下的表示矩阵为分块对角矩阵$\boldsymbol{A} = \operatorname{diag}\{\boldsymbol{A}_1, \boldsymbol{A}_2, \cdots, \boldsymbol{A}_m\}$

于是$f(\lambda) = |\lambda \boldsymbol{I}_n - \boldsymbol{A}| = |\lambda \boldsymbol{I} - \boldsymbol{A}_1| |\lambda \boldsymbol{I} - \boldsymbol{A}_2| \cdots |\lambda \boldsymbol{I} - \boldsymbol{A}_m|$ 即$f(\lambda) = f_1(\lambda) f_2(\lambda) \cdots f_m(\lambda)$.

(2)任取$\boldsymbol{\alpha} \in V_0$, 设$\boldsymbol{\alpha} = \boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_2 + \cdots + \boldsymbol{\alpha}_m$, 其中$\boldsymbol{\alpha}_i \in V_i$, 则

$\varphi(\alpha_1) + \varphi(\alpha_2) + \cdots + \varphi(\alpha_m) = \varphi(\boldsymbol{\alpha}) = \lambda_0 \boldsymbol{\alpha} = \lambda_0 \alpha_1 + \lambda_0 \alpha_2 + \cdots + \lambda_0 \alpha_m$注意到$\boldsymbol{\varphi}(\boldsymbol{\alpha}_i) \in V_i$

故由直和的充要条件 (分解唯一性) 可得$\boldsymbol{\varphi}(\boldsymbol{\alpha}_i) = \lambda_0 \alpha_i$ 即$\boldsymbol{\alpha}_i \in V_{i,0}$, 从而$V_0 = V_{1,0} + V_{2,0} + \cdots + V_{m,0}$.

注意到$V_{i,0} \subseteq V_i$, 故$V_{i,0} \cap (V_{1,0} + \cdots + V_{i-1,0}) \subseteq V_i \cap (V_1 + \cdots + V_{i-1}) = 0, \quad 2 \le i \le m$于是上述和为直和

Theorem 7.11.2 (特征值降阶公式)

设$\boldsymbol{A}$是$m \times n$矩阵, $\boldsymbol{B}$是$n \times m$矩阵, 且$m \ge n$.

求证 : $|\lambda \boldsymbol{I}_m - \boldsymbol{AB}| = \lambda^{m-n} |\lambda \boldsymbol{I}_n - \boldsymbol{BA}|$ 特别地, 若$\boldsymbol{A}, \boldsymbol{B}$都是$n$阶矩阵, 则$\boldsymbol{AB}$与$\boldsymbol{BA}$有相同的特征多项式.

Proof 当$\lambda \neq 0$时, 考虑下列分块矩阵 : $\begin{pmatrix} \lambda \boldsymbol{I}_m & \boldsymbol{A} \\ \boldsymbol{B} & \boldsymbol{I}_n \end{pmatrix}$ 因为$\lambda \boldsymbol{I}_m, \boldsymbol{I}_n$都是可逆矩阵, 故由行列式的降阶公式可得

$|\boldsymbol{I}_n| \cdot \left|\lambda \boldsymbol{I}_m - \boldsymbol{A} (\boldsymbol{I}_n)^{-1} \boldsymbol{B}\right| = |\lambda \boldsymbol{I}_m| \cdot \left|\boldsymbol{I}_n - \boldsymbol{B} (\lambda \boldsymbol{I}_m)^{-1} \boldsymbol{A}\right| \Longrightarrow |\lambda \boldsymbol{I}_m - \boldsymbol{AB}| = \lambda^{m-n} |\lambda \boldsymbol{I}_n - \boldsymbol{BA}|$ 成立.

当$\lambda = 0$时 若$m > n$ $WTS : |AB| = 0$, 则 $\mathrm{r}(\boldsymbol{AB}) \le \min\{\mathrm{r}(\boldsymbol{A}), \mathrm{r}(\boldsymbol{B})\} \le \min\{m, n\} = n < m$, 故$|-\boldsymbol{AB}| = 0$, 结论成立

当$\lambda = 0$时 若$m = n$, $WTS : |-AB| = |-BA|$ 则$|-\boldsymbol{AB}| = (-1)^n |\boldsymbol{A}||\boldsymbol{B}| = |-\boldsymbol{BA}|$, 结论也成立.

Theorem 7.11.3 (无公共特征值性质)

设$f(x)$和$m(x)$分别是m阶矩阵$\boldsymbol{A}$的特征多项式和极小多项式, $g(x)$和$n(x)$分别是n阶矩阵$\boldsymbol{B}$的特征多项式和极小多项式

证明以下结论等价 :

(1)$\boldsymbol{A}, \boldsymbol{B}$没有公共的特征值

(2)$(f(x), g(x)) = 1$或 $(f(x), n(x)) = 1$或 $(m(x), g(x)) = 1$或 $(m(x), n(x)) = 1$

(3)$f(\boldsymbol{B})$或$m(\boldsymbol{B})$或$g(\boldsymbol{A})$或$n(\boldsymbol{A})$是可逆矩阵.

Proposition 7.6 (AXB 形)

1.设V为n阶矩阵全体构成的线性空间, V上的线性变换$\boldsymbol{\varphi}$定义为$\boldsymbol{\varphi}(\boldsymbol{X}) = \boldsymbol{AXA}$, 其中$\boldsymbol{A} \in V$.

证明 : 若$\boldsymbol{A}$可对角化, 则$\boldsymbol{\varphi}$也可对角化.

2.设$\boldsymbol{A}, \boldsymbol{B}$分别为$m, n$阶矩阵, V为$m \times n$矩阵全体构成的线性空间, V上的线性变换$\boldsymbol{\varphi}$定义为 : $\boldsymbol{\varphi}(\boldsymbol{X}) = \boldsymbol{AXB}$.

设$\boldsymbol{A}$的特征值为$\lambda_i (1 \le i \le m)$, $\boldsymbol{B}$的特征值为$\mu_j (1 \le j \le n)$.

求证 : 线性变换$\boldsymbol{\varphi}$的特征值为$\lambda_i \mu_j (1 \le i \le m; 1 \le j \le n)$.

3.设$\boldsymbol{A},\boldsymbol{B}$分别为$m,n$阶矩阵,$V$为$m\times n$矩阵全体构成的线性空间,$V$上的线性变换$\varphi$定义为:$\varphi(\boldsymbol{X})=\boldsymbol{AXB}$

证明:φ是线性自同构的充要条件是$\boldsymbol{A},\boldsymbol{B}$都是可逆矩阵.

4.设$\boldsymbol{A},\boldsymbol{B}$分别为$m,n$阶矩阵,$V$为$m\times n$矩阵全体构成的线性空间,$V$上的线性变换$\varphi$定义为:$\varphi(X)=AXB$.

证明:φ是幂零线性变换的充要条件是A,B至少有一个是幂零矩阵.

Proof 法一: 证明设$\boldsymbol{P}$为n阶可逆矩阵,使得$\boldsymbol{P}^{-1}\boldsymbol{AP}=\Lambda=\operatorname{diag}\{\lambda_1,\lambda_2,\cdots,\lambda_n\}$,则$\boldsymbol{P}'\boldsymbol{A}'(\boldsymbol{P}')^{-1}=\Lambda$,即$\boldsymbol{A}'$也可对角化.

设$\boldsymbol{P}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)$, $(\boldsymbol{P}')^{-1}=(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n)$分别为两个矩阵的列分块,则

$\boldsymbol{A\alpha}_i=\lambda_i\boldsymbol{\alpha}_i$, $\boldsymbol{A}'\boldsymbol{\beta}_j=\lambda_j\boldsymbol{\beta}_j$, $1\le i,j\le n$且$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$线性无关,$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n$线性无关.

可知,$\left\{\boldsymbol{\alpha}_i\boldsymbol{\beta}_j',1\le i,j\le n\right\}$是$V$中$n^2$个线性无关的矩阵.注意到$\boldsymbol{\varphi}\left(\boldsymbol{\alpha}_i\boldsymbol{\beta}_j'\right)=\boldsymbol{A\alpha}_i\boldsymbol{\beta}_j'\boldsymbol{A}=(\boldsymbol{A\alpha}_i)\left(\boldsymbol{A}'\boldsymbol{\beta}_j\right)'=\lambda_i\lambda_j\boldsymbol{\alpha}_i\boldsymbol{\beta}_j'$

法二: 由于$\boldsymbol{A}$可对角化,故存在可逆矩阵$\boldsymbol{P}$,使得$\boldsymbol{P}^{-1}\boldsymbol{AP}=\Lambda$为对角矩阵.

可知,$\boldsymbol{\varphi}$在基础矩阵这组基下的表示矩阵为$\boldsymbol{A}\otimes\boldsymbol{A}'$,于是$\left(\boldsymbol{P}\otimes(\boldsymbol{P}')^{-1}\right)^{-1}(\boldsymbol{A}\otimes\boldsymbol{A}')\left(\boldsymbol{P}\otimes(\boldsymbol{P}')^{-1}\right)=\Lambda\otimes\Lambda$为对角矩阵

即$\boldsymbol{A}\otimes\boldsymbol{A}'$可对角化,从而$\varphi$可对角化.

Proof

取V的一组基为$m\times n$基础矩阵$\boldsymbol{E}_{11},\cdots,\boldsymbol{E}_{1n},\boldsymbol{E}_{21},\cdots,\boldsymbol{E}_{2n},\cdots,\boldsymbol{E}_{m1},\cdots,\boldsymbol{E}_{mn}$

我们首先证明φ在这组基下的表示矩阵为$\boldsymbol{A}\otimes\boldsymbol{B}'$.

事实上,$\boldsymbol{\varphi}(\boldsymbol{E}_{ij})=\boldsymbol{AE}_{ij}\boldsymbol{B}=\boldsymbol{Ae}_i\boldsymbol{f}_j'\boldsymbol{B}=\sum_{k=1}^m\sum_{l=1}^n a_{ki}b_{jl}\boldsymbol{E}_{kl}$,其中$\boldsymbol{e}_i,\boldsymbol{f}_j$分别是$m,n$维标准单位列向量,故$\boldsymbol{\varphi}$的表示矩阵为

$$\begin{pmatrix} a_{11}\boldsymbol{B}' & a_{12}\boldsymbol{B}' & \cdots & a_{1m}\boldsymbol{B}' \\ a_{21}\boldsymbol{B}' & a_{22}\boldsymbol{B}' & \cdots & a_{2m}\boldsymbol{B}' \\ \vdots & \vdots & & \vdots \\ a_{m1}\boldsymbol{B}' & a_{m2}\boldsymbol{B}' & \cdots & a_{mm}\boldsymbol{B}' \end{pmatrix}=\boldsymbol{A}\otimes\boldsymbol{B}'.$$

注意到$\boldsymbol{B}'$与$\boldsymbol{B}$有相同的特征值,故由可知,$\boldsymbol{\varphi}$的特征值为$\lambda_i\mu_j$.

Proof 法一:证明容易验证φ是线性变换.若$\boldsymbol{A},\boldsymbol{B}$都是可逆矩阵,则$\boldsymbol{\psi}(\boldsymbol{X})=\boldsymbol{A}^{-1}\boldsymbol{XB}^{-1}$是$\varphi$的逆线性变换

来证明必要性.若$\boldsymbol{A}$是不可逆矩阵,则我们可证明$\boldsymbol{\varphi}$不是单映射,即存在$\boldsymbol{X}\ne\boldsymbol{O}$,使得$\boldsymbol{\varphi}(\boldsymbol{X})=\boldsymbol{AXB}=\boldsymbol{O}$,从而$\boldsymbol{\varphi}$不是可逆变换.

事实上,若$\boldsymbol{A}$的秩等于$r<m$,则存在可逆矩阵$\boldsymbol{P}^{m\times m}$和$\boldsymbol{Q}^{m\times m}$,使得$\boldsymbol{PAQ}=\begin{pmatrix} I_r & O \\ O & O\end{pmatrix}_{m\times m}$.

令$\boldsymbol{C}=\begin{pmatrix} O & O \\ O & I_{n-r}\end{pmatrix}_{m\times m}$,则$\boldsymbol{PAQC}=\boldsymbol{O}$,而$\boldsymbol{P}$是可逆矩阵,故$\boldsymbol{AQC}=\boldsymbol{O}$,再令$\boldsymbol{X}=\boldsymbol{QC}$即可.

同理,若$\boldsymbol{B}$的秩小于n,也可以证明φ不是可逆变换.

法二:已知φ的特征值为$\lambda_i\mu_j\ne0$ φ是自同构当且仅当φ所有的特征值$\lambda_i\mu_j\ne0$,这当且仅当所有的$\lambda_i\ne0$以及所有的$\mu_j\ne0$

这也当且仅当$\boldsymbol{A},\boldsymbol{B}$都是可逆矩阵.

Proof 先证充分性.不妨设$\boldsymbol{A}$是幂零矩阵,即存在正整数k,使得$\boldsymbol{A}^k=\boldsymbol{O}$,则$\varphi^k(\boldsymbol{X})=\boldsymbol{A}^k\boldsymbol{X}B^k=\boldsymbol{O}$,即$\varphi^k=0$,于是$\varphi$是幂零线性变换.

再证必要性.设$\boldsymbol{A},\boldsymbol{B}$都不是幕零矩阵,即对任意给定的正整数$k$,$\boldsymbol{A}^k\ne\boldsymbol{O}$,$B^k\ne O$,只要证明$\boldsymbol{\varphi}^k\ne0$即可.我们给出以下4种证法.

证法1不妨设$\boldsymbol{A}^k$的第i列非零,$\boldsymbol{B}^k$的第j行非零,即有列向量$\boldsymbol{A}^k\boldsymbol{e}_i\ne0$,行向量$\boldsymbol{f}_j'\boldsymbol{B}^k\ne0$,其中$\boldsymbol{e}_i,\boldsymbol{f}_j$分别是$m,n$维标准单位列向量

于是$\varphi^k(\boldsymbol{E}_{ij})=\boldsymbol{A}^k\boldsymbol{E}_{ij}\boldsymbol{B}^k=\boldsymbol{A}^ke_i\boldsymbol{f}_j'\boldsymbol{B}^k=\left(\boldsymbol{A}^ke_i\right)\left(\boldsymbol{f}_j'\boldsymbol{B}^k\right)\ne\boldsymbol{O}$.

证法2设$\boldsymbol{P}_i,\boldsymbol{Q}_i$为可逆矩阵,使得$\boldsymbol{P}_1\boldsymbol{A}^k\boldsymbol{Q}_1=\operatorname{diag}\{\boldsymbol{I}_r,\boldsymbol{O}\}$,$\boldsymbol{P}_2\boldsymbol{B}^k\boldsymbol{Q}_2=\operatorname{diag}\{\boldsymbol{I}_s,\boldsymbol{O}\}$,不妨设$r\ge s\ge1$,于是

$\boldsymbol{\varphi}^k(\boldsymbol{Q}_1\boldsymbol{P}_2)=\boldsymbol{P}_1^{-1}\operatorname{diag}\{\boldsymbol{I}_r,\boldsymbol{O}\}\operatorname{diag}\{\boldsymbol{I}_s,\boldsymbol{O}\}\boldsymbol{Q}_2^{-1}=\boldsymbol{P}_1^{-1}\operatorname{diag}\{\boldsymbol{I}_s,\boldsymbol{O}\}\boldsymbol{Q}_2^{-1}\ne\boldsymbol{O}$

证法3可知,φ^k在基础矩阵这组基下的表示矩阵为$\boldsymbol{A}^k\otimes(\boldsymbol{B}^k)'$, 再由$Kronecker$积的定义可知$\boldsymbol{A}^k\otimes(\boldsymbol{B}^k)'\ne\boldsymbol{O}$, 于是$\boldsymbol{\varphi}^k\ne0$

证法4可知,φ是幂零线性变换当且仅当φ的所有特征值都等于零.

由于$\boldsymbol{A},\boldsymbol{B}$都不是幂零矩阵,故$\boldsymbol{A}$的特征值$\lambda_i$不全为零,$\boldsymbol{B}$的特征值$\mu_j$不全为零.可知,$\varphi$的特征值$\lambda_i\mu_j$也不全为零,从而$\boldsymbol{\varphi}$不是幂零线性变换

Proposition 7.7 ($AX-XB$形问题)

(1) 设$\boldsymbol{A},\boldsymbol{B},\boldsymbol{C}$分别是$m\times m,n\times n,m\times n$矩阵, 满足：$\boldsymbol{AC}=\boldsymbol{CB},\mathrm{r}(\boldsymbol{C})=r$.
求证：$\boldsymbol{A}$和$\boldsymbol{B}$至少有r个相同的特征值.
(2) 设A,B分别是数域P上的n级、m级方阵
证明：若A,B有$r(0<r\le\min\{n,m\})$个两两不等的公共特征值, 那么矩阵方程$AX-XB=O$有秩为r的矩阵解
(3) 设A,B,C是n阶矩阵, 其中$C=AB-BA$.若它们满足条件$AC=CA,BC=CB$
求证：C的特征值全为零.又若将条件减弱为$ABC=CAB,BAC=CBA$, 则上述结论不再成立.
(4) 设V为n阶矩阵全体构成的线性空间，V上的线性变换φ定义为$\varphi(X)=AX-XA$, 其中$A\in V$.
证明：若$\boldsymbol{A}$可对角化, 则φ也可对角化.
(5) 设$\boldsymbol{A}$为m阶矩阵, $\boldsymbol{B}$为n阶矩阵, 求证：若$\boldsymbol{A},\boldsymbol{B}$没有公共的特征值, 则矩阵方程$\boldsymbol{AX}=\boldsymbol{XB}$只有零解$\boldsymbol{X}=\boldsymbol{O}$
(6) 设$\boldsymbol{A},\boldsymbol{B}$分别为$m,n$阶矩阵　V为$m\times n$矩阵全体构成的线性空间　V上的线性变换φ定义为：$\varphi(\boldsymbol{X})=\boldsymbol{AX}-\boldsymbol{X}$.
求证　：　φ是线性自同构的充要条件是$\boldsymbol{A},\boldsymbol{B}$没有公共的特征值. $\Longrightarrow$ 此时, 对任一$m\times n$矩阵$\boldsymbol{C}$, 矩阵方程$\boldsymbol{AX}-\boldsymbol{XB}=\boldsymbol{C}$存在唯一解.
(7) 设$\boldsymbol{A},\boldsymbol{B}$分别为$m,n$阶矩阵, V为$m\times n$矩阵全体构成的线性空间, V上的线性变换$\boldsymbol{\varphi}$定义为$\boldsymbol{\varphi}(\boldsymbol{X})=\boldsymbol{AX}-\boldsymbol{XB}$
设$\boldsymbol{A}$的特征值为$\lambda_i(1\le i\le m)$，$\boldsymbol{B}$的特征值为$\mu_j(1\le j\le n)$
求证：线性变换$\boldsymbol{\varphi}$的特征值为$\lambda_i-\mu_j(1\le i\le m;1\le j\le n)$.
(8) 设$\boldsymbol{A},\boldsymbol{B}$分别为$m,n$阶矩阵, V为$m\times n$矩阵全体构成的线性空间, V上的线性变换φ定义为：$\boldsymbol{\varphi}(\boldsymbol{X})=\boldsymbol{AX}-\boldsymbol{XB}$.
证明：若$\boldsymbol{A},\boldsymbol{B}$都是幂零矩阵, 则$\varphi$是幂零线性变换.

Proof (1) 设P为m阶非异阵, Q为n阶非异阵, 使得$PCQ=\begin{pmatrix} I_r & O\\ O & O\end{pmatrix}$

$PAP^{-1}PCQ=PCQQ^{-1}BQ\Longrightarrow PAP^{-1}\begin{pmatrix} I_r & O\\ O & O\end{pmatrix}=\begin{pmatrix} I_r & O\\ O & O\end{pmatrix}PBP^{-1}$

设$PAP^{-1}=\begin{pmatrix}\boldsymbol{A}_{11} & \boldsymbol{A}_{12}\\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22}\end{pmatrix}\quad PBP^{-1}=\begin{pmatrix}\boldsymbol{B}_{11} & \boldsymbol{B}_{12}\\ \boldsymbol{B}_{21} & \boldsymbol{B}_{22}\end{pmatrix}$为对应的分块,
$\Longrightarrow \boldsymbol{A}_{11}=\boldsymbol{B}_{11},\boldsymbol{A}_{21}=\boldsymbol{O},\boldsymbol{B}_{12}=\boldsymbol{O}$.
$\Longrightarrow PAP^{-1}=\begin{pmatrix}\boldsymbol{A}_{11} & \boldsymbol{A}_{12}\\ O & \boldsymbol{A}_{22}\end{pmatrix}\quad PBP^{-1}=\begin{pmatrix}A_{11} & O\\ \boldsymbol{B}_{21} & \boldsymbol{B}_{22}\end{pmatrix}$
于是$|\lambda\boldsymbol{I}_n-\boldsymbol{A}|=|\lambda\boldsymbol{I}_r-\boldsymbol{A}_{11}|\cdot|\lambda\boldsymbol{I}_{n-r}-\boldsymbol{A}_{22}|$，$|\lambda\boldsymbol{I}_n-\boldsymbol{B}|=|\lambda\boldsymbol{I}_r-\boldsymbol{B}_{11}|\cdot|\lambda\boldsymbol{I}_{n-r}-\boldsymbol{B}_{22}|$, 从而$\boldsymbol{A},\boldsymbol{B}$至少有$r$个相同的特征值 (即$\boldsymbol{A}_{11}=\boldsymbol{B}_{11}$的特征值).

(2) 设$\lambda_1,\lambda_2,\cdots,\lambda_r$两两不等, 它们是$A$与$B$的公共特征值, 则它们也是$B'$的特征值.
从而在P^n中有非零向量$\alpha_1,\alpha_2,\cdots,\alpha_r$使得$A\alpha_i=\lambda_i\alpha_i(i=1,2,\cdots,r)$.
同时, 在P^m中有非零向量$\beta_1,\beta_2,\cdots,\beta_r$使得$B'\beta_i=\lambda_i\beta_i$, 于是$\beta_i'B=\lambda_i\beta_i'(i=1,2,\cdots,r)$

现在令$C=\begin{pmatrix}\alpha_1 & \alpha_2 & \cdots & \alpha_r\end{pmatrix}\begin{pmatrix}\beta_1'\\ \beta_2'\\ \vdots\\ \beta_r'\end{pmatrix}$则$C$是$n\times m$矩阵, 满足

$$AC-CB=A\begin{pmatrix}\alpha_1&\alpha_2&\cdots&\alpha_r\end{pmatrix}\begin{pmatrix}\beta_1'\\\beta_2'\\\vdots\\\beta_r'\end{pmatrix}-\begin{pmatrix}\alpha_1&\alpha_2&\cdots&\alpha_r\end{pmatrix}\begin{pmatrix}\beta_1'\\\beta_2'\\\vdots\\\beta_r'\end{pmatrix}$$

$$=\begin{pmatrix}\lambda_1\alpha_1&\lambda_2\alpha_2&\cdots&\lambda_r\alpha_r\end{pmatrix}\begin{pmatrix}\beta_1'\\\beta_2'\\\vdots\\\beta_r'\end{pmatrix}-\begin{pmatrix}\alpha_1&\alpha_2&\cdots&\alpha_r\end{pmatrix}\begin{pmatrix}\lambda_1\beta_1'\\\lambda_2\beta_2'\\\vdots\\\lambda_r\beta_r'\end{pmatrix}$$

$$=(\lambda_1\alpha_1\beta_1'+\lambda_2\alpha_2\beta_2'+\cdots+\lambda_r\alpha_r\beta_r')-(\lambda_1\alpha_1\beta_1'+\lambda_2\alpha_2\beta_2'+\cdots+\lambda_r\alpha_r\beta_r')$$

$$=0$$

这说明C是矩阵方程$AX-XB=O$的解.下面证明$r(C)=r$:

注意到$\alpha_1,\alpha_2,\cdots,\alpha_r$是线性无关的,同时$\beta_1,\beta_2,\cdots,\beta_r$也是线性无关的,所以$r\begin{pmatrix}\alpha_1&\alpha_2&\cdots&\alpha_r\end{pmatrix}=r\begin{pmatrix}\beta_1'\\\beta_2'\\\vdots\\\beta_r'\end{pmatrix}=r$

一方面我们有$r(C)\le r\begin{pmatrix}\alpha_1&\alpha_2&\cdots&\alpha_r\end{pmatrix}=r$

另一方面,由$slverster$不等式有$r(C)\ge r\begin{pmatrix}\alpha_1&\alpha_2&\cdots&\alpha_r\end{pmatrix}+r\begin{pmatrix}\beta_1'\\\beta_2'\\\vdots\\\beta_r'\end{pmatrix}-r=r$所以$r(C)=r$.

Proof

由$\boldsymbol{AC}=\boldsymbol{CA}$可知,对任意的正整数$k$,$\boldsymbol{C}^k=\boldsymbol{C}^{k-1}\boldsymbol{AB}-\boldsymbol{C}^{k-1}\boldsymbol{BA}=\boldsymbol{A}\left(\boldsymbol{C}^{k-1}\boldsymbol{B}\right)-\left(\boldsymbol{C}^{k-1}\boldsymbol{B}\right)\boldsymbol{A}$.

由迹的线性和交换性可得 $\operatorname{tr}\left(\boldsymbol{C}^k\right)=0(k\ge1)$,可知$\boldsymbol{C}$为幂零矩阵,从而$\boldsymbol{C}$的特征值全为零

注从上述证明不难看出：只需要$\boldsymbol{AC}=\boldsymbol{CA}$和$\boldsymbol{BC}=\boldsymbol{CB}$这两个条件中的一个就能证明本题的结论.

如果这两个条件都有，我们将在后面给出另外两个证明：一个是利用矩阵乘法交换性诱导的同时性质，另一个是$Jordan$标准型理论

另外,我们还将证明A,B,C可同时上三角化.

如将条件减弱为如题所述,则结论不再成立,可参考下面的反例:$\boldsymbol{A}=\begin{pmatrix}1&0\\0&-1\end{pmatrix}$, $\boldsymbol{B}=\begin{pmatrix}0&\frac{1}{2}\\-\frac{1}{2}&0\end{pmatrix}$, $\boldsymbol{C}=\begin{pmatrix}0&1\\1&0\end{pmatrix}$

由计算可得$C=AB-BA,ABC=CAB,CBA=BAC$，但C的特征值为1和-1

Proof

法一： 设$\boldsymbol{P}$为n阶可逆矩阵，使得$\boldsymbol{P}^{-1}\boldsymbol{AP}=\Lambda=\operatorname{diag}\{\lambda_1,\lambda_2,\cdots,\lambda_n\}$，则$\boldsymbol{P}'\boldsymbol{A}'(\boldsymbol{P}')^{-1}=\Lambda$，即$\boldsymbol{A}'$也可对角化.

设$\boldsymbol{P}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)$， $(\boldsymbol{P}')^{-1}=(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n)$分别为两个矩阵的列分块,则

$\boldsymbol{A\alpha}_i=\lambda_i\boldsymbol{\alpha}_i$, $\boldsymbol{A}'\boldsymbol{\beta}_j=\lambda_j\boldsymbol{\beta}_j$, $1\le i,j\le n$且$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$线性无关,$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n$线性无关.

可知,$\left\{\boldsymbol{\alpha}_i\boldsymbol{\beta}_j',1\le i,j\le n\right\}$是$V$中$n^2$个线性无关的矩阵.

注意到$\boldsymbol{\varphi}\left(\boldsymbol{\alpha}_i\boldsymbol{\beta}_j'\right)=\boldsymbol{A\alpha}_i\boldsymbol{\beta}_j'-\boldsymbol{\alpha}_i\boldsymbol{\beta}_j'\boldsymbol{A}=(\boldsymbol{A\alpha}_i)\boldsymbol{\beta}_j'-\boldsymbol{\alpha}_i\left(\boldsymbol{A}'\boldsymbol{\beta}_j\right)'=(\lambda_i-\lambda_j)\boldsymbol{\alpha}_i\boldsymbol{\beta}_j'$故$\varphi$有$n^2$个线性无关的特征向量,从而可对角化.

法二： 由于$\boldsymbol{A}$可对角化,故存在可逆矩阵$\boldsymbol{P}$,使得$\boldsymbol{P}^{-1}\boldsymbol{AP}=\Lambda$为对角矩阵.

可知,$\boldsymbol{\varphi}$在基础矩阵这组基下的表示矩阵为$\boldsymbol{A}\otimes\boldsymbol{I}_n-\boldsymbol{I}_n\otimes\boldsymbol{A}'$,于是

$\left(\boldsymbol{P}\otimes(\boldsymbol{P}')^{-1}\right)^{-1}(\boldsymbol{A}\otimes\boldsymbol{I}_n-\boldsymbol{I}_n\otimes\boldsymbol{A}')\left(\boldsymbol{P}\otimes(\boldsymbol{P}')^{-1}\right)=\Lambda\otimes\boldsymbol{I}_n-\boldsymbol{I}_n\otimes\Lambda$为对角矩阵

即$A\otimes I_n-I_n\otimes A'$可对角化,从而$\varphi$可对角化.

Proof

(5) 证法1设$f(\lambda)=|\lambda \boldsymbol{I}_m-\boldsymbol{A}|$为$\boldsymbol{A}$的特征多项式,则由$CH$定理可知$f(\boldsymbol{A})=\boldsymbol{O}$,再由$\boldsymbol{AX}=\boldsymbol{XB}$可得$O=f(A)X=Xf(B)$

因为$\boldsymbol{A},\boldsymbol{B}$没有公共的特征值,故可知,$f(\boldsymbol{B})$是可逆矩阵,从而由上式即得$\boldsymbol{X}=\boldsymbol{O}$

证法2任取矩阵方程的一个解$\boldsymbol{X}=\boldsymbol{C}$,若$\boldsymbol{C}\neq\boldsymbol{O}$,则 $\mathrm{r}(\boldsymbol{C})=r\geq 1$.

可知,$\boldsymbol{A},\boldsymbol{B}$至少有$r$个相同的特征值,这与$\boldsymbol{A},\boldsymbol{B}$没有公共的特征值相矛盾.因此$C=O$,即矩阵方程只有零解.

Proof

若$\boldsymbol{A},\boldsymbol{B}$没有公共的特征值,则由例可知,$\boldsymbol{\varphi}$是$V$上的单映射,从而是线性自同构.

若$\boldsymbol{A},\boldsymbol{B}$有公共的特征值$\lambda_0$,则$\lambda_0$也是$\boldsymbol{B}'$的特征值.设$\boldsymbol{\alpha},\boldsymbol{\beta}$为对应的特征向量,即$\boldsymbol{A\alpha}=\lambda_0\boldsymbol{\alpha},\boldsymbol{B}'\boldsymbol{\beta}=\lambda_0\boldsymbol{\beta}$,则$\boldsymbol{\alpha\beta}'\neq\boldsymbol{O}$且

$\boldsymbol{\varphi}(\boldsymbol{\alpha\beta}')=(\boldsymbol{A\alpha})\boldsymbol{\beta}'-\boldsymbol{\alpha}(\boldsymbol{B}'\boldsymbol{\beta})'=\lambda_0\boldsymbol{\alpha\beta}'-\lambda_0\boldsymbol{\alpha\beta}'=\boldsymbol{O}$

于是 $\mathrm{Ker}\varphi\neq 0$,从而$\varphi$不是线性自同构.

Proof 取V的一组基为$m\times n$基础矩阵:$\boldsymbol{E}_{11},\cdots,\boldsymbol{E}_{1n},\boldsymbol{E}_{21},\cdots,\boldsymbol{E}_{2n},\cdots,\boldsymbol{E}_{m1},\cdots,\boldsymbol{E}_{mn}$

类似讨论可得,$\boldsymbol{\varphi}$在上述基下的表示矩阵为$\boldsymbol{A}\otimes\boldsymbol{I}_n-\boldsymbol{I}_m\otimes\boldsymbol{B}'$

可知,存在m阶可逆矩阵$\boldsymbol{P}$以及n阶可逆矩阵$\boldsymbol{Q}$,使得

$$\boldsymbol{P}^{-1}\boldsymbol{AP}=\begin{pmatrix}\lambda_1 & * & * & * \\ & \lambda_2 & * & * \\ & & \ddots & \vdots \\ & & & \lambda_m\end{pmatrix},\quad \boldsymbol{Q}^{-1}\boldsymbol{B}'\boldsymbol{Q}=\begin{pmatrix}\mu_1 & * & * & * \\ & \mu_2 & * & * \\ & & \ddots & \vdots \\ & & & \mu_n\end{pmatrix}$$

注意到$(\boldsymbol{P}\otimes\boldsymbol{Q})^{-1}(\boldsymbol{A}\otimes\boldsymbol{I}_n-\boldsymbol{I}_m\otimes\boldsymbol{B}')(\boldsymbol{P}\otimes\boldsymbol{Q})=\left(\boldsymbol{P}^{-1}\boldsymbol{AP}\right)\otimes\boldsymbol{I}_n-\boldsymbol{I}_m\otimes\left(\boldsymbol{Q}^{-1}\boldsymbol{B}'\boldsymbol{Q}\right)$是一个上三角矩阵,其主对角元素依次为

$\lambda_1-\mu_1,\cdots,\lambda_1-\mu_n,\lambda_2-\mu_1,\cdots,\lambda_2-\mu_n,\cdots,\lambda_m-\mu_1,\cdots,\lambda_m-\mu_n$由此即得结论

Proof 因为$\boldsymbol{A},\boldsymbol{B}$都是幂零矩阵,所以它们的特征值都为零.可知,$\boldsymbol{\varphi}$的特征值也都为零,于是$\varphi$是幂零线性变换.也可由矩阵的运算直接证明

Theorem 7.11.4 (特征多项式与主子式联系)

设n阶矩阵$\boldsymbol{A}$的特征多项式为$f(\lambda)=\lambda^n+a_1\lambda^{n-1}+\cdots+a_{n-1}\lambda+a_n$.

求证:a_r等于$(-1)^r$乘以$\boldsymbol{A}$的所有r阶主子式之和,即$a_r=(-1)^r\sum\limits_{1\leq i_1<i_2<\cdots<i_r\leq n}A\begin{pmatrix}i_1 & i_2 & \cdots & i_r\\ i_1 & i_2 & \cdots & i_r\end{pmatrix},\quad 1\leq r\leq n$

进一步,若设$\boldsymbol{A}$的特征值为$\lambda_1,\lambda_2,\cdots,\lambda_n$,则 $\sum\limits_{1\leq i_1<i_2<\cdots<i_r\leq n}\lambda_{i_1}\lambda_{i_2}\cdots\lambda_{i_r}=\sum\limits_{1\leq i_1<i_2<\cdots<i_r\leq n}\boldsymbol{A}\begin{pmatrix}i_1 & i_2 & \cdots & i_r\\ i_1 & i_2 & \cdots & i_r\end{pmatrix},\quad 1\leq r\leq n$

Proof A,B为n阶矩阵。

$$|A+B|=|A|+|B|+\sum_{1\leqslant k\leqslant n-1}\left(\sum_{\substack{1\leqslant i_1<i_2<\cdots<i_k\leqslant n\\ 1\leqslant j_1<j_2<\cdots<j_k\leqslant n}}A\begin{pmatrix}i_1 & i_2 & \cdots & i_k\\ j_1 & j_2 & \cdots & j_k\end{pmatrix}\widehat{B}\begin{pmatrix}i_1 & i_2 & \cdots & i_k\\ j_1 & j_2 & \cdots & j_k\end{pmatrix}\right)$$

$$\Longrightarrow|\lambda I_n-A|=\lambda^n+(-1)^n|A|+\sum_{1\leqslant k\leqslant n-1}\left(\sum_{1\leqslant i_1<i_2<\cdots<i_k\leqslant n}(\lambda I_n)\begin{pmatrix}i_1 & i_2 & \cdots & i_k\\ i_1 & i_2 & \cdots & i_k\end{pmatrix}\widehat{(-A)}\begin{pmatrix}i_1 & i_2 & \cdots & i_k\\ i_1 & i_2 & \cdots & i_k\end{pmatrix}\right)$$

$$\Longrightarrow |\lambda I_n - A| = \lambda^n + (-1)^n |A| + \sum_{1\leqslant k\leqslant n-1} \lambda^{n-k}\left(\sum_{1\leqslant i_1<i_2<\cdots<i_k\leqslant n} (-1)^k \cdot A\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ i_1 & i_2 & \cdots & i_k \end{pmatrix}\right)$$

$$\Longrightarrow f(\lambda)\text{的}\lambda^{n-k}\text{的系数即为}a_k = \sum_{1\leqslant i_1<i_2<\cdots<i_k\leqslant n} (-1)^k \cdot A\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ i_1 & i_2 & \cdots & i_k \end{pmatrix}$$

第二个结论由$Vieta$定理即得.

Theorem 7.11.5 (矩阵可交换—同时结论)

1.设φ,ψ是复线性空间V上乘法可交换的线性变换，即$\varphi\psi=\psi\varphi$

求证：φ的特征子空间是ψ的不变子空间，ψ的特征子空间是φ的不变子空间.

2.设φ,ψ是复线性空间V上乘法可交换的线性变换，求证：φ,ψ至少有一个公共的特征向量.

3.设φ,ψ是数域$\mathbb{F}$上线性空间V上的乘法可交换的线性变换，且φ,ψ的特征值都在$\mathbb{F}$中

求证：φ,ψ的特征子空间互为不变子空间，并且φ,ψ至少有一个公共的特征向量.

代数版本：若数域$\mathbb{F}$上的n阶矩阵$\boldsymbol{A},\boldsymbol{B}$乘法可交换，且它们的特征值都在$\mathbb{F}$中

则$\boldsymbol{A},B$的特征子空间互为不变子空间，并且A,B在$\mathbb{F}^n$中至少有一个公共的特征向量.

4.设$\boldsymbol{A},\boldsymbol{B}$是数域$\mathbb{F}$上的$n$阶矩阵，满足：$\boldsymbol{AB}=\boldsymbol{BA}$且$\boldsymbol{A},\boldsymbol{B}$的特征值都在$\mathbb{F}$中

求证：$\boldsymbol{A},\boldsymbol{B}$在$\mathbb{F}$上可同时上三角化，即存在$\mathbb{F}$上的可逆矩阵$\boldsymbol{P}$，使得$\boldsymbol{P}^{-1}\boldsymbol{AP}$和$\boldsymbol{P}^{-1}\boldsymbol{BP}$都是上三角矩阵.

5.设$\boldsymbol{\varphi},\boldsymbol{\psi}$是数域$\mathbb{F}$上$n$维线性空间$V$上的线性变换，满足：$\varphi\psi=\psi\varphi$且$\varphi,\psi$都可对角化

求证：φ,ψ可同时对角化，即存在V的一组基，使得φ,ψ在这组基下的表示矩阵都是对角矩阵.

代数版本：设$\boldsymbol{A},\boldsymbol{B}$是数域$\mathbb{F}$上的$n$阶矩阵，满足：$\boldsymbol{AB}=\boldsymbol{BA}$且$A,B$都在$\mathbb{F}$上可对角化

则$\boldsymbol{A},\boldsymbol{B}$在$\mathbb{F}$上可同时对角化，即存在$\mathbb{F}$上的可逆矩阵$\boldsymbol{P}$，使得$P^{-1}\boldsymbol{AP}$和$\boldsymbol{P}^{-1}\boldsymbol{BP}$都是对角矩阵

Proof 1.由代数基本定理以及线性方程组的求解理论可知，$n(n\geq 1)$维复线性空间上的线性变换或n阶复矩阵至少有一个特征值和特征

任取线性变换φ的一个特征值λ_0，设V_0是特征值λ_0的特征子空间，则对任意的$\alpha\in V_0$，有

$\varphi\psi(\alpha)=\psi\varphi(\alpha)=\psi\left(\lambda_0\alpha\right)=\lambda_0\psi(\alpha)$即$\boldsymbol{\psi}(\alpha)\in V_0$，因此$V_0$是$\boldsymbol{\psi}$的不变子空间.同理可证$\boldsymbol{\psi}$的特征子空间是$\varphi$的不变子空间.

2.证明任取$\boldsymbol{\varphi}$的特征值λ_0及其特征子空间V_0，可知，V_0是$\psi-$不变子空间.

将线性变换$\boldsymbol{\psi}$限制在V_0上，由于V_0是维数大于零的复线性空间，故 $\boldsymbol{\psi}|_{V_0}$ 至少有一个特征值μ_0及其特征向量$\boldsymbol{\alpha}\in V_0$

从而$\boldsymbol{\varphi}(\boldsymbol{\alpha})=\lambda_0\boldsymbol{\alpha},\boldsymbol{\psi}(\boldsymbol{\alpha})=\mu_0\boldsymbol{\alpha}$，于是$\alpha$就是$\varphi,\psi$的公共特征向量.

3.由线性方程组的求解理论可知

若数域$\mathbb{F}$上的线性变换或$\mathbb{F}$上的矩阵在$\mathbb{F}$中有一个特征值，则在$\mathbb{F}$上的线性空间或$\mathbb{F}$上的列向量空间中必存在对应的特征向量.

任取线性变换φ的一个特征值$\lambda_0\in\mathbb{F}$，设V_0是特征值λ_0的特征子空间，则

$\forall\alpha\in V_0$，有$\varphi\psi(\boldsymbol{\alpha})=\psi\varphi(\boldsymbol{\alpha})=\psi\left(\lambda_0\boldsymbol{\alpha}\right)=\lambda_0\psi(\boldsymbol{\alpha})\Longrightarrow\boldsymbol{\psi}(\boldsymbol{\alpha})\in V_0$，因此$V_0$是$\boldsymbol{\psi}-$不变子空间.

取V_0的一组基并扩张为V的一组基，则ψ在这组基下的表示矩阵为分块对角矩阵$\begin{pmatrix}A & C\\ O & B\end{pmatrix}$，其中$A$是 $\psi|_{V_0}$ 在给定基下的表示矩阵

于是$|\lambda I_V-\psi|=|\lambda I-\boldsymbol{A}||\lambda\boldsymbol{I}-\boldsymbol{B}|$.因为$\boldsymbol{\psi}$的特征值都在$\mathbb{F}$中，故$\boldsymbol{A}$的特征值都在$\mathbb{F}$中，于是 $\psi|_{V_0}$ 的特征值都在$\mathbb{F}$中.

任取 $\psi|_{V_0}$ 的一个特征值$\mu_0\in\mathbb{F}$及其特征向量$\boldsymbol{\alpha}\in V_0$，则$\boldsymbol{\varphi}(\boldsymbol{\alpha})=\lambda_0\boldsymbol{\alpha},\boldsymbol{\psi}(\boldsymbol{\alpha})=\mu_0\boldsymbol{\alpha}$，于是$\boldsymbol{\alpha}$就是$\boldsymbol{\varphi},\boldsymbol{\psi}$的公共特征向量.

4.对阶数进行归纳.当$n=1$时结论显然成立，设对$n-1$阶矩阵结论成立，现对n阶矩阵进行证明.

因为$\boldsymbol{AB}=\boldsymbol{BA}$且$\boldsymbol{A},\boldsymbol{B}$的特征值都在$\mathbb{F}$中，故可知，$\boldsymbol{A},\boldsymbol{B}$有公共的特征向量$\boldsymbol{e}_1\in\mathbb{F}^n$

不妨设$\boldsymbol{Ae}_1=\lambda_1\boldsymbol{e}_1$， $\boldsymbol{Be}_1=\mu_1\boldsymbol{e}_1$其中$\lambda_1,\mu_1\in\mathbb{F}$分别是$\boldsymbol{A},\boldsymbol{B}$的特征值.由基扩张定理，可将$\boldsymbol{e}_1$扩张为$\mathbb{F}^n$的一组基$\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\}$

令$\boldsymbol{P}=(\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n)$，则$\boldsymbol{P}$是$\mathbb{F}$上的$n$阶可逆矩阵，且有

$\boldsymbol{P}^{-1}\boldsymbol{AP}=\begin{pmatrix}\lambda_1 & *\\ \boldsymbol{O} & \boldsymbol{A}_1\end{pmatrix}$， $\boldsymbol{P}^{-1}\boldsymbol{BP}=\begin{pmatrix}\mu_1 & *\\ \boldsymbol{O} & \boldsymbol{B}_1\end{pmatrix}$ 其中$\boldsymbol{A}_1,\boldsymbol{B}_1$是$\mathbb{F}$上的$n-1$阶矩阵

从$\boldsymbol{AB}=\boldsymbol{BA}$不难推出$\boldsymbol{A}_1\boldsymbol{B}_1=\boldsymbol{B}_1\boldsymbol{A}_1$，又容易验证$\boldsymbol{A}_1,\boldsymbol{B}_1$的特征值都在$\mathbb{F}$中

故由归纳假设，存在$\mathbb{F}$上的$n-1$阶可逆矩阵Q，使得$Q^{-1}A_1Q$和$Q^{-1}B_1Q$都是上三角矩阵

令$R=P\begin{pmatrix}1 & O\\ O & Q\end{pmatrix}$则$\boldsymbol{R}$是$\mathbb{F}$上的$n$阶可逆矩阵，且

$$\boldsymbol{R}^{-1}\boldsymbol{A}\boldsymbol{R}=\begin{pmatrix}1 & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{Q}\end{pmatrix}^{-1}\begin{pmatrix}\lambda_1 & *\\ O & A_1\end{pmatrix}\begin{pmatrix}1 & \boldsymbol{O}\\ O & Q\end{pmatrix}=\begin{pmatrix}\lambda_1 & *\\ O & Q^{-1}A_1Q\end{pmatrix}$$

$$\boldsymbol{R}^{-1}\boldsymbol{B}\boldsymbol{R}=\begin{pmatrix}1 & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{Q}\end{pmatrix}^{-1}\begin{pmatrix}\mu_1 & *\\ \boldsymbol{O} & B_1\end{pmatrix}\begin{pmatrix}1 & \boldsymbol{O}\\ O & Q\end{pmatrix}=\begin{pmatrix}\mu_1 & *\\ O & Q^{-1}B_1Q\end{pmatrix}$$

都是上三角矩阵.

Proof 5.证明对空间维数进行归纳.当$n=1$时结论显然成立,设对维数小于n的线性空间结论成立,现对n维线性空间进行证明

设φ的全体不同特征值为$\lambda_1,\cdots,\lambda_s\in\mathbb{F}$，对应的特征子空间分别为$V_1,\cdots,V_s$，则由$\varphi$可对角化可知$V=V_1\oplus\cdots\oplus V_s$

若$s=1$,则$\boldsymbol{\varphi}=\lambda_1\boldsymbol{I}_V$为纯量变换,此时只要取$V$的一组基,使得$\boldsymbol{\psi}$在这组基下的表示矩阵为对角矩阵

则$\boldsymbol{\varphi}$在这组基下的表示矩阵为$\lambda_1\boldsymbol{I}_n$,结论成立.

若$s>1$,则 $\dim V_i<n$.注意到$\boldsymbol{\varphi\psi}=\boldsymbol{\psi\varphi}$且$\boldsymbol{\varphi},\boldsymbol{\psi}$的特征值都在$\mathbb{F}$中,可知$V_i$都是$\boldsymbol{\psi}-$不变子空间.

考虑线性变换的限制 $\varphi|_{V_i},\psi|_{V_i}$：它们乘法可交换,且由可对角化线性变换的性质可知它们都可对角化

故由归纳假设可知,$\varphi|_{V_i},\psi|_{V_i}$ 可同时对角化

即存在V_i的一组基，使得 $\boldsymbol{\varphi}|_{V_i},\boldsymbol{\psi}|_{V_i}$ 在这组基下的表示矩阵都是对角矩阵

将V_i的基拼成V的一组基,则φ,ψ在这组基下的表示矩阵都是对角矩阵,即φ,ψ可同时对角化.

代数版本证明：

由于A可对角化,所以存在可逆矩阵P使得

$P^{-1}AP=\operatorname{diag}\{\lambda_1E_1,\lambda_2E_2,\cdots,\lambda_sE_s\}$,其中$\lambda_1,\lambda_2,\cdots,\lambda_s$是$A$的所有互异特征值,$E_1,E_2,\cdots,E_s$分别是$r_1,r_2,\cdots,r_s$级单位矩阵,$r_1+r_2+\cdots+r_s=n$.

由$AB=BA$可得$P^{-1}APP^{-1}BP=P^{-1}BPP^{-1}AP$

可得$P^{-1}BP=\operatorname{diag}\{B_1,B_2,\cdots,B_s\}$，其中$B_1,B_2,\cdots,B_s$分别为$r_1,r_2,\cdots,r_s$级方阵

由于B可对角化,所以$P^{-1}BP$也可以对角化,从而$B_1,B_2,\cdots,B_s$都可以对角化

所以对任意的$i=1,2,\cdots,s$,存在可逆矩阵Q_i使得$Q_i^{-1}B_iQ_i$为对角矩阵

取$Q=\operatorname{diag}\{Q_1,Q_2,\cdots,Q_s\}$,则$Q^{-1}P^{-1}BPQ=\operatorname{diag}\left\{Q_1^{-1}B_1Q_1,Q_2^{-1}B_2Q_2,\cdots,Q_s^{-1}B_sQ_s\right\}$为对角矩阵,同时

$Q^{-1}P^{-1}APQ=\operatorname{diag}\left\{\lambda_1Q_1^{-1}Q_1,\lambda_2Q_2^{-1}Q_2,\cdots,\lambda_sQ_s^{-1}Q_s\right\}=\{\lambda_1E_1,\lambda_2E_2,\cdots,\lambda_sE_s\}$仍为对角矩阵.

所以取$T=PQ$就有$T^{-1}AT,T^{-1}BT$同时为对角矩阵.

注释：B可对角化$B=\begin{pmatrix}B_1 & & & \\ & B_2 & & \\ & & \cdots & \\ & & & B_s\end{pmatrix}$ B可对角化说明$m(x)$作为B的最小多项式无重根

且$m(B)=\operatorname{diag}\{m(B_1)\cdots m(B_s)\}=O\Longrightarrow m(B_i)=O$故$B_i$的最小多项式都无重根故都可对角化

Theorem 7.11.6 (分块上三角矩阵对角化, 全空间可对角化在不变子空间上限制仍然可对角化)

1.设m阶矩阵$\boldsymbol{A}$与n阶矩阵$\boldsymbol{B}$没有公共的特征值, 且$\boldsymbol{A},\boldsymbol{B}$均可对角化, 又$\boldsymbol{C}$为$m\times n$矩阵

求证 : $M=\begin{pmatrix} A & C \\ O & B \end{pmatrix}$也可对角化.

2.设$\boldsymbol{A}$为m阶矩阵, $\boldsymbol{B}$为n阶矩阵, $\boldsymbol{C}$为$m\times n$矩阵, $\boldsymbol{M}=\begin{pmatrix} \boldsymbol{A} & \boldsymbol{C} \\ \boldsymbol{O} & \boldsymbol{B} \end{pmatrix}$

求证 : 若$\boldsymbol{M}$可对角化, 则$A,\boldsymbol{B}$均可对角化.证明任取$\boldsymbol{M}$的特征值λ_0

Proof 1.任取$\boldsymbol{A}$的特征值λ_0, 记其代数重数为$m_{\boldsymbol{A}}(\lambda_0)$, 几何重数为$t_{\boldsymbol{A}}(\lambda_0)$

首先注意到$\boldsymbol{A},\boldsymbol{B}$没有公共的特征值, 故$\lambda_0$不是$\boldsymbol{B}$的特征值, 又$|\lambda\boldsymbol{I}-\boldsymbol{M}|=|\lambda\boldsymbol{I}-\boldsymbol{A}||\lambda\boldsymbol{I}-\boldsymbol{B}|$, 从而$m_{\boldsymbol{M}}(\lambda_0)=m_{\boldsymbol{A}}(\lambda_0)$.

由于$\lambda_0\boldsymbol{I}-\boldsymbol{B}$是非异阵, 故有如下分块矩阵的初等变换:

$$\lambda_0\boldsymbol{I}-\boldsymbol{M}=\begin{pmatrix} \lambda_0\boldsymbol{I}-\boldsymbol{A} & -\boldsymbol{C} \\ \boldsymbol{O} & \lambda_0\boldsymbol{I}-\boldsymbol{B} \end{pmatrix}\to\begin{pmatrix} \lambda_0\boldsymbol{I}-\boldsymbol{A} & \boldsymbol{O} \\ \boldsymbol{O} & \lambda_0\boldsymbol{I}-\boldsymbol{B} \end{pmatrix}$$

因为矩阵的秩在分块初等变换下不变, 故由矩阵秩的等式可得

$\mathrm{r}(\lambda_0\boldsymbol{I}-\boldsymbol{M})=\mathrm{r}(\lambda_0\boldsymbol{I}-\boldsymbol{A})+\mathrm{r}(\lambda_0\boldsymbol{I}-\boldsymbol{B})=\mathrm{r}(\lambda_0\boldsymbol{I}-\boldsymbol{A})+n$于是

$t_{\boldsymbol{M}}(\lambda_0)=(m+n)-\mathrm{r}(\lambda_0\boldsymbol{I}-\boldsymbol{M})=m-\mathrm{r}(\lambda_0\boldsymbol{I}-\boldsymbol{A})=t_{\boldsymbol{A}}(\lambda_0)$.

因为$\boldsymbol{A}$可对角化， 所以$\boldsymbol{A}$有完全的特征向量系, 从而$m_{\boldsymbol{A}}(\lambda_0)=t_{\boldsymbol{A}}(\lambda_0)$, 于是$m_{\boldsymbol{M}}(\lambda_0)=t_{\boldsymbol{M}}(\lambda_0)$

同理可证, 对$\boldsymbol{B}$的任一特征值μ_0, 成立$m_{\boldsymbol{M}}(\mu_0)=t_{\boldsymbol{M}}(\mu_0)$。因此$\boldsymbol{M}$有完全的特征向量系, 从而可对角化.

2.由$|\lambda\boldsymbol{I}-\boldsymbol{M}|=|\lambda\boldsymbol{I}-\boldsymbol{A}||\lambda\boldsymbol{I}-\boldsymbol{B}|$可得$m_{\boldsymbol{M}}(\lambda_0)=m_{\boldsymbol{A}}(\lambda_0)+m_{\boldsymbol{B}}(\lambda_0)$.

考虑如下分块矩阵 : $\lambda_0\boldsymbol{I}-\boldsymbol{M}=\begin{pmatrix} \lambda_0\boldsymbol{I}-\boldsymbol{A} & -\boldsymbol{C} \\ \boldsymbol{O} & \lambda_0\boldsymbol{I}-\boldsymbol{B} \end{pmatrix}$

由矩阵秩的不等式可得 $\mathrm{r}(\lambda_0\boldsymbol{I}-\boldsymbol{M})\geq\mathrm{r}(\lambda_0\boldsymbol{I}-\boldsymbol{A})+\mathrm{r}(\lambda_0\boldsymbol{I}-\boldsymbol{B})$ 于是

$t_{\boldsymbol{M}}(\lambda_0)=(m+n)-\mathrm{r}(\lambda_0\boldsymbol{I}-\boldsymbol{M})\leq(m-\mathrm{r}(\lambda_0\boldsymbol{I}-\boldsymbol{A}))+(n-\mathrm{r}(\lambda_0\boldsymbol{I}-\boldsymbol{B}))=t_{\boldsymbol{A}}(\lambda_0)+t_{\boldsymbol{B}}(\lambda_0)$.

由于几何重数总是小于等于代数重数, 故有$t_{\boldsymbol{M}}(\lambda_0)\leq t_{\boldsymbol{A}}(\lambda_0)+t_{\boldsymbol{B}}(\lambda_0)\leq m_{\boldsymbol{A}}(\lambda_0)+m_{\boldsymbol{B}}(\lambda_0)=m_{\boldsymbol{M}}(\lambda_0)$

因为$\boldsymbol{M}$可对角化, 所以$\boldsymbol{M}$有完全的特征向量系, 从而$t_{\boldsymbol{M}}(\lambda_0)=m_{\boldsymbol{M}}(\lambda_0)$

再由上述不等式可得$t_{\boldsymbol{A}}(\lambda_0)=m_{\boldsymbol{A}}(\lambda_0)$, $t_{\boldsymbol{B}}(\lambda_0)=m_{\boldsymbol{B}}(\lambda_0)$.由$\lambda_0$的任意性即知, $\boldsymbol{A},\boldsymbol{B}$均有完全的特征向量系, 从而均可对角化.

实际上我们利用最小多项式也能快速得到答案

Proposition 7.8 (矩阵的 Kronecker 乘积)

1.设$\boldsymbol{A},\boldsymbol{B}$分别是$m,n$阶矩阵, $\boldsymbol{A}$的特征值为$\lambda_i(1\leq i\leq m)$, $\boldsymbol{B}$的特征值为$\mu_j(1\leq j\leq n)$

求证 : $\boldsymbol{A}\otimes\boldsymbol{B}$的特征值为$\lambda_i\mu_j(1\leq i\leq m;1\leq j\leq n)$.

Proof 可知, 存在m阶可逆矩阵$\boldsymbol{P}$以及n阶可逆矩阵$\boldsymbol{Q}$, 使得

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\begin{pmatrix} \lambda_1 & * & * & * \\ & \lambda_2 & * & * \\ & & \ddots & \vdots \\ & & & \lambda_m \end{pmatrix},\quad \boldsymbol{Q}^{-1}\boldsymbol{B}\boldsymbol{Q}=\begin{pmatrix} \mu_1 & * & * & * \\ & \mu_2 & * & * \\ & & \ddots & \vdots \\ & & & \mu_n \end{pmatrix}$$

容易验证上三角矩阵的$Kronecker$积仍是上三角矩阵

$\Longrightarrow(\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P})\otimes(\boldsymbol{Q}^{-1}\boldsymbol{B}\boldsymbol{Q})$ 的主对角元素依次为$\lambda_1\mu_1,\cdots,\lambda_1\mu_n,\lambda_2\mu_1,\cdots,\lambda_2\mu_n,\cdots,\lambda_m\mu_1\cdots,\lambda_m\mu_n$

注意到 $(\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P})\otimes(\boldsymbol{Q}^{-1}\boldsymbol{B}\boldsymbol{Q})=(\boldsymbol{P}\otimes\boldsymbol{Q})^{-1}(\boldsymbol{A}\otimes\boldsymbol{B})(\boldsymbol{P}\otimes\boldsymbol{Q})$, 故结论得证.

第 8 章 二次型

8.1 二次型基础定义

Definition 8.1

设f是数域$\mathbb{K}$上的n元二次齐次多项式：

$f(x_1,x_2,\cdots,x_n)=a_{11}x_1^2+2a_{12}x_1x_2+\cdots+2a_{1n}x_1x_n+a_{22}x_2^2+\cdots+2a_{2n}x_2x_n+\cdots+a_{nn}x_n^2$,

称f为数域$\mathbb{K}$上的n元二次型, 简称二次型.

笔记 我们现在要用矩阵为工具来处理二次型.用矩阵的乘法我们可以把式写成矩阵相乘的形式：

$f(x_1,x_2,\cdots,x_n)=\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}$, 其中$\boldsymbol{A}=\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&&\vdots\\a_{n1}&a_{n2}&\cdots&a_{nn}\end{pmatrix}$,$\boldsymbol{x}=\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}$.

在矩阵$\boldsymbol{A}$中,$a_{ij}=a_{ji}$对一切i,j成立, 也就是说矩阵$\boldsymbol{A}$是一个对称阵.

由此可知, 给定一个$\mathbb{K}$上的n元二次型, 我们就得到了一个$\mathbb{K}$上的n阶对称阵$\boldsymbol{A}$(称为该二次型的相伴矩阵或系数矩阵).

反过来, 若给定数域$\mathbb{K}$上的n阶对称阵$\boldsymbol{A}$, 则由式, 我们可以得到一个$\mathbb{K}$上二次型, 称为对称阵$\boldsymbol{A}$的相伴二次型.

现在的问题是：用这样的方法得到的矩阵是否和二次型一一对应?回答这一点并不难.

设$f=\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}=\boldsymbol{x}'\boldsymbol{B}\boldsymbol{x}$, 我们要证明$\boldsymbol{A}=\boldsymbol{B}$.

这等价于证明下面的结论：设$\boldsymbol{A}$是对称阵, 若$\alpha'A\alpha=0$对一切α成立, 则$A=O$.

设$\boldsymbol{A}=(a_{ij})$是n阶对称阵, 则$a_{ii}=\boldsymbol{e}_i'\boldsymbol{A}\boldsymbol{e}_i=0$且

$0=(e_i+e_j)'\boldsymbol{A}(e_i+e_j)=e_i'\boldsymbol{A}e_i+\boldsymbol{e}_j'\boldsymbol{A}e_j+\boldsymbol{e}_i'\boldsymbol{A}e_j+e_j'\boldsymbol{A}e_i=a_{ij}+a_{ji}$.

因为$a_{ij}=a_{ji}$, 故$a_{ij}=0$, 于是$\boldsymbol{A}=\boldsymbol{O}$.这表明用对称阵来表示二次型时, 表示矩阵是唯一的.

笔记 二次型理论的基本问题是要寻找一个线性变换把它变成只含平方项.

由上面我们知道, 二次型与对称阵一一对应, 而线性变换可以用矩阵来表示.

自然地, 二次型的变换与矩阵有着密切的关系.现在我们来探讨这个关系.

设V是n维线性空间, 二次型$f(x_1,x_2,\cdots,x_n)$可以看成是V上的二次函数.

即若设V的一组基为$\{e_1,e_2,\cdots,e_n\}$, 向量x在这组基下的坐标为$x_1,x_2,\cdots,x_n$, 则f便是向量$\boldsymbol{x}$的函数.

现假设$\{\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n\}$是$V$的另一组基, 向量$\boldsymbol{x}$在$\{\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n\}$下的坐标为$y_1,y_2,\cdots,y_n$.

记$\boldsymbol{C}=(c_{ij})$是从基$\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\}$到基$\{\boldsymbol{f}_1,\boldsymbol{f}_2,\cdots,\boldsymbol{f}_n\}$的过渡矩阵, 则$\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}=\begin{pmatrix}c_{11}&c_{12}&\cdots&c_{1n}\\c_{21}&c_{22}&\cdots&c_{2n}\\\vdots&\vdots&&\vdots\\c_{n1}&c_{n2}&\cdots&c_{nn}\end{pmatrix}\begin{pmatrix}y_1\\y_2\\\vdots\\y_n\end{pmatrix}$,

或简记为$\boldsymbol{x}=\boldsymbol{C}\boldsymbol{y}$, 其中$y$为$n$维列向量$\begin{pmatrix}y_1\\y_2\\\vdots\\y_n\end{pmatrix}$.

将上式代入$f(x_1,x_2,\cdots,x_n)=\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}$, 得$f(x_1,x_2,\cdots,x_n)=\boldsymbol{y}\boldsymbol{C}'\boldsymbol{A}\boldsymbol{C}\boldsymbol{y}$.

显然,$\boldsymbol{C}'\boldsymbol{A}\boldsymbol{C}$仍是一个对称阵, 故$\boldsymbol{y}'\boldsymbol{C}'\boldsymbol{A}\boldsymbol{C}\boldsymbol{y}$是以$\boldsymbol{y}$为变元的二次型, 记为$g(y_1,y_2,\cdots,y_n)$.

由此我们可看出：若二次型$f(x_1,x_2,\cdots,x_n)$所对应的对称阵为$\boldsymbol{A}$

则经过变量代换之后得到的二次型$g(y_1,y_2,\cdots,y_n)$所对应的对称阵为$C'AC$.

Definition 8.2 (合同)

设A,B是数域$\mathbb{K}$上的n阶矩阵,若存在n阶非异阵C,使$B=C'AC$,则称$\boldsymbol{B}$与$\boldsymbol{A}$是合同的,或称$\boldsymbol{B}$与$\boldsymbol{A}$具有合同关系.

不难证明,合同关系是一个等价关系,即

(1)任一矩阵$\boldsymbol{A}$与自己合同,因为$\boldsymbol{A}=\boldsymbol{I}'\boldsymbol{A}\boldsymbol{I}$;

(2)若$\boldsymbol{B}$与$\boldsymbol{A}$合同,则$\boldsymbol{A}$与$\boldsymbol{B}$合同.这是因为若$\boldsymbol{B}=\boldsymbol{C}'\boldsymbol{A}\boldsymbol{C}$,则$\boldsymbol{A}=(C')^{-1}BC^{-1}=\left(C^{-1}\right)'BC^{-1}$;

(3)若$\boldsymbol{B}$与$\boldsymbol{A}$合同,$\boldsymbol{D}$与$\boldsymbol{B}$合同,则$\boldsymbol{D}$与$\boldsymbol{A}$合同.事实上,若$\boldsymbol{B}=\boldsymbol{C}'\boldsymbol{A}\boldsymbol{C},\boldsymbol{D}=\boldsymbol{H}'\boldsymbol{B}\boldsymbol{H}$,则$\boldsymbol{D}=\boldsymbol{H}'\boldsymbol{C}'\boldsymbol{A}\boldsymbol{C}\boldsymbol{H}=(\boldsymbol{C}\boldsymbol{H})'A(\boldsymbol{C}\boldsymbol{H})$.

Definition 8.3 (合同变换)

对称阵$\boldsymbol{A}$的下列变换都是合同变换:

(1)对换$\boldsymbol{A}$的第i行与第j行,再对换第i列与第j列;

(2)将非零常数k乘以$\boldsymbol{A}$的第i行,再将k乘以第i列;

(3)将$\boldsymbol{A}$的第i行乘以k加到第j行上,再将第i列乘以k加到第j列上.

上述变换相当于将一个初等矩阵左乘以$\boldsymbol{A}$后再将这个初等矩阵的转置右乘之,因此是合同变换.

Lemma 8.1

设$\boldsymbol{A}$是数域$\mathbb{K}$上的非零对称阵,则必存在非异阵C,使$C'AC$的第$(1,1)$元素不等于零.

Proof 若$a_{11}=0$,而$a_{ii}\neq 0$,则将$\boldsymbol{A}$的第一行与第i行对换,再将第一列与第i列对换,得到的矩阵的第$(1,1)$元素不为零.
根据上述引理,这样得到的矩阵和原矩阵合同.
若所有的$a_{ii}=0(i=1,2,\cdots,n)$.设$a_{ji}\neq 0(i\neq j)$,将$\boldsymbol{A}$的第j行加到第i行上,再将第j列加到第i列上.
因为$\boldsymbol{A}$是对称阵,$a_{ji}=a_{ij}\neq 0$,于是第(i,i)元素是$2a_{ij}$且不为零.
再用前面的办法使第$(1,1)$元素不等于零.显然我们得到的矩阵和原矩阵仍合同.这就证明了结论.

Lemma 8.2

设$\boldsymbol{A}$是数域$\mathbb{K}$上的n阶对称阵,则必存在$\mathbb{K}$上的n阶非异阵C,使$C'AC$为对角阵.

Proof 由上述引理
不妨设$\boldsymbol{A}=\left(a_{ij}\right)$中$a_{11}\neq 0$.若$a_{i1}\neq 0$,则可将第一行乘以$-a_{11}^{-1}a_{i1}$加到第$i$行上,再将第一列乘以$-a_{11}^{-1}a_{i1}$加到第$i$列上.
由于$a_{i1}=a_{1i}$,故得到的矩阵的第$(1,i)$元素及第$(i,1)$元素均等于零.
由引理可知,新得到的矩阵与$\boldsymbol{A}$是合同的.这样,可依次把$\boldsymbol{A}$的第一行与第一列除a_{11}外的元素都消去.
于是$\boldsymbol{A}$合同于下列矩阵:

$$\begin{pmatrix} a_{11} & 0 & 0 & \cdots & 0 \\ 0 & b_{22} & b_{23} & \cdots & b_{2n} \\ 0 & b_{32} & b_{33} & \cdots & b_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & b_{n2} & b_{n3} & \cdots & b_{nn} \end{pmatrix}.$$

上式右下角是一个$n-1$阶对称阵,记为$\boldsymbol{A}_1$.
因此可归纳地假设存在非异的$n-1$阶矩阵D,使$D'A_1D$为对角阵,于是

$$\begin{pmatrix} 1 & O \\ O & D' \end{pmatrix}\begin{pmatrix} a_{11} & O \\ O & A_1 \end{pmatrix}\begin{pmatrix} 1 & O \\ O & D \end{pmatrix}=\begin{pmatrix} a_{11} & O \\ O & D'A_1D \end{pmatrix}$$

是一个对角阵.显然$\begin{pmatrix} 1 & O \\ O & D' \end{pmatrix}=\begin{pmatrix} 1 & O \\ O & D \end{pmatrix}'$,因此$\boldsymbol{A}$合同于一对角阵.

8.2 惯性定理与合同标准型

Theorem 8.2.1 (合同标准型)

任意一个实对称阵$\boldsymbol{A}$必合同于一个对角阵 : $\mathrm{diag}\{1,\cdots,1;-1,\cdots,-1;0,\cdots,0\}$

Proof 我们从上节的引理可以知道$\boldsymbol{C}'\boldsymbol{A}\boldsymbol{C}=\mathrm{diag}\{d_1,d_2,\cdots,d_r,0,\cdots,0\}$,其中$d_i\neq 0(i=1,\cdots,r)$.
注意到$\boldsymbol{C}$是可逆阵,所以$r=\mathrm{r}(\boldsymbol{C}'\boldsymbol{A}\boldsymbol{C})=\mathrm{r}(\boldsymbol{A})$.故秩$r$是矩阵合同关系下的一个不变量.
如同相似标准型一样,我们的目的是要找实对称阵在合同关系下的全系不变量,即找出一组足以判断两个实对称阵是否合同的合同不
我们不妨设实对称阵已具有下列对角阵的形状 :
$\boldsymbol{A}=\mathrm{diag}\{d_1,d_2,\cdots,d_r,0,\cdots,0\}$.
由引理不难知道,任意调换$\boldsymbol{A}$的主对角线上的元素得到的矩阵仍与$\boldsymbol{A}$合同.
因此我们可把零放在一起,把正项与负项放在一起,即可设$d_1>0,\cdots,d_p>0;d_{p+1}<0,\cdots,d_r<0$.
$\boldsymbol{A}$所代表的二次型为$f(x_1,x_2,\cdots,x_n)=d_1x_1^2+d_2x_2^2+\cdots+d_rx_r^2$.
令$y_1=\sqrt{d_1}x_1,\cdots,y_p=\sqrt{d_p}x_p;y_{p+1}=\sqrt{-d_{p+1}}x_{p+1},\cdots,y_r=\sqrt{-d_r}x_r;y_j=x_j(j=r+1,\cdots,n)$,
则式变为$f=y_1^2+\cdots+y_p^2-y_{p+1}^2-\cdots-y_r^2$.
这一事实等价于说$\boldsymbol{A}$合同于下列对角阵 : $\mathrm{diag}\{1,\cdots,1;-1,\cdots,-1;0,\cdots,0\}$,其中有$p$个1,$q$个$-1$,$n-r$个零.

Theorem 8.2.2 (惯性定理)

设$f(x_1,x_2,\cdots,x_n)$是一个n元实二次型,且f可化为两个标准型 :
$c_1y_1^2+\cdots+c_py_p^2-c_{p+1}y_{p+1}^2-\cdots-c_ry_r^2$
$d_1z_1^2+\cdots+d_kz_k^2-d_{k+1}z_{k+1}^2-\cdots-d_rz_r^2$
其中$c_i>0,d_i>0$,则必有$p=k$.

Proof 用反证法,设$p>k$.由前面的说明知道可设c_i及d_i均为1.
因此$y_1^2+\cdots+y_p^2-y_{p+1}^2-\cdots-y_r^2=z_1^2+\cdots+z_k^2-z_{k+1}^2-\cdots-z_r^2$.
又设$\boldsymbol{x}=\boldsymbol{B}\boldsymbol{y},\boldsymbol{x}=\boldsymbol{C}\boldsymbol{z}$

其中$\boldsymbol{x}=\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix},\boldsymbol{y}=\begin{pmatrix}y_1\\y_2\\\vdots\\y_n\end{pmatrix},\boldsymbol{z}=\begin{pmatrix}z_1\\z_2\\\vdots\\z_n\end{pmatrix}$.

于是$\boldsymbol{z}=\boldsymbol{C}^{-1}\boldsymbol{B}\boldsymbol{y}$.

令$\boldsymbol{C}^{-1}\boldsymbol{B}=\begin{pmatrix}c_{11}&c_{12}&\cdots&c_{1n}\\c_{21}&c_{22}&\cdots&c_{2n}\\\vdots&\vdots&&\vdots\\c_{n1}&c_{n2}&\cdots&c_{nn}\end{pmatrix}$

$$\text{则}\begin{cases}z_1=c_{11}y_1+c_{12}y_2+\cdots+c_{1n}y_n,\\z_2=c_{21}y_1+c_{22}y_2+\cdots+c_{2n}y_n,\\\qquad\cdots\cdots\cdots\\z_n=c_{n1}y_1+c_{n2}y_2+\cdots+c_{nn}y_n\end{cases}$$

因为$p>k$,齐次线性方程组
$$\begin{cases} c_{11}y_1+c_{12}y_2+\cdots+c_{1n}y_n=0 \\ \cdots\cdots\cdots\cdots \\ c_{k1}y_1+c_{k2}y_2+\cdots+c_{kn}y_n=0 \\ y_{p+1}=0 \\ \cdots\cdots\cdots\cdots \\ y_n=0. \end{cases}$$
必有非零解 (n个未知数,$n-(p-k)$个方程式).

令其中一个非零解为$y_1=a_1,\cdots,y_p=a_p,y_{p+1}=0,\cdots,y_n=0$.

把这组解代入$y_1^2+\cdots+y_p^2-y_{p+1}^2-\cdots-y_r^2=z_1^2+\cdots+z_k^2-z_{k+1}^2-\cdots-z_r^2$.式

左边得到$a_1^2+\cdots+a_p^2>0$但这时$z_1=\cdots=z_k=0$,故右边将小于等于零.引出了矛盾.同理可证$p<k$也不可能.

Definition 8.4 (正负惯性指数与符号差)

设$f(x_1,x_2,\cdots,x_n)$是一个实二次型,若它能化为形如实合同标准型的形状,则称

r是该二次型的秩,p是它的正惯性指数,$q=r-p$是它的负惯性指数,$s=p-q$称为f的符号差.

显然,若已知秩r与符号差s,则$p=\frac{1}{2}(r+s),q=\frac{1}{2}(r-s)$.

事实上,在p,q,r,s中只需知道其中两个数,其余两个数也就知道了.

由于实对称阵与实二次型之间的等价关系

我们将实二次型的秩、惯性指数及符号差也称为相应的实对称阵的秩、惯性指数及符号差.

Theorem 8.2.3

秩与符号差 (或正负惯性指数) 是实对称阵在合同关系下的全系不变量.

Proof 由上面的定理知道,秩r与符号差s是实对称阵合同关系的不变量.

反之,若n阶实对称阵$\boldsymbol{A},\boldsymbol{B}$的秩都为$r$,符号差都是$s$,则它们都合同于$\mathrm{diag}\{1,\cdots,1;-1,\cdots,-1;0,\cdots,0\}$,

其中有$p=\dfrac{1}{2}(r+s)$个1,$q=\dfrac{1}{2}(r-s)$个-1及$n-r$个零,因此$\boldsymbol{A}$与$\boldsymbol{B}$合同.对正负惯性指数的结论也同样成立.

Theorem 8.2.4 (复二次型的标准型)

因为复二次型$f(x_1,x_2,\cdots,x_n)=d_1x_1^2+d_2x_2^2+\cdots+d_rx_r^2$必可化为$z_1^2+z_2^2+\cdots+z_r^2$

其中$z_i=\sqrt{d_i}x_i(i=1,2,\cdots,r),z_j=x_j(j=r+1,\cdots,n)$.所以复对称阵的合同关系只有一个全系不变量,那就是秩r.

8.3 正定型与正定矩阵

Definition 8.5

设$f(x_1,x_2,\cdots,x_n)=\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}$是$n$元实二次型.

(1)若对任意n维非零列向量$\boldsymbol{\alpha}$均有$\boldsymbol{\alpha}'\boldsymbol{A}\boldsymbol{\alpha}>0$, 则称$f$是正定二次型(简称正定型), 矩阵$\boldsymbol{A}$称为正定矩阵(简称正定阵);

(2)若对任意n维非零列向量$\boldsymbol{\alpha}$均有$\boldsymbol{\alpha}'\boldsymbol{A}\boldsymbol{\alpha}<0$, 则称$f$是负定二次型(简称负定型), 矩阵$\boldsymbol{A}$称为负定矩阵(简称负定阵);

(3)若对任意n维非零列向量$\boldsymbol{\alpha}$均有$\boldsymbol{\alpha}'\boldsymbol{A}\boldsymbol{\alpha}\geq 0$, 则称$f$是半正定二次型(简称半正定型), 矩阵$\boldsymbol{A}$称为半正定矩阵(简称半正定阵);

(4)若对任意n维非零列向量$\boldsymbol{\alpha}$均有$\boldsymbol{\alpha}'\boldsymbol{A}\boldsymbol{\alpha}\leq 0$, 则称$f$是半负定二次型(简称半负定型), 矩阵$\boldsymbol{A}$称为半负定矩阵(简称半负定阵);

(5)若存在$\boldsymbol{\alpha}$, 使$\boldsymbol{\alpha}'\boldsymbol{A}\boldsymbol{\alpha}>0$; 又存在$\boldsymbol{\beta}$, 使$\boldsymbol{\beta}'\boldsymbol{A}\boldsymbol{\beta}<0$, 则称$f$是不定型.

Theorem 8.3.1

实二次型$f(x_1,x_2,\cdots,x_n)$是正定型的充分必要条件是f的正惯性指数等于n;

$f(x_1,x_2,\cdots,x_n)$是负定型的充分必要条件是f的负惯性指数等于n;

$f(x_1,x_2,\cdots,x_n)$是半正定型的充分必要条件是f的正惯性指数等于f的秩r;

$f(x_1,x_2,\cdots,x_n)$是半负定型的充分必要条件是f的负惯性指数等于f的秩r

Proof 若f的正惯性指数等于n, 则f可化为下列标准型：$f=y_1^2+y_2^2+\cdots+y_n^2$.显然$f$是正定型.
反之, 若f是正定型, 如果f的正惯性指数$p<n$
则f可化为如下标准型：$f=y_1^2+\cdots+y_p^2-c_{p+1}y_{p+1}^2-\cdots-c_ny_n^2$, 其中$c_j\geq 0(j=p+1,\cdots,n)$.
这时令$b_1=\cdots=b_p=0, b_{p+1}=\cdots=b_n=1$, 则$b_1,b_2,\cdots,b_n$不全为零.

假设这时$\boldsymbol{x}=\boldsymbol{C}\boldsymbol{y}$, 其中$\boldsymbol{x}=\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}$, $\boldsymbol{y}=\begin{pmatrix}y_1\\y_2\\\vdots\\y_n\end{pmatrix}$, C是非异阵.

故从$y_i=b_i(i=1,\cdots,n)$可得$x_i=a_i(i=1,\cdots,n)$是一组不全为零的实数.
于是$f(a_1,a_2,\cdots,a_n)\leq 0$.这与$f$是正定型矛盾, 其余结论的证明类似.

Corollary 8.1

实对称阵$\boldsymbol{A}$是正定阵当且仅当它合同于单位阵$\boldsymbol{I}_n$

$\boldsymbol{A}$是负定阵当且仅当它合同于$-\boldsymbol{I}_n$

$\boldsymbol{A}$是半正定阵当且仅当$\boldsymbol{A}$合同于下列对角阵：$\begin{pmatrix}\boldsymbol{I}_r&\boldsymbol{O}\\\boldsymbol{O}&\boldsymbol{O}\end{pmatrix}$

$\boldsymbol{A}$是半负定阵当且仅当$\boldsymbol{A}$合同于下列对角阵：$\begin{pmatrix}-\boldsymbol{I}_r&\boldsymbol{O}\\\boldsymbol{O}&\boldsymbol{O}\end{pmatrix}$.

Definition 8.6

设$\boldsymbol{A}=(a_{ij})$是n阶矩阵, $\boldsymbol{A}$的n个子式：$\begin{vmatrix}a_{11}&a_{12}&\cdots&a_{1k}\\a_{21}&a_{22}&\cdots&a_{2k}\\\vdots&\vdots&&\vdots\\a_{k1}&a_{k2}&\cdots&a_{kk}\end{vmatrix}$$(k=1,2,\cdots,n)$称为$\boldsymbol{A}$的顺序主子式.

Theorem 8.3.2

n阶实对称阵A是正定阵的充分必要条件是它的n个顺序主子式全大于零.

Proof 设与$\boldsymbol{A}$对应的实二次型为$f(x_1,x_2,\cdots,x_n)$.

若$\boldsymbol{A}=(a_{ij})$且f是正定型,则f可表示为$f(x_1,x_2,\cdots,x_n)=\sum\limits_{j=1}^{n}\sum\limits_{i=1}^{n}a_{ij}x_ix_j$.

令$f_k(x_1,x_2,\cdots,x_k)=\sum\limits_{j=1}^{k}\sum\limits_{i=1}^{k}a_{ij}x_ix_j$

则对任意一组不全为零的实数$c_1,c_2,\cdots,c_k$,有$f_k(c_1,c_2,\cdots,c_k)=f(c_1,c_2,\cdots,c_k,0,\cdots,0)>0$

所以,f_k是一个正定二次型.它的相伴矩阵$\boldsymbol{A}_k$由$\boldsymbol{A}$的前k行及前k列组成,因此$\boldsymbol{A}_k$是一个正定阵.

$\boldsymbol{A}_k$合同于$\boldsymbol{I}_k$,即k阶单位阵.也就是说存在k阶非异阵$\boldsymbol{B}$,使$\boldsymbol{B}'\boldsymbol{A}_k\boldsymbol{B}=\boldsymbol{I}_k$

因此$\det(\boldsymbol{B}'\boldsymbol{A}_k\boldsymbol{B})=(\det\boldsymbol{B})^2\det\boldsymbol{A}_k=1$,即有$\det\boldsymbol{A}_k>0$.这证明了必要性.

我们对$\boldsymbol{A}$的阶n用数学归纳法来证明充分性.

当$n=1$时,$\boldsymbol{A}=(a),a>0$,因此$f=ax_1^2$是正定型.

设结论对$n-1$成立,现要证明对n阶实对称阵$\boldsymbol{A}$,若它的n个顺序主子式全大于零,则$\boldsymbol{A}$必是正定阵.

记$\boldsymbol{A}_{n-1}$是$\boldsymbol{A}$的$n-1$阶顺序主子式所在的矩阵,$\boldsymbol{A}$可写为$\begin{pmatrix}\boldsymbol{A}_{n-1} & \boldsymbol{\alpha}\\ \boldsymbol{\alpha}' & a_{nn}\end{pmatrix}$.

因为$\boldsymbol{A}$的顺序主子式全大于零,$\boldsymbol{A}_{n-1}$的顺序主子式也全大于零,由归纳假设,$\boldsymbol{A}_{n-1}$是正定阵.

于是$\boldsymbol{A}_{n-1}$合同于$n-1$阶单位阵,即存在$n-1$阶非异阵$\boldsymbol{B}$,使$B'A_{n-1}B=I_{n-1}$.

令C是下列分块矩阵:$\boldsymbol{C}=\begin{pmatrix}\boldsymbol{B} & \boldsymbol{O}\\ \boldsymbol{O} & 1\end{pmatrix}$,则

$$\boldsymbol{C}'\boldsymbol{A}\boldsymbol{C}=\begin{pmatrix}\boldsymbol{B}' & \boldsymbol{O}\\ \boldsymbol{O} & 1\end{pmatrix}\begin{pmatrix}\boldsymbol{A}_{n-1} & \boldsymbol{\alpha}\\ \boldsymbol{\alpha}' & a_{nn}\end{pmatrix}\begin{pmatrix}\boldsymbol{B} & \boldsymbol{O}\\ \boldsymbol{O} & 1\end{pmatrix}=\begin{pmatrix}\boldsymbol{I}_{n-1} & \boldsymbol{B}'\boldsymbol{\alpha}\\ \boldsymbol{\alpha}'\boldsymbol{B} & a_{nn}\end{pmatrix}.$$

这是一个实对称阵,其形式为

$$\boldsymbol{C}'\boldsymbol{A}\boldsymbol{C}=\begin{pmatrix}1 & \cdots & 0 & c_1\\ \vdots & & \vdots & \vdots\\ 0 & \cdots & 1 & c_{n-1}\\ c_1 & \cdots & c_{n-1} & a_{nn}\end{pmatrix}.$$

用第三类初等行及列变换可将上述矩阵化为对角阵.

这相当于对$\boldsymbol{C}'\boldsymbol{A}\boldsymbol{C}$右乘一个非异阵$Q$后再左乘$Q'$得到一个对角阵,亦即$Q'C'ACQ$等于$\mathrm{diag}\{1,\cdots,1,c\}$.

由于$|\boldsymbol{A}|>0$,因此$c>0$.这就证明了$\boldsymbol{A}$是一个正定阵.

Corollary 8.2

若A是正定阵,证明:

(1)$\boldsymbol{A}$的任一k阶主子阵,即由$\boldsymbol{A}$的第$i_1,i_2,\cdots,i_k$行及$\boldsymbol{A}$的第$i_1,i_2,\cdots,i_k$列交点上元素组成的矩阵必是正定阵;

(2)$\boldsymbol{A}$的所有主子式全大于零,特别,$\boldsymbol{A}$的主对角元素全大于零;

(3)$\boldsymbol{A}$中绝对值最大的元素仅在主对角线上.

Proof (1)设$\boldsymbol{A}_k$是矩阵$\boldsymbol{A}$的第k个顺序主子式所在的矩阵

则$\boldsymbol{A}_k$是实对称阵且其顺序主子式都大于零,因此$\boldsymbol{A}_k$是正定阵.

经过若干次行对换以及相同的列对换,我们不难将$\boldsymbol{A}$的第$i_1,i_2,\cdots,i_k$行及$i_1,i_2,\cdots,i_k$列分别换成第$1,2,\cdots,k$行和第$1,2,\cdots,k$列.

利用上面的结论即知 (1)成立.(2)是 (1)的推论.

(3)用反证法.假定$a_{ij}(i\neq j)$是$\boldsymbol{A}$的绝对值最大的元素.根据 (1),我们只需证明第i,j行与第i,j列交点上元素组成的矩阵不是正定阵即可.

考虑矩阵 (由$\boldsymbol{A}$的对称性不妨设$i<j$)$\begin{pmatrix}a_{ii} & a_{ij}\\ a_{ji} & a_{jj}\end{pmatrix}$

注意到$a_{ij}=a_{ji}$且$|a_{ij}|\geq|a_{ii}|,|a_{ij}|\geq|a_{jj}|$,上述矩阵的行列式值$a_{ii}a_{jj}-a_{ij}^2\leq 0$.所以这个矩阵一定不是正定阵.

8.4 亚 (半) 正定阵与半正定阵性质

Definition 8.7 (亚正定阵)

设M为n阶实矩阵, 若对任意的非零实列向量α, 总有$\alpha' M\alpha > 0$, 则称M是亚正定阵.证明下列3个结论等价 :

(1)M是亚正定阵

(2)$M + M'$是正定阵

(3)$\boldsymbol{M} = \boldsymbol{A} + \boldsymbol{S}$, 其中$\boldsymbol{A}$是正定实对称矩阵, $\boldsymbol{S}$是实反对称矩阵.

Proof

(1) $\Rightarrow$ (2) : 将$\boldsymbol{\alpha}'\boldsymbol{M}\boldsymbol{\alpha} > 0$转置得$\boldsymbol{\alpha}'\boldsymbol{M}'\boldsymbol{\alpha} > 0$, 再将两式相加得$\boldsymbol{\alpha}'(\boldsymbol{M}+\boldsymbol{M}')\boldsymbol{\alpha} > 0$对任意的非零实列向量$\boldsymbol{\alpha}$都成立, 因此$M + M'$是正定阵.

(2) $\Rightarrow$ (3) : 令$\boldsymbol{A} = \frac{1}{2}(\boldsymbol{M}+\boldsymbol{M}')$ 为$\boldsymbol{M}$的对称化, $\boldsymbol{S} = \frac{1}{2}(\boldsymbol{M}-\boldsymbol{M}')$ 为$\boldsymbol{M}$的反对称化, 则结论成立.

(3) $\Rightarrow$ (1) : 可知, 对任意的非零实列向量$\boldsymbol{\alpha}$, 总有$\boldsymbol{\alpha}'\boldsymbol{M}\boldsymbol{\alpha} = \boldsymbol{\alpha}'\boldsymbol{A}\boldsymbol{\alpha} + \boldsymbol{\alpha}'\boldsymbol{S}\boldsymbol{\alpha} = \boldsymbol{\alpha}'\boldsymbol{A}\boldsymbol{\alpha} > 0$, 即$\boldsymbol{M}$为亚正定阵.

Proposition 8.1

1.$\boldsymbol{A}$为亚正定阵则其特征值的实部大于0

2. $|\boldsymbol{A}|$ 行列式大于0

Proof 1.设$\lambda_0 = a + bi$ 是$\boldsymbol{A}$的特征值, $\boldsymbol{\eta}$是属于λ_0的特征向量.

将$\boldsymbol{\eta}$的实部和虚部分开, 记为$\boldsymbol{\eta} = \boldsymbol{\alpha} + \mathrm{i}\boldsymbol{\beta}$, 则$\boldsymbol{A}(\boldsymbol{\alpha}+\mathrm{i}\boldsymbol{\beta}) = (a+b\mathrm{i})(\boldsymbol{\alpha}+\mathrm{i}\boldsymbol{\beta})$.分开实部和虚部可得

$\boldsymbol{A}\boldsymbol{\alpha} = a\boldsymbol{\alpha} - b\boldsymbol{\beta} \quad \boldsymbol{A}\boldsymbol{\beta} = b\boldsymbol{\alpha} + a\boldsymbol{\beta}$

于是$\boldsymbol{\alpha}'\boldsymbol{A}\boldsymbol{\alpha} = a\boldsymbol{\alpha}'\boldsymbol{\alpha} - b\boldsymbol{\alpha}'\boldsymbol{\beta} \quad \boldsymbol{\beta}'\boldsymbol{A}\boldsymbol{\beta} = b\boldsymbol{\beta}'\boldsymbol{\alpha} + a\boldsymbol{\beta}'\boldsymbol{\beta}$.

因此$\boldsymbol{\alpha}'\boldsymbol{A}\boldsymbol{\alpha} + \boldsymbol{\beta}'\boldsymbol{A}\boldsymbol{\beta} = a(\boldsymbol{\alpha}'\boldsymbol{\alpha} + \boldsymbol{\beta}'\boldsymbol{\beta})$ 因为$\boldsymbol{\alpha}, \boldsymbol{\beta}$中至少有一个是非零列向量, 故由假设可知, 式左边大于零, 又$\boldsymbol{\alpha}'\boldsymbol{\alpha} + \boldsymbol{\beta}'\boldsymbol{\beta} > 0$, 因此$a > 0$.

2.A的特征多项式为实系数多项式则其特征根如果为实数则大于0

如果是成对的虚根$a + bi$与$a - bi$则 $(a+bi)(a-bi) = a^2 + b^2 > 0$则 $|A| > 0$

也可以从亚正定阵$M = A + S$得到 $|M| = |A + S| \geqslant |A| + |S| > 0$

Theorem 8.4.1 (半正定与正定阵之间的摄动法)

n阶实对称矩阵$\boldsymbol{A}$是半正定阵的充要条件是对任意的正实数t, $\boldsymbol{A} + t\boldsymbol{I}_n$都是正定阵

Proof 先证必要性.对任一非零实列向量$\boldsymbol{\alpha}$, 有$\boldsymbol{\alpha}'(\boldsymbol{A} + t\boldsymbol{I}_n)\boldsymbol{\alpha} = \boldsymbol{\alpha}'\boldsymbol{A}\boldsymbol{\alpha} + t\boldsymbol{\alpha}'\boldsymbol{\alpha}$

因为$\boldsymbol{A}$半正定, 故$\boldsymbol{\alpha}'\boldsymbol{A}\boldsymbol{\alpha} \geq 0$.又$t > 0$且$\boldsymbol{\alpha}'\boldsymbol{\alpha} > 0$, 从而$\boldsymbol{\alpha}'(\boldsymbol{A} + t\boldsymbol{I}_n)\boldsymbol{\alpha} > 0$, 因此$\boldsymbol{A} + t\boldsymbol{I}_n$是正定阵

再证充分性。由假设对任一非零实列向量$\boldsymbol{\alpha}$和正实数t, 有$\boldsymbol{\alpha}'(\boldsymbol{A} + t\boldsymbol{I}_n)\boldsymbol{\alpha} = \boldsymbol{\alpha}'\boldsymbol{A}\boldsymbol{\alpha} + t\boldsymbol{\alpha}'\boldsymbol{\alpha} > 0$

令$t \to 0+$, 上式两边同取极限可得$\boldsymbol{\alpha}'\boldsymbol{A}\boldsymbol{\alpha} \geq 0$, 即$\boldsymbol{A}$是半正定阵.

Theorem 8.4.2

若主对角元为零, 则同行同列的所有元素都为零

Proof 设$\boldsymbol{A} = (a_{ij})$ 为n阶半正定实对称矩阵, 若$a_{ii} = 0$, 则$\boldsymbol{A}$的第i行和第i列的所有元素都等于零.

证明任取$j \neq i$, 考虑$\boldsymbol{A}$的第i, j行和列构成的主子式, 由可得$\begin{vmatrix} a_{ii} & a_{ij} \\ a_{ji} & a_{jj} \end{vmatrix} = a_{ii}a_{jj} - a_{ij}a_{ji} = -a_{ij}^2 \geq 0$

从而$a_{ij} = a_{ji} = 0 (j \neq i)$, 结论得证.

Theorem 8.4.3

设$\boldsymbol{A}$为n阶实对称矩阵

求证：$\boldsymbol{A}$为半正定阵或半负定阵的充要条件是对任一满足$\boldsymbol{\alpha}'\boldsymbol{A}\boldsymbol{\alpha}=0$的$n$维实列向量$\boldsymbol{\alpha}$,均有$\boldsymbol{A}\boldsymbol{\alpha}=0$

Proof 先证必要性.若$\boldsymbol{A}$是半正定阵,则存在实矩阵$\boldsymbol{C}$,使得$\boldsymbol{A}=\boldsymbol{C}'\boldsymbol{C}$,从而$0=\boldsymbol{\alpha}'\boldsymbol{A}\boldsymbol{\alpha}=\boldsymbol{\alpha}'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{\alpha}=(\boldsymbol{C}\boldsymbol{\alpha})'(\boldsymbol{C}\boldsymbol{\alpha})$于是$\boldsymbol{C}\boldsymbol{\alpha}=0$

因此$\boldsymbol{A}\boldsymbol{\alpha}=\boldsymbol{C}'(\boldsymbol{C}\boldsymbol{\alpha})=0$同理可证$\boldsymbol{A}$是半负定阵的情形.

再证充分性.用反证法,设$\boldsymbol{A}$既不是半正定阵,也不是半负定阵,则$\boldsymbol{A}$的正惯性指数$p>0$,负惯性指数$q>0$.

设$\boldsymbol{C}$是非异实矩阵,使得$\boldsymbol{B}=\boldsymbol{C}'\boldsymbol{A}\boldsymbol{C}=\operatorname{diag}\left\{\boldsymbol{I}_p,-\boldsymbol{I}_q,\boldsymbol{O}\right\}$ 为$\boldsymbol{A}$的合同标准型.

令$b_1=1,b_{p+1}=1$,其他b_i全为零,则$\boldsymbol{\beta}=(b_1,b_2,\cdots,b_n)'$ 是非零列向量,并且满足$\boldsymbol{\beta}'\boldsymbol{B}\boldsymbol{\beta}=0$,但$\boldsymbol{B}\boldsymbol{\beta}\neq 0$

从而$\boldsymbol{\alpha}=\boldsymbol{C}\boldsymbol{\beta}=(a_1,a_2,\cdots,a_n)'$ 也是非零列向量,并且满足$\boldsymbol{\alpha}'\boldsymbol{A}\boldsymbol{\alpha}=0$,但$\boldsymbol{A}\boldsymbol{\alpha}=\boldsymbol{A}\boldsymbol{C}\boldsymbol{\beta}=(\boldsymbol{C}')^{-1}\boldsymbol{B}\boldsymbol{\beta}\neq 0$,这就推出了矛盾.

Theorem 8.4.4

设$\boldsymbol{M}=\begin{pmatrix}\boldsymbol{A}&\boldsymbol{B}\\\boldsymbol{B}'&\boldsymbol{D}\end{pmatrix}$为半正定实对称矩阵：则有：$\mathrm{r}(\boldsymbol{A}\mid\boldsymbol{B})=\mathrm{r}(\boldsymbol{A})$进而$\boldsymbol{A}\boldsymbol{X}=\boldsymbol{B}$有解

Proof 证法1根据线性方程组的求解理论,要证明 $\mathrm{r}(\boldsymbol{A};\boldsymbol{B})=\mathrm{r}(\boldsymbol{A})$,只要证明线性方程组$\begin{pmatrix}\boldsymbol{A}\\\boldsymbol{B}'\end{pmatrix}\boldsymbol{x}=0$与$\boldsymbol{A}\boldsymbol{x}=0$同解即可.

显然前面线性方程组的解是后面线性方程组的解,下面证明反之也成立.

设$\boldsymbol{A}\boldsymbol{x}_0=0$,其中$\boldsymbol{x}_0$是实列向量,则有$\begin{pmatrix}\boldsymbol{x}_0'&0\end{pmatrix}\begin{pmatrix}\boldsymbol{A}&\boldsymbol{B}\\\boldsymbol{B}'&\boldsymbol{D}\end{pmatrix}\begin{pmatrix}\boldsymbol{x}_0\\0\end{pmatrix}=\boldsymbol{x}_0'\boldsymbol{A}x_0=0\Longrightarrow$ 可知$\begin{pmatrix}\boldsymbol{A}&\boldsymbol{B}\\\boldsymbol{B}'&\boldsymbol{D}\end{pmatrix}\begin{pmatrix}\boldsymbol{x}_0\\0\end{pmatrix}=0$,即有$\begin{pmatrix}\boldsymbol{A}\\\boldsymbol{B}'\end{pmatrix}\boldsymbol{x}_0=0$成立,从而结论得证.

证法2由$\boldsymbol{M}$的半正定性可得$\boldsymbol{A}$的半正定性,因此存在非异实矩阵$\boldsymbol{C}$,使得$C'AC=\begin{pmatrix}I_r&O\\O&O\end{pmatrix}$.

考虑如下合同变换：$\begin{pmatrix}\boldsymbol{C}'&\boldsymbol{O}\\\boldsymbol{O}&\boldsymbol{I}\end{pmatrix}\begin{pmatrix}\boldsymbol{A}&\boldsymbol{B}\\\boldsymbol{B}'&\boldsymbol{D}\end{pmatrix}\begin{pmatrix}\boldsymbol{C}&\boldsymbol{O}\\\boldsymbol{O}&\boldsymbol{I}\end{pmatrix}=\begin{pmatrix}\boldsymbol{C}'\boldsymbol{A}\boldsymbol{C}&\boldsymbol{C}'\boldsymbol{B}\\\boldsymbol{B}'\boldsymbol{C}&\boldsymbol{D}\end{pmatrix}=\begin{pmatrix}\boldsymbol{I}_r&\boldsymbol{O}&\boldsymbol{B}_1\\\boldsymbol{O}&\boldsymbol{O}&\boldsymbol{B}_2\\\boldsymbol{B}_1'&\boldsymbol{B}_2'&\boldsymbol{D}\end{pmatrix}$

由半正定阵对角元性质可知$\boldsymbol{B}_2=\boldsymbol{O}$.

对分块矩阵$(\boldsymbol{A}:\boldsymbol{B})$左乘$\boldsymbol{C}'$,相当于实施初等行变换,再对左边的分块$\boldsymbol{A}$右乘$\boldsymbol{C}$,相当于实施初等列变换

注意到矩阵的秩在初等变换下不改变,故有$\mathrm{r}\left(\boldsymbol{A}',\boldsymbol{B}\right)=\mathrm{r}\left(\boldsymbol{C}'\boldsymbol{A}\boldsymbol{C}:\boldsymbol{C}'\boldsymbol{B}\right)=\mathrm{r}\begin{pmatrix}\boldsymbol{I}_r&\boldsymbol{O}&\boldsymbol{B}_1\\\boldsymbol{O}&\boldsymbol{O}&\boldsymbol{O}\end{pmatrix}=r=\mathrm{r}(\boldsymbol{A})$

8.5 重点结论与知识

Theorem 8.5.1 (实对称矩阵的 Cholesky 分解)

下列关于n阶实对称矩阵$\boldsymbol{A}=(a_{ij})$的命题等价：

(1)$\boldsymbol{A}$是正定阵

(2)存在主对角元全等于1的上三角矩阵$\boldsymbol{B}$和主对角元全为正数的对角矩阵D，使得$\boldsymbol{A}=\boldsymbol{B}'D\boldsymbol{B}$

(3)存在主对角元全为正数的上三角矩阵$\boldsymbol{C}$, 使得$\boldsymbol{A}=\boldsymbol{C}'\boldsymbol{C}$.

Proof 证明 (1) $\Rightarrow$ (2) : 只要证明存在主对角元全为1的上三角矩阵$\boldsymbol{T}$, 使得$\boldsymbol{T}'\boldsymbol{A}\boldsymbol{T}=D$是正定对角矩阵即可.
因为一旦得证, $\boldsymbol{B}=\boldsymbol{T}^{-1}$也是主对角元全为1的上三角矩阵, 并且$\boldsymbol{A}=\boldsymbol{B}'D\boldsymbol{B}$
对阶数n进行归纳, 当$n=1$时结论显然成立.假设对$n-1$阶正定阵结论成立, 现证明n阶正定阵的情形
设$\boldsymbol{A}=\begin{pmatrix}\boldsymbol{A}_{n-1} & \boldsymbol{\alpha}\\ \boldsymbol{\alpha}' & a_{nn}\end{pmatrix}$，其中$\boldsymbol{A}_{n-1}$是$n-1$阶矩阵, $\boldsymbol{\alpha}$是$n-1$维列向量.
因为$\boldsymbol{A}$正定, 所以$\boldsymbol{A}_{n-1}$是$n-1$阶正定阵, 从而是可逆矩阵.考虑如下对称分块初等变换：
$\begin{pmatrix}\boldsymbol{I}_{n-1} & \boldsymbol{O}\\ -\boldsymbol{\alpha}'\boldsymbol{A}_{n-1}^{-1} & 1\end{pmatrix}\begin{pmatrix}\boldsymbol{A}_{n-1} & \boldsymbol{\alpha}\\ \boldsymbol{\alpha}' & a_{nn}\end{pmatrix}\begin{pmatrix}\boldsymbol{I}_{n-1} & -\boldsymbol{A}_{n-1}^{-1}\boldsymbol{\alpha}\\ \boldsymbol{O} & 1\end{pmatrix}=\begin{pmatrix}\boldsymbol{A}_{n-1} & \boldsymbol{O}\\ \boldsymbol{O} & a_{nn}-\boldsymbol{\alpha}'\boldsymbol{A}_{n-1}^{-1}\boldsymbol{\alpha}\end{pmatrix}$ 由$\boldsymbol{A}$的正定性可得$a_{nn}-\boldsymbol{\alpha}'\boldsymbol{A}_{n-1}^{-1}\boldsymbol{\alpha}>0$.
再由归纳假设, 存在主对角元全为1的$n-1$阶上三角矩阵$\boldsymbol{T}_{n-1}$, 使得$\boldsymbol{T}_{n-1}'\boldsymbol{A}_{n-1}\boldsymbol{T}_{n-1}=\boldsymbol{D}_{n-1}$是$n-1$阶正定对角矩阵.
令$T=\begin{pmatrix}I_{n-1} & -A_{n-1}^{-1}\alpha\\ O & 1\end{pmatrix}\begin{pmatrix}T_{n-1} & O\\ O & 1\end{pmatrix}$ 则$\boldsymbol{T}$是一个主对角全为1的n阶上三角阵, 使$\boldsymbol{T}'\boldsymbol{A}\boldsymbol{T}=\begin{pmatrix}\boldsymbol{D}_{n-1} & \boldsymbol{O}\\ \boldsymbol{O} & a_{nn}-\boldsymbol{\alpha}'\boldsymbol{A}_{n-1}^{-1}\alpha\end{pmatrix}$是$n$阶正定对角矩阵.

(2) $\Rightarrow$ (3) : 由 (2) 可设$D=\operatorname{diag}\{d_1,d_2,\cdots,d_n\}$, 令$s_i=\sqrt{d_i}>0$, $\boldsymbol{S}=\operatorname{diag}\{s_1,s_2,\cdots,s_n\}$
设$\boldsymbol{C}=\boldsymbol{S}\boldsymbol{B}$, 则$\boldsymbol{A}=\boldsymbol{C}'\boldsymbol{C}$.显然$\boldsymbol{C}=\boldsymbol{S}\boldsymbol{B}$是主对角元全为正数的上三角矩阵.

(3) $\Rightarrow$ (1) : 这时$\boldsymbol{A}=\boldsymbol{C}'\boldsymbol{I}_n\boldsymbol{C}$, 故$\boldsymbol{A}$和$\boldsymbol{I}_n$合同, 从而$\boldsymbol{A}$正定.

注 注设$C=(c_{ij})$为主对角元全为正数的上三角矩阵, 使得$\boldsymbol{A}=\boldsymbol{C}'\boldsymbol{C}$，则$c_{11}c_{1j}=a_{1j}$, 从而$c_{11}=\sqrt{a_{11}}>0, c_{1j}=\dfrac{a_{1j}}{\sqrt{a_{11}}}(2\le j\le n)$
即$\boldsymbol{C}$的第一行元素被唯一确定.同理不断地讨论下去, 可得这样的$\boldsymbol{C}$存在并被正定阵$\boldsymbol{A}$唯一确定.
因为S是由C的主对角元构成的对角矩阵，故由C的唯一性可得S的唯一性，从而可得$\boldsymbol{D}=\boldsymbol{S}^2$以及$\boldsymbol{B}=\boldsymbol{S}^{-1}\boldsymbol{C}$的唯一性
因此，关于正定阵$\boldsymbol{A}$的两种分解 (2)和 (3)都是存在且唯一的, 其中分解 (3)通常称为正定阵$\boldsymbol{A}$的$Cholesky$分解.

Theorem 8.5.2 (反对称矩阵的合同标准型)

设A是数域P上的一个反称矩阵, 则A合同于准对角矩阵$\operatorname{diag}\left\{\begin{pmatrix}0&1\\-1&0\end{pmatrix},\cdots,\begin{pmatrix}0&1\\-1&0\end{pmatrix},0,\cdots,0\right\}$

Proof 数学归纳法.
当$n=1$时, $A=0$, 命题显然成立
当$n=2$时, 可设$A=\begin{pmatrix}0&a\\-a&0\end{pmatrix}$, 若$a=0$, 则结论也成立, 若$a\neq 0$, 则将$A$的第一行与第一节均乘以$\frac{1}{a}$, 即得$A\simeq\begin{pmatrix}0&1\\-1&0\end{pmatrix}$, 所以此时命题也成
接下来设命题对级数小于等于$n-1(n\ge 3)$的反称矩阵成立, 现在考虑n级反称矩阵$A=(a_{ij})$的情况：
(i)若$a_{12}\neq 0$, 则将A做如下分块$A=\begin{pmatrix}A_1&A_2\\-A_2'&A_4\end{pmatrix}$其中$A_1=\begin{pmatrix}0&a_{12}\\-a_{12}&0\end{pmatrix}$
显然A_1可逆, 所以取$P=\begin{pmatrix}E_2&-A_1^{-1}A_2\\O&E_{n-2}\end{pmatrix}$, 就有$P'AP=\begin{pmatrix}A_1&O\\O&A_4+A_2'A_1^{-1}A_2\end{pmatrix}$

根据$n=2$的情况可知矩阵$A_1 \simeq \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$

而显然$A_4 + A_2' A_1^{-1} A_2$为$n-2$级反称矩阵,所以由归纳假设可知$A_4 + A_2' A_1^{-1} A_2 \simeq \operatorname{diag}\left\{\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \cdots, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, 0, \cdots, 0\right\}$

从而$A \simeq P'AP \simeq \operatorname{diag}\left\{\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \cdots, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, 0, \cdots, 0\right\}$

(ii)若$a_{12}=0$,但存在$a_{1j} \neq 0 (3 \le j \le n)$,此时可以将$A$的第$j$列加到第2列,将第$j$行加到第2行,这就将$A$通过合同的方式转化为了$(i)$,于

(iii)若$a_{11}=a_{12}=\cdots=a_{1n}=0$,那么$A=\begin{pmatrix} 0 & 0 \\ 0 & A_{n-1} \end{pmatrix}$,其中$A_{n-1}$为$n-1$级反称矩阵.

先取分块初等矩阵$Q=\begin{pmatrix} & 1 \\ E_{n-1} & \end{pmatrix}$,有$Q'AQ=\begin{pmatrix} A_{n-1} & \\ & 0 \end{pmatrix}$

而根据归纳假设可知$A_{n-1} \simeq \operatorname{diag}\left\{\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \cdots, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, 0, \cdots, 0\right\}$

综上

所以$A \simeq Q'AQ \simeq \operatorname{diag}\left\{\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \cdots, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, 0, \cdots, 0\right\}$

Corollary 8.3

n阶实反对称矩阵$\boldsymbol{A}$的行列式值总是非负实数.

Proof 可知,存在非异实矩阵$\boldsymbol{C}$,使得$\boldsymbol{C}'\boldsymbol{A}\boldsymbol{C} = \operatorname{diag}\{\boldsymbol{S}, \cdots, \boldsymbol{S}, 0, \cdots, 0\}$其中$\boldsymbol{S}=\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

若$\boldsymbol{A}$是奇异阵,则$|\boldsymbol{A}|=0$,结论显然成立.若$\boldsymbol{A}$是非异阵(蕴含了A的阶数为偶数),则由上式可得$|\boldsymbol{A}| \cdot |\boldsymbol{C}|^2 = |\boldsymbol{S}|^{n/2} = 1$,从而$|\boldsymbol{A}| > 0$.

Proposition 8.2 (实反对称矩阵的行列式估计)

设$\boldsymbol{A}$为n阶实反对称矩阵,$\boldsymbol{Q}$为正定阵.$\boldsymbol{P}$为半正定阵那么:

(1) $|\boldsymbol{I}_n + \boldsymbol{A}| \ge 1 + |\boldsymbol{A}|$,且等号成立当且仅当$n \le 2$或当$n \ge 3$时$\boldsymbol{A}=\boldsymbol{O}$

(2) $|\boldsymbol{I}_n + \boldsymbol{A}| \ge 1$,且等号成立当且仅当$\boldsymbol{A}=\boldsymbol{O}$

(3) $|\boldsymbol{Q}+\boldsymbol{A}| \geqslant |\boldsymbol{Q}| + |\boldsymbol{A}|$且等号成立当且仅当$n \le 2$或当$n \ge 3$时$\boldsymbol{A}=\boldsymbol{O}$

(4) $|\boldsymbol{P}+\boldsymbol{A}| \geqslant |\boldsymbol{P}| + |\boldsymbol{A}| \geqslant |\boldsymbol{P}| \geqslant 0$ (利用摄动法证明)

Proof 可知$|\boldsymbol{I}_n + \boldsymbol{A}| = |\boldsymbol{I}_n| + |\boldsymbol{A}| + \sum\limits_{1 \le k \le n-1} \left(\sum\limits_{1 \le i_1 < i_2 < \cdots < i_k \le n} \boldsymbol{A}\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ i_1 & i_2 & \cdots & i_k \end{pmatrix} \right)$

注意到$\boldsymbol{A}\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ i_1 & i_2 & \cdots & i_k \end{pmatrix}$是$k$阶实反对称行列式,故可知其值大于等于零,于是$|\boldsymbol{I}_n + \boldsymbol{A}| \ge 1 + |\boldsymbol{A}|$成立.

当$n \le 2$时,容易验证不等式的等号成立.

当$n \ge 3$时,若不等式的等号成立,则必有$\boldsymbol{A}\begin{pmatrix} i & j \\ i & j \end{pmatrix} = \begin{vmatrix} 0 & a_{ij} \\ -a_{ij} & 0 \end{vmatrix} = a_{ij}^2 = 0$.即有$a_{ij}=0 (1 \le i < j \le n)$,从而$\boldsymbol{A}=\boldsymbol{O}$.

(2)同理可证

(3) 对于正定阵$\boldsymbol{Q}=\boldsymbol{C}'\boldsymbol{C}$ ($\boldsymbol{C}$为一可逆阵),那么此时$\left(\boldsymbol{C}^{-1}\right)' \boldsymbol{A}\boldsymbol{C}^{-1}$同样为反对称矩阵则由(1)

$\left|\boldsymbol{I}_n + \left(\boldsymbol{C}^{-1}\right)' \boldsymbol{A}\boldsymbol{C}^{-1}\right| \ge 1 + \left|\left(\boldsymbol{C}^{-1}\right)' \boldsymbol{A}\boldsymbol{C}^{-1}\right| \xrightarrow{\text{两边同时乘}\boldsymbol{C}'\boldsymbol{C}} |\boldsymbol{Q}+\boldsymbol{A}| = |\boldsymbol{Q}| + |\boldsymbol{A}|$ 等号成立当且仅当$n \le 2$或当$n \ge 3$时$\boldsymbol{A}=\boldsymbol{O}$

Proposition 8.3

证明：一个秩大于1的实二次型可以分解为两个实系数一次多项式之积的充要条件是它的秩等于2，且符号差等于零

Proof 先证必要性.设秩大于1的实二次型$f(x_1,x_2,\cdots,x_n)=(a_1x_1+a_2x_2+\cdots+a_nx_n)(b_1x_1+b_2x_2+\cdots+b_nx_n)$

令$y_1=a_1x_1+a_2x_2+\cdots+a_nx_n,\quad y_2=b_1x_1+b_2x_2+\cdots+b_nx_n$

如果向量$(a_1,a_2,\cdots,a_n)$和$(b_1,b_2,\cdots,b_n)$线性相关,则它们的元素成比例

不妨假设$(b_1,b_2,\cdots,b_n)=k(a_1,a_2,\cdots,a_n)$,于是(不妨设$a_1\neq 0$)

构造可逆线性替换$\begin{cases} y_1=a_1x_1+a_2x_2+\cdots+a_nx_n \\ y_2=x_2 \\ \cdots \\ y_n=x_n \end{cases}\implies f=ky_1^2$,这与$f$的秩大于1相矛盾

因此向量$(a_1,a_2,\cdots,a_n)$和$(b_1,b_2,\cdots,b_n)$线性无关.不妨设$\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}\neq 0$

定义可逆线性变换如下：$\begin{cases} y_1=a_1x_1+a_2x_2+\cdots+a_nx_n \\ y_2=b_1x_1+b_2x_2+\cdots+b_nx_n \\ y_i=x_i(3\le i\le n) \end{cases}$ 得到$f=y_1y_2$.

再令$y_1=z_1+z_2,y_2=z_1-z_2,y_i=z_i(3\le i\le n)$,可得$f=z_1^2-z_2^2$,显然$f$的秩为2且符号差为零

再证充分性.假设$f=y_1^2-y_2^2$,显然有$f=(y_1+y_2)(y_1-y_2)$,即f可以分解为两个一次多项式之积.

Theorem 8.5.3 (线性替换后的正负指数只会减小)

1.设实二次型$f(x_1,x_2,\cdots,x_n)=y_1^2+\cdots+y_k^2-y_{k+1}^2-\cdots-y_{k+s}^2$其中$y_i=a_{i1}x_1+a_{i2}x_2+\cdots+a_{in}x_n(1\le i\le k+s)$

求证：f的正惯性指数$p\le k$,负惯性指数$q\le s$

2.设$\boldsymbol{A}$为m阶实对称矩阵,$\boldsymbol{C}$为$m\times n$实矩阵,证明：$\boldsymbol{C}'\boldsymbol{A}\boldsymbol{C}$的正惯性指数小于等于$\boldsymbol{A}$的正惯性指数

$\boldsymbol{C}'\boldsymbol{A}\boldsymbol{C}$的负惯性指数小于等于$\boldsymbol{A}$的负惯性指数

Proof 假设经过可逆线性变换$\boldsymbol{x}=\boldsymbol{C}\boldsymbol{z}$后$f$变为规范标准型：$f(x_1,x_2,\cdots,x_n)=z_1^2+\cdots+z_p^2-z_{p+1}^2-\cdots-z_{p+q}^2$于是有

$y_1^2+\cdots+y_k^2-y_{k+1}^2-\cdots-y_{k+s}^2=z_1^2+\cdots+z_p^2-z_{p+1}^2-\cdots-z_{p+q}^2\quad(*)$

若$p>k$,作线性方程组$\begin{cases} y_i=0(1\le i\le k) \\ z_j=0(p+1\le j\le n) \end{cases}$

这是一个未知数个数（以$x_1,x_2,\cdots,x_n$为未知数）超过方程式个数$(n-(p-k))$的齐次线性方程组

故必有非零解$\boldsymbol{x}=\boldsymbol{\alpha}$此时$0\neq\boldsymbol{z}=\boldsymbol{C}^{-1}\boldsymbol{x}$而$z_{p+1}\sim z_n=0$故$z_1\sim z_p$不全为0

将$\boldsymbol{x}=\boldsymbol{\alpha}$代入$(*)$式,其左边小于等于零,而其右边严格大于零,矛盾,于是只能$p\leqslant k$

同理可证$q\le s$.

2.由于正负惯性指数是合同不变量，故不妨假设$\boldsymbol{A}=\operatorname{diag}\{\boldsymbol{I}_k,-\boldsymbol{I}_s,\boldsymbol{O}\}$是合同标准型,其中$k,s$分别是$\boldsymbol{A}$的正负惯性指数.

设$\boldsymbol{x}=(x_1,x_2,\cdots,x_n)'$,$f(\boldsymbol{x})=\boldsymbol{x}'\boldsymbol{C}'\boldsymbol{A}\boldsymbol{C}\boldsymbol{x}$是相伴于$\boldsymbol{C}'\boldsymbol{A}\boldsymbol{C}$的二次型,$\boldsymbol{C}=(a_{ij})_{m\times n}$,$\boldsymbol{y}=(y_1,y_2,\cdots,y_m)'=\boldsymbol{C}\boldsymbol{x}$

即$y_i=a_{i1}x_1+a_{i2}x_2+\cdots+a_{in}x_n(1\le i\le m)$,则$f(\boldsymbol{x})=(\boldsymbol{C}\boldsymbol{x})'\boldsymbol{A}(\boldsymbol{C}\boldsymbol{x})=\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y}=y_1^2+\cdots+y_k^2-y_{k+1}^2-\cdots-y_{k+s}^2$

由(1)可知,$f(\boldsymbol{x})$的正惯性指数$p\le k$,负惯性指数$q\le s$,结论得证.

Theorem 8.5.4

设实二次型$f\left(x_1,x_2,\cdots,x_n\right)=\sum\limits_{i=1}^{k}\left(a_{i1}x_1+a_{i2}x_2+\cdots+a_{in}x_n\right)^2$其中$a_{ij}$都是实数

求证f是半正定型且f的秩等于下列矩阵的秩：$\boldsymbol{A}=\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&&\vdots\\a_{k1}&a_{k2}&\cdots&a_{kn}\end{pmatrix}$

Proof 设$\alpha_i=\begin{pmatrix}a_{i1}&a_{i2}&\cdots&a_{in}\end{pmatrix}$ $\boldsymbol{x}=\begin{pmatrix}x_1&x_2&\cdots&x_n\end{pmatrix}'$ 则$f=\sum\limits_{i=1}^{k}\boldsymbol{x}'\alpha_i'\alpha_i\boldsymbol{x}=\boldsymbol{x}'A'A\boldsymbol{x}$

$A=\begin{pmatrix}\alpha_1&\alpha_2&\cdots&\alpha_n\end{pmatrix}'$ f的半正定性由定义即得

注意到$f(\boldsymbol{x})=(\boldsymbol{A}\boldsymbol{x})'(\boldsymbol{A}\boldsymbol{x})=\boldsymbol{x}'\left(\boldsymbol{A}'\boldsymbol{A}\right)\boldsymbol{x}$, 故$f$的相伴矩阵为$\boldsymbol{A}'\boldsymbol{A}$, 于是可知, $\mathrm{r}(f)=\mathrm{r}\left(\boldsymbol{A}'\boldsymbol{A}\right)=\mathrm{r}(\boldsymbol{A})$.

Theorem 8.5.5 (利用顺序主子式的值确定标准型与规范型)

设$f(\boldsymbol{x})=\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}$是实二次型, 相伴矩阵$\boldsymbol{A}$的前$n-1$个顺序主子式$P_1,\cdots,P_{n-1}$非零

求证：经过可逆线性变换f可化为下列标准型：$f=P_1y_1^2+\dfrac{P_2}{P_1}y_2^2+\cdots+\dfrac{P_n}{P_{n-1}}y_n^2$其中$P_n=|\boldsymbol{A}|$.

Proof 对n用归纳法. 当$n=1$时结论显然成立, 假设结论对$n-1$成立.

设$\boldsymbol{A}=\begin{pmatrix}\boldsymbol{A}_{n-1}&\boldsymbol{\alpha}\\\boldsymbol{\alpha}'&a_{nn}\end{pmatrix}$ 由于$\left|\boldsymbol{A}_{n-1}\right|=P_{n-1}\neq0$

故可对$\boldsymbol{A}$进行下列对称分块初等变换：$\boldsymbol{A}=\begin{pmatrix}\boldsymbol{A}_{n-1}&\boldsymbol{\alpha}\\\boldsymbol{\alpha}'&a_{nn}\end{pmatrix}\to\begin{pmatrix}\boldsymbol{A}_{n-1}&\boldsymbol{\alpha}\\\boldsymbol{O}&a_{nn}-\boldsymbol{\alpha}'\boldsymbol{A}_{n-1}^{-1}\boldsymbol{\alpha}\end{pmatrix}\to\begin{pmatrix}\boldsymbol{A}_{n-1}&\boldsymbol{O}\\\boldsymbol{O}&a_{nn}-\boldsymbol{\alpha}'\boldsymbol{A}_{n-1}^{-1}\boldsymbol{\alpha}\end{pmatrix}=\boldsymbol{B}$

显然这是一个合同变换.

又因为第三类分块初等变换不改变行列式的值, 故$|\boldsymbol{A}|=\left|\boldsymbol{A}_{n-1}\right|\left(a_{nn}-\boldsymbol{\alpha}'\boldsymbol{A}_{n-1}^{-1}\boldsymbol{\alpha}\right)$ 即$a_{nn}-\boldsymbol{\alpha}'\boldsymbol{A}_{n-1}^{-1}\boldsymbol{\alpha}=\dfrac{P_n}{P_{n-1}}$

由归纳假设, 存在可逆矩阵$\boldsymbol{M}$, 使得$\boldsymbol{M}'\boldsymbol{A}_{n-1}\boldsymbol{M}=\operatorname{diag}\left\{P_1,\dfrac{P_2}{P_1},\cdots,\dfrac{P_{n-1}}{P_{n-2}}\right\}$.

作矩阵$\boldsymbol{C}=\begin{pmatrix}\boldsymbol{M}&\boldsymbol{O}\\\boldsymbol{O}&1\end{pmatrix}$, 则$\boldsymbol{C}'\boldsymbol{B}\boldsymbol{C}=\operatorname{diag}\left\{P_1,\dfrac{P_2}{P_1},\cdots,\dfrac{P_n}{P_{n-1}}\right\}$

Theorem 8.5.6 (正定矩阵与半正定矩阵和的行列式估计)

1.已知A是n级正定矩阵, B是n级半正定矩阵, 则$|A+B|\geq|A|+|B|$, 并且仅当$B=O$时等号成立

2.已知A,B都是n级半正定矩阵, 则$|A+B|\geq|A|+|B|$.

Proof 1.由于A正定，所以$|A|>0$, $|A+B|\geq|A|+|B|$就等价于$\left|E+A^{-1}B\right|\geq1+\left|A^{-1}B\right|$.

由于A正定, 所以A^{-1}也正定, 从而存在可逆矩阵C使得$A^{-1}=CC'$, 所以

$A^{-1}B=CC'B=CC'BCC^{-1}\sim C'BC\simeq B$

由B半正定得到$C'BC$半正定, 设$C'BC$的特征值为$\lambda_1,\lambda_2,\cdots,\lambda_n$, 则$\lambda_i\geq0(i=1,2,\cdots,n)$且$\left|A^{-1}B\right|=\lambda_1\lambda_2\cdots\lambda_n$.

由于$A^{-1}B\sim C'BC$，所以$A^{-1}B$的特征值也是$\lambda_1,\lambda_2,\cdots,\lambda_n$, 从而$E+A^{-1}B$的特征值为$1+\lambda_1,1+\lambda_2,\cdots,1+\lambda_n$, 故

$\left|E+A^{-1}B\right|=\left(1+\lambda_1\right)\left(1+\lambda_2\right)\cdots\left(1+\lambda_n\right)\geq1+\lambda_1\lambda_2\cdots\lambda_n=1+\left|A^{-1}B\right|$ 且仅当$\lambda_1=\lambda_2=\cdots=\lambda_n=0$时等号成立

等号成立 $\Longleftrightarrow\lambda_1=\lambda_2=\cdots=\lambda_n=0\xLongleftrightarrow{\text{又}C'BC\text{半正定}}C'BC$正交相似标准型为$O\Longleftrightarrow C'BC=O\Longleftrightarrow B=O$

2.利用摄动法

Theorem 8.5.7

已知A, D分别是n级, m级实方阵, B是$n \times m$的实矩阵, 且$\begin{pmatrix} A & B \\ B' & D \end{pmatrix}$是正定矩阵, 则

(1)$A, D, D - B'A^{-1}B, A - BD^{-1}B'$都是正定矩阵

(2)$\begin{vmatrix} A & B \\ B' & D \end{vmatrix} \le |A||D|$，且仅当$B = O$时等号成立

Proof (1) 由于$\begin{pmatrix} A & B \\ B' & D \end{pmatrix}$正定, 所以它的所有主子式都大于零

而A, D的顺序主子式都是$\begin{pmatrix} A & B \\ B' & D \end{pmatrix}$的主子式, 所以$A, D$的顺序主子式都大于零, 这就得到$A, D$都是正定矩阵.

同时, 利用打洞原理可得$\begin{pmatrix} A & O \\ O & D - B'A^{-1}B \end{pmatrix}$与$\begin{pmatrix} A - BD^{-1}B' & O \\ O & D \end{pmatrix}$都是正定矩阵

它们的所有主子式都大于零, 从而$D - B'A^{-1}B$与$A - BD^{-1}B'$的顺序主子式都大于零, 即$D - B'A^{-1}B, A - BD^{-1}B'$也都是正定矩阵.

(2)利用打洞原理, 我们知道$\begin{vmatrix} A & B \\ B' & D \end{vmatrix} = |A| \left|D - B'A^{-1}B\right|$.

注意到$|A| > 0$, 所以$\begin{vmatrix} A & B \\ B' & D \end{vmatrix} = |A| \left|D - B'A^{-1}B\right| \le |A||D|$等价于$\left|D - B'A^{-1}B\right| \le |D|$.

由于$D - B'A^{-1}B$是一个正定矩阵, 且$B'A^{-1}B$是一个半正定矩阵

$\Longrightarrow$ 我们有$|D| = \left|(D - B'A^{-1}B) + B'A^{-1}B\right| \ge \left|D - B'A^{-1}B\right|$且等号成立的充要条件是$B'A^{-1}B = O$, 这就等价于$B = O$

原因是：如果$B'A^{-1}B = O$，可设$A^{-1} = C'C$，其中C是实可逆方阵, 则 $(BC)'(BC) = O$

从而$r(B) = r(BC) = r\left((BC)'(BC)\right) = 0$, 即$B = O$

所以, 我们就得到$\begin{vmatrix} A & B \\ B' & D \end{vmatrix} \le |A||D|$，且仅当$B = O$时等号成立.

Corollary 8.4

设$A = \begin{pmatrix} A_{n-1} & \alpha \\ \alpha' & a_{nn} \end{pmatrix}$是一个正定矩阵, 则$|A| \le |A_{n-1}| \, a_{nn}$，递推就有$|A| \le a_{11}a_{22} \cdots a_{nn}$

再推广, 如果A是一个实方阵, 当A可逆时, 则AA'是一个正定矩阵, 从而 $|AA'| \le \prod_{i=1}^{n} \left(a_{i1}^2 + a_{i2}^2 + \cdots + a_{in}^2\right)$

当A不可逆时, $|AA'| = 0$, 当然也有 $|AA'| \le \prod_{i=1}^{n} \left(a_{i1}^2 + a_{i2}^2 + \cdots + a_{in}^2\right)$

所以对任意的实方阵$A = (a_{ij})_{n \times n}$ 都有 $|AA'| \le \prod_{i=1}^{n} \left(a_{i1}^2 + a_{i2}^2 + \cdots + a_{in}^2\right)$这就是所谓的$Hadamard$不等式.

第 9 章 内积空间

9.1 内积空间的性质

Definition 9.1 (内积定义与欧式空间与酉空间)

设V是实数域上的线性空间

若存在某种规则,使对V中任意一组有序向量$\{\boldsymbol{\alpha},\boldsymbol{\beta}\}$,都唯一地对应一个实数,记为 $(\boldsymbol{\alpha},\boldsymbol{\beta})$,且适合如下规则:

(1)$(\boldsymbol{\beta},\boldsymbol{\alpha})=(\boldsymbol{\alpha},\boldsymbol{\beta})$;

(2)$(\boldsymbol{\alpha}+\boldsymbol{\beta},\boldsymbol{\gamma})=(\boldsymbol{\alpha},\boldsymbol{\gamma})+(\boldsymbol{\beta},\boldsymbol{\gamma})$;

(3)$(c\boldsymbol{\alpha},\boldsymbol{\beta})=c(\boldsymbol{\alpha},\boldsymbol{\beta})$,$c$为任一实数;

(4)$(\boldsymbol{\alpha},\boldsymbol{\alpha})\geq 0$且等号成立当且仅当$\boldsymbol{\alpha}=0$

则称在V上定义了一个内积.实数 $(\boldsymbol{\alpha},\boldsymbol{\beta})$称为$\boldsymbol{\alpha}$与$\boldsymbol{\beta}$的内积.线性空间$V$称为实内积空间

有限维实内积空间称为$Euclid$空间,简称为欧氏空间

设V是复数域上的线性空间

若存在某种规则,使对V中任意一组有序向量$\{\boldsymbol{\alpha},\boldsymbol{\beta}\}$,都唯一地对应一个复数,记为 $(\boldsymbol{\alpha},\boldsymbol{\beta})$,且适合如下规则:

(1)$(\boldsymbol{\beta},\boldsymbol{\alpha})=\overline{(\boldsymbol{\alpha},\boldsymbol{\beta})}$;

(2)$(\boldsymbol{\alpha}+\boldsymbol{\beta},\boldsymbol{\gamma})=(\boldsymbol{\alpha},\boldsymbol{\gamma})+(\boldsymbol{\beta},\boldsymbol{\gamma})$;

(3)$(c\boldsymbol{\alpha},\boldsymbol{\beta})=c(\boldsymbol{\alpha},\boldsymbol{\beta})$,$c$为任一复数;

(4)$(\boldsymbol{\alpha},\boldsymbol{\alpha})\geq 0$且等号成立当且仅当$\boldsymbol{\alpha}=0$

则称在V上定义了一个内积.复数 $(\boldsymbol{\alpha},\boldsymbol{\beta})$称为$\boldsymbol{\alpha}$与$\boldsymbol{\beta}$的内积.线性空间$V$称为复内积空间.有限维复内积空间称为酉空间.

注意酉空间上关于第二变量的数乘提出是共轭$(\alpha,c\beta)=\overline{c}\,(\alpha,\beta)$

Example 9.1 内积空间的例子

1.设$\mathbb{R}_n$是n维实列向量空间,$\boldsymbol{\alpha}=(x_1,x_2,\cdots,x_n)'$,$\boldsymbol{\beta}=(y_1,y_2,\ \cdots,y_n)'$

定义 $(\boldsymbol{\alpha},\boldsymbol{\beta})=x_1y_1+x_2y_2+\cdots+x_ny_n$,则我们定义了一个内积,这个内积称为$\mathbb{R}_n$中的标准内积.

2.设$\mathbb{C}_n$是n维复列向量空间,$\boldsymbol{\alpha}=(x_1,x_2,\cdots,x_n)'$,$\boldsymbol{\beta}=(y_1,y_2,\ \cdots,y_n)'$

定义 $(\boldsymbol{\alpha},\boldsymbol{\beta})=x_1\bar{y}_1+x_2\bar{y}_2+\cdots+x_n\bar{y}_n$,则在此定义下$\mathbb{C}_n$成为一个酉空间,上述内积称为$\mathbb{C}_n$的标准内积.

3.设V是n维实列向量空间,$\boldsymbol{G}$是n阶正定实对称阵,对$\boldsymbol{\alpha},\boldsymbol{\beta}\in V$,定义 $(\boldsymbol{\alpha},\boldsymbol{\beta})=\boldsymbol{\alpha}'\boldsymbol{G}\boldsymbol{\beta}$

当$\boldsymbol{G}=\boldsymbol{I}_n$为单位阵时,$V$上内积就是$\mathbb{R}_n$中的标准内积.

4.对n维复列向量空间U,若有正定$Hermite$矩阵$\boldsymbol{H}$,也可定义U上内积:$(\boldsymbol{\alpha},\boldsymbol{\beta})=\boldsymbol{\alpha}'\boldsymbol{H}\overline{\boldsymbol{\beta}}$

当$\boldsymbol{H}=\boldsymbol{I}_n$时就是$\mathbb{C}_n$中的标准内积.

Proof 我们来证明V在上式的定义下成为欧氏空间

定义 (2), (3)显然成立.

对 (1),注意到$\boldsymbol{\alpha}'\boldsymbol{G}\boldsymbol{\beta}$是实数,其转置仍是它自己,而$\boldsymbol{G}$是对称阵,故 $(\alpha,\beta)=\alpha'G\beta=(\alpha'G\beta)'=\beta'G'\alpha=\boldsymbol{\beta}'G\alpha=(\beta,\alpha)$.

又从$\boldsymbol{G}$是正定阵即可知道 (4)成立.

Definition 9.2 (范数与距离定义)

1.设V是内积空间 (实或复),α是V中的向量,定义α的长度 (或范数) 为 $\|\boldsymbol{\alpha}\|=(\boldsymbol{\alpha},\boldsymbol{\alpha})^{\frac{1}{2}}$,即实数 $(\boldsymbol{\alpha},\boldsymbol{\alpha})$的算术根.

2.设$\boldsymbol{\alpha},\boldsymbol{\beta}\in V$,定义$\boldsymbol{\alpha}$与$\boldsymbol{\beta}$的距离为$d(\boldsymbol{\alpha},\boldsymbol{\beta})=\|\boldsymbol{\alpha}-\boldsymbol{\beta}\|$.显然$d(\boldsymbol{\alpha},\boldsymbol{\beta})=d(\boldsymbol{\beta},\boldsymbol{\alpha})$.

注意,当V是复内积空间时,由于规则 (4), $(\boldsymbol{\alpha},\boldsymbol{\alpha})$总是实数.从长度的定义知, $\|\boldsymbol{\alpha}\|=0$当且仅当$\boldsymbol{\alpha}=0$.

Proposition 9.1 (范数不等式)

设V是实或复的内积空间, $\boldsymbol{\alpha},\boldsymbol{\beta}\in V$, c是任一常数 (实数或复数), 则

(1) $\|c\boldsymbol{\alpha}\|=|c|\,\|\boldsymbol{\alpha}\|$

(2)$|(\boldsymbol{\alpha},\boldsymbol{\beta})|\leq\|\boldsymbol{\alpha}\|\cdot\|\boldsymbol{\beta}\|$ $(Cauchy-Schwarz$不等式$)$

(3) $\|\boldsymbol{\alpha}+\boldsymbol{\beta}\|\leq\|\boldsymbol{\alpha}\|+\|\boldsymbol{\beta}\|$. $($三角不等式$)$

Proof (1) $\|c\boldsymbol{\alpha}\|^2=(c\boldsymbol{\alpha},c\boldsymbol{\alpha})=c\bar{c}\|\boldsymbol{\alpha}\|^2=|c|^2\|\boldsymbol{\alpha}\|^2$, 故 $\|c\boldsymbol{\alpha}\|=|c|\,\|\boldsymbol{\alpha}\|$.

(2)若$\boldsymbol{\alpha}=0$, 则 $(0,\boldsymbol{\beta})=(0+0,\boldsymbol{\beta})=2(0,\boldsymbol{\beta})$, 故 $(0,\boldsymbol{\beta})=0$.因此 (2)成立.

若$\alpha\neq 0$, 令$v=\boldsymbol{\beta}-\dfrac{(\boldsymbol{\beta},\boldsymbol{\alpha})}{\|\boldsymbol{\alpha}\|^2}\boldsymbol{\alpha}$则 $(\boldsymbol{v},\boldsymbol{\alpha})=0$, 且

$$
\begin{aligned}
0\leq\|\boldsymbol{v}\|^2&=\left(\boldsymbol{\beta}-\frac{(\boldsymbol{\beta},\boldsymbol{\alpha})}{\|\boldsymbol{\alpha}\|^2}\boldsymbol{\alpha},\boldsymbol{\beta}-\frac{(\boldsymbol{\beta},\boldsymbol{\alpha})}{\|\boldsymbol{\alpha}\|^2}\boldsymbol{\alpha}\right)\\
&=(\boldsymbol{\beta},\boldsymbol{\beta})-\frac{(\boldsymbol{\beta},\boldsymbol{\alpha})}{\|\boldsymbol{\alpha}\|^2}(\boldsymbol{\alpha},\boldsymbol{\beta})\\
&=\|\boldsymbol{\beta}\|^2-\frac{|(\boldsymbol{\alpha},\boldsymbol{\beta})|^2}{\|\boldsymbol{\alpha}\|^2}.
\end{aligned}
$$

由此即可得 (2).

(3)我们有

$\|\boldsymbol{\alpha}+\boldsymbol{\beta}\|^2=(\boldsymbol{\alpha}+\boldsymbol{\beta},\boldsymbol{\alpha}+\boldsymbol{\beta})=\|\boldsymbol{\alpha}\|^2+(\boldsymbol{\alpha},\boldsymbol{\beta})+(\boldsymbol{\beta},\boldsymbol{\alpha})+\|\boldsymbol{\beta}\|^2=\|\boldsymbol{\alpha}\|^2+\|\boldsymbol{\beta}\|^2+(\boldsymbol{\alpha},\boldsymbol{\beta})+\overline{(\boldsymbol{\alpha},\boldsymbol{\beta})}$.

由 (2)得$|\overline{(\boldsymbol{\alpha},\boldsymbol{\beta})}=|(\boldsymbol{\alpha},\boldsymbol{\beta})|\leq\|\boldsymbol{\alpha}\|\|\boldsymbol{\beta}\|$

故 $\|\boldsymbol{\alpha}+\boldsymbol{\beta}\|^2\leq\|\boldsymbol{\alpha}\|^2+\|\boldsymbol{\beta}\|^2+2\|\boldsymbol{\alpha}\|\|\boldsymbol{\beta}\|=(\|\boldsymbol{\alpha}\|+\|\boldsymbol{\beta}\|)^2$.

Definition 9.3 (夹角定义)

当V是实内积空间时,定义非零向量$\boldsymbol{\alpha},\boldsymbol{\beta}$的夹角$\theta$之余弦为 $\cos\theta=\dfrac{(\boldsymbol{\alpha},\boldsymbol{\beta})}{\|\boldsymbol{\alpha}\|\|\boldsymbol{\beta}\|}$.

当V是复内积空间时,定义非零向量$\boldsymbol{\alpha},\boldsymbol{\beta}$的夹角$\theta$之余弦为 $\cos\theta=\dfrac{|(\boldsymbol{\alpha},\boldsymbol{\beta})|}{\|\boldsymbol{\alpha}\|\|\boldsymbol{\beta}\|}$.

内积空间中两个向量$\boldsymbol{\alpha},\boldsymbol{\beta}$若适合 $(\boldsymbol{\alpha},\boldsymbol{\beta})=0$, 则称$\boldsymbol{\alpha}$与$\boldsymbol{\beta}$垂直或正交, 我们用记号$\alpha\perp\beta$来表示.

显然,零向量和任何向量都正交; 若$\boldsymbol{\alpha}$与$\boldsymbol{\beta}$正交, 则$\boldsymbol{\beta}$与$\boldsymbol{\alpha}$也正交; 两个非零向量$\boldsymbol{\alpha},\boldsymbol{\beta}$正交时夹角为$90^\circ$.

注意在定义夹角式中要使θ有意义, 必须保证$\cos\theta\leq 1$.而这就是柯西施瓦兹不等式保证.

Proposition 9.2 (勾股定理)

若$\boldsymbol{\alpha}$与$\boldsymbol{\beta}$正交, 则 $(\boldsymbol{\alpha},\boldsymbol{\beta})=(\boldsymbol{\beta},\boldsymbol{\alpha})=0$, 因此 $\|\boldsymbol{\alpha}+\boldsymbol{\beta}\|^2=\|\boldsymbol{\alpha}\|^2+\|\boldsymbol{\beta}\|^2$.

Proof $\|\boldsymbol{\alpha}+\boldsymbol{\beta}\|^2=(\boldsymbol{\alpha}+\boldsymbol{\beta},\boldsymbol{\alpha}+\boldsymbol{\beta})=\|\boldsymbol{\alpha}\|^2+(\boldsymbol{\alpha},\boldsymbol{\beta})+(\boldsymbol{\beta},\boldsymbol{\alpha})+\|\boldsymbol{\beta}\|^2=\|\boldsymbol{\alpha}\|^2+\|\boldsymbol{\beta}\|^2+(\boldsymbol{\alpha},\boldsymbol{\beta})+\overline{(\boldsymbol{\alpha},\boldsymbol{\beta})}.=\|\boldsymbol{\alpha}\|^2+\|\boldsymbol{\beta}\|^2$

9.2 内积与正交基

Definition 9.4 (度量矩阵 Gram 矩阵)

设V^n是欧氏空间 (酉空间), $\{\boldsymbol{v}_1,\boldsymbol{v}_2,\cdots,\boldsymbol{v}_n\}$是$V$的一组基　如果 $(\boldsymbol{v}_i,\boldsymbol{v}_j)=g_{ij}(i,j=1,2,\cdots,n)$

令$\boldsymbol{G}=\begin{pmatrix} g_{11} & g_{12} & \cdots & g_{1n} \\ g_{21} & g_{22} & \cdots & g_{2n} \\ \vdots & \vdots & & \vdots \\ g_{n1} & g_{n2} & \cdots & g_{nn} \end{pmatrix}=\begin{pmatrix} (\boldsymbol{v}_1,\boldsymbol{v}_1) & (\boldsymbol{v}_1,\boldsymbol{v}_2) & \cdots & (\boldsymbol{v}_1,\boldsymbol{v}_n) \\ (\boldsymbol{v}_2,\boldsymbol{v}_1) & (\boldsymbol{v}_2,\boldsymbol{v}_2) & \cdots & (\boldsymbol{v}_2,\boldsymbol{v}_n) \\ \vdots & \vdots & & \vdots \\ (\boldsymbol{v}_n,\boldsymbol{v}_1) & (\boldsymbol{v}_n,\boldsymbol{v}_2) & \cdots & (\boldsymbol{v}_n,\boldsymbol{v}_n) \end{pmatrix}$

称为基向量$\{v_1,v_2,\cdots,v_n\}$的$Gram$(格列姆)矩阵或内积空间V在给定基下的度量矩阵

Proposition 9.3 (内积空间上内积的表示)

设V^n是欧氏空间 (酉空间), $\{\boldsymbol{v}_1,\boldsymbol{v}_2,\cdots,\boldsymbol{v}_n\}$是$V$的一组基　设度量矩阵为$\boldsymbol{G}=(g_{ij})_{n\times n}$

设$\boldsymbol{\alpha}=a_1\boldsymbol{v}_1+a_2\boldsymbol{v}_2+\cdots+a_n\boldsymbol{v}_n$　　$\boldsymbol{\beta}=b_1\boldsymbol{v}_1+b_2\boldsymbol{v}_2+\cdots+b_n\boldsymbol{v}_n$

当V是欧氏空间时$\Longrightarrow(\boldsymbol{\alpha},\boldsymbol{\beta})=x'\boldsymbol{G}y$其中$\boldsymbol{x},\boldsymbol{y}$分别是向量$\boldsymbol{\alpha},\boldsymbol{\beta}$在给定基下的坐标向量.

当V是酉空间时$\Longrightarrow(\boldsymbol{\alpha},\boldsymbol{\beta})=\boldsymbol{x}'\boldsymbol{G}\overline{\boldsymbol{y}}$其中$\boldsymbol{x},\boldsymbol{y}$分别是向量$\boldsymbol{\alpha},\boldsymbol{\beta}$在给定基下的坐标向量.

Proof 当V是欧氏空间时

$$(\boldsymbol{\alpha},\boldsymbol{\beta})=\left(\sum_{i=1}^n a_i\boldsymbol{v}_i,\sum_{j=1}^n b_j\boldsymbol{v}_j\right)=\sum_{i,j=1}^n a_ig_{ij}b_j\Longrightarrow(\boldsymbol{\alpha},\boldsymbol{\beta})=(a_1,a_2,\cdots,a_n)\begin{pmatrix} g_{11} & g_{12} & \cdots & g_{1n} \\ g_{21} & g_{22} & \cdots & g_{2n} \\ \vdots & \vdots & & \vdots \\ g_{n1} & g_{n2} & \cdots & g_{nn} \end{pmatrix}\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

$\Longrightarrow(\boldsymbol{\alpha},\boldsymbol{\beta})=x'\boldsymbol{G}y$其中$\boldsymbol{x},\boldsymbol{y}$分别是向量$\boldsymbol{\alpha},\boldsymbol{\beta}$在给定基下的坐标向量.

当V是酉空间时

$$(\boldsymbol{\alpha},\boldsymbol{\beta})=\left(\sum_{i=1}^n a_i\boldsymbol{v}_i,\sum_{j=1}^n b_j\boldsymbol{v}_j\right)=\sum_{i,j=1}^n a_ig_{ij}\overline{b_j}\Longrightarrow(\boldsymbol{\alpha},\boldsymbol{\beta})=(a_1,a_2,\cdots,a_n)\begin{pmatrix} g_{11} & g_{12} & \cdots & g_{1n} \\ g_{21} & g_{22} & \cdots & g_{2n} \\ \vdots & \vdots & & \vdots \\ g_{n1} & g_{n2} & \cdots & g_{nn} \end{pmatrix}\begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \\ \vdots \\ \bar{b}_n \end{pmatrix}$$

$\Longrightarrow(\boldsymbol{\alpha},\boldsymbol{\beta})=\boldsymbol{x}'\boldsymbol{G}\overline{\boldsymbol{y}}$其中$\boldsymbol{x},\boldsymbol{y}$分别是向量$\boldsymbol{\alpha},\boldsymbol{\beta}$在给定基下的坐标向量.

Definition 9.5 (正交基与标准正交基定义)

设$\{e_1,e_2,\cdots,e_n\}$是n维内积空间V的一组基.

若$e_i\bot e_j$对一切$i\neq j$成立, 则称这组基是V的一组正交基. 又若V的一组正交基中每个向量的长度等于1, 则称这组正交基为标准正交

Proposition 9.4 (正交向量组的性质与正交基)

1.R^n中, 正交向量组一定是线性无关的

2.R^n中, n个向量组成的正交向量组一定是R^n的一个基, 称为正交基, n个单位向量组成的正交向量组称为R^n的一个标准正交基

3.在欧几里得空间R^n中, 如果向量$\boldsymbol{\alpha}$与R^n的一个正交基$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n$的每个向量都正交, 那么$\boldsymbol{\alpha}=0$

Proof 1.设$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$是正交向量组.

设$k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_s\boldsymbol{\alpha}_s=0$, 则$(k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_s\boldsymbol{\alpha}_s,\boldsymbol{\alpha}_i)=(0,\boldsymbol{\alpha}_i),1\leqslant i\leqslant s$

由于$(\boldsymbol{\alpha}_j,\boldsymbol{\alpha}_i)=0$, 当$j\neq i$, 因此由上式得$k_i(\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_i)=0$

由于$\boldsymbol{\alpha}_i\neq 0$, 因此$(\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_i)>0$, 从而由上式得出, $k_i=0$, 其中$i=1,2,\cdots,s$, 因此$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$线性无关

2.由1.得

3.设$\alpha = a_1\beta_1 + a_2\beta_2 + \cdots + a_n\beta_n$, 则由$\left(\alpha, \beta_j\right) = 0$, 得$0 = \left(\alpha, \beta_j\right) = \left(\sum_{i=1}^n a_i\beta_i, \beta_j\right) = \sum_{i=1}^n a_i\left(\beta_i, \beta_j\right) = a_j\left(\beta_j, \beta_j\right)$
由于$\left(\beta_j, \beta_j\right) \neq 0$, 因此$a_j = 0, j = 1, 2, \cdots, n$, 从而$\alpha = 0$

Theorem 9.2.1 (施密特正交化定理)

设$\alpha_1, \alpha_2, \cdots, \alpha_s$是欧几里得空间$\mathrm{R}^n$中一个线性无关的向量组, 令

$$\beta_1 = \alpha_1$$
$$\beta_2 = \alpha_2 - \frac{(\alpha_2, \beta_1)}{(\beta_1, \beta_1)}\beta_1$$
$$\cdots\cdots\cdots$$
$$\beta_s = \alpha_s - \sum_{j=1}^{s-1}\frac{\left(\alpha_s, \beta_j\right)}{\left(\beta_j, \beta_j\right)}\beta_j,$$

则$\beta_1, \beta_2, \cdots, \beta_s$是正交向量组, 并且$\beta_1, \beta_2, \cdots, \beta_s$与$\alpha_1, \alpha_2, \cdots, \alpha_s$等价

Proof 对线性无关的向量组所含向量的个数作数学归纳法

$s = 1$时, 令$\beta_1 = \alpha_1$, 由于$\alpha_1 \neq 0$, 因此β_1是正交向量组, 且β_1与α_1等价

假设$s = k$时命题为真.现在来看$s = k + 1$的情形

由于$\beta_{k+1} = \alpha_{k+1} - \sum_{j=1}^{k}\frac{\left(\alpha_{k+1}, \beta_j\right)}{\left(\beta_j, \beta_j\right)}\beta_j,$

因此当$1 \leqslant i \leqslant k$时, 有$(\beta_{k+1}, \beta_i) = (\alpha_{k+1}, \beta_i) - \sum_{j=1}^{k}\frac{\left(\alpha_{k+1}, \beta_j\right)}{\left(\beta_j, \beta_j\right)}\left(\beta_j, \beta_i\right) = (\alpha_{k+1}, \beta_i) - \frac{(\alpha_{k+1}, \beta_i)}{(\beta_i, \beta_i)}(\beta_i, \beta_i) = 0.$

这表明β_{k+1}与β_i正交$(i = 1, 2, \cdots, k)$.

从式子以及归纳假设可以看出, β_{k+1}可以由$\alpha_1, \alpha_2, \cdots, \alpha_k, \alpha_{k+1}$线性表出, 并且表出式中$\alpha_{k+1}$的系数为1, 因此$\beta_{k+1} \neq 0$

于是$\beta_1, \cdots, \beta_k, \beta_{k+1}$是正交向量组

从式以及归纳假设立即得出$\beta_1, \cdots, \beta_k, \beta_{k+1}$与$\alpha_1, \alpha_k, \cdots, \alpha_{k+1}$等价, 因此当$s = k + 1$时, 命题也为真

根据数学归纳法原理, 命题为真

注 给出了在欧几里得空间R^n中从一个线性无关的向量组$\alpha_1, \alpha_2, \cdots, \alpha_s$出发, 构造出与它等价的一个正交向量组的方法

这种方法称为施密特$(Schmidt)$正交化过程

只要将$\beta_1, \beta_2, \cdots, \beta_s$中每个向量单位化, 即$\eta_i = \frac{1}{|\beta_i|}\beta_i, \quad i = 1, 2, \cdots, s$, 则$\eta_1, \eta_2, \cdots, \eta_s$是与$\alpha_1, \alpha_2, \cdots, \alpha_s$等价的正交单位向量组

R^n中, 如果给了一个基$\alpha_1, \alpha_2, \cdots, \alpha_n$, 则先经过施密特正交化过程, 后经单位化, 得到向量组$\eta_1, \eta_2, \cdots, \eta_n$就是$\mathrm{R}^n$的一个标准正交基

注 现在我们设V是一个n维欧氏空间, $\alpha_1, \alpha_2, \cdots, \alpha_n$是它的一组基, 我们构造正交基$\beta_1, \beta_2, \cdots, \beta_n$(它和$\alpha_1, \alpha_2, \cdots, \alpha_n$等价)

如下 : $\begin{cases} \beta_1 = \alpha_1 \\ \beta_2 = \alpha_2 - \dfrac{(\alpha_2, \beta_1)}{(\beta_1, \beta_1)}\beta_1 \\ \quad\vdots \\ \beta_n = \alpha_n - \dfrac{(\alpha_n, \beta_1)}{(\beta_1, \beta_1)}\beta_1 - \cdots - \dfrac{(\alpha_n, \beta_{n-1})}{(\beta_{n-1}, \beta_{n-1})}\beta_{n-1} \end{cases}$

$$\Longrightarrow \begin{pmatrix} \beta_1 & \beta_2 & \cdots & \beta_n \end{pmatrix} = \begin{pmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_n \end{pmatrix}\begin{pmatrix} 1 & * & * & * \\ & 1 & * & * \\ & & \ddots & \vdots \\ & & & 1 \end{pmatrix}$$

不难得到$\beta_1, \beta_2, \cdots, \beta_n$(它和$\alpha_1, \alpha_2, \cdots, \alpha_n$等价)

从中把$\alpha_1,\alpha_2,\cdots,\alpha_n$解出来，笼统地写为$\begin{cases}\alpha_1=\beta_1\\ \alpha_2=a_{12}\beta_1+\beta_2\\ \quad\vdots\\ \alpha_n=a_{1n}\beta_1+\cdots+a_{n-1,n}\beta_{n-1}+\beta_n\end{cases}$

$$\Longrightarrow\begin{pmatrix}\alpha_1 & \alpha_2 & \cdots & \alpha_n\end{pmatrix}=\begin{pmatrix}\beta_1 & \beta_2 & \cdots & \beta_n\end{pmatrix}\begin{pmatrix}1 & * & * & *\\ & 1 & * & *\\ & & \ddots & \vdots\\ & & & 1\end{pmatrix}$$

把$\beta_1,\beta_2,\cdots,\beta_n$单位化一下,我们记$\eta_i=\dfrac{\beta_i}{|\beta_i|}(i=1,2,\cdots,n)$,就有$\begin{cases}\alpha_1=b_{11}\eta_1\\ \alpha_2=b_{12}\eta_1+b_{22}\eta_2\\ \quad\vdots\\ \alpha_n=b_{1n}\eta_1+\cdots+b_{n-1,n}\eta_{n-1}+b_{nn}\eta_n\end{cases}$ 其中$b_{ii}=|\beta_i|>0$

$$\Longrightarrow \text{写成分块矩阵乘法的样子}:(\alpha_1,\alpha_2,\cdots,\alpha_n)=(\eta_1,\eta_2,\cdots,\eta_n)\begin{pmatrix}b_{11} & b_{12} & \cdots & b_{1n}\\ & b_{22} & \cdots & b_{2n}\\ & & \ddots & \vdots\\ & & & b_{nn}\end{pmatrix}$$

Definition 9.6 (正交补空间)

设U是内积空间V的子空间.令$U^{\perp}=\{\boldsymbol{v}\in V\mid(\boldsymbol{v},U)=0\}$这里$(v,U)=0$表示对一切$u\in U$，均有$(\boldsymbol{v},\boldsymbol{u})=0$

容易验证$U^{\perp}$是V的子空间,称为U的正交补空间

Proposition 9.5

设V是n维内积空间,U是V的子空间,则

(1)$V=U\oplus U^{\perp}$

(2)U上的任一组标准正交基均可扩张为V上的标准正交基.

Proof (1)若$\boldsymbol{x}\in U\cap U^{\perp}$,则$(\boldsymbol{x},\boldsymbol{x})=0$,因此$\boldsymbol{x}=0$,即$U\cap U^{\perp}=0$ 另一方面,存在U的一组标准正交基$\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_m\}$.

此时对任意的$\boldsymbol{v}\in V$，令$\boldsymbol{u}=(\boldsymbol{v},\boldsymbol{e}_1)\boldsymbol{e}_1+(\boldsymbol{v},\boldsymbol{e}_2)\boldsymbol{e}_2+\cdots+(\boldsymbol{v},\boldsymbol{e}_m)\boldsymbol{e}_m$则$\boldsymbol{u}\in U$

又令$\boldsymbol{w}=\boldsymbol{v}-\boldsymbol{u}$则对任一$\boldsymbol{e}_i(i=1,2,\cdots,m)$,有$(\boldsymbol{w},\boldsymbol{e}_i)=(\boldsymbol{v},\boldsymbol{e}_i)-(\boldsymbol{u},\boldsymbol{e}_i)=(\boldsymbol{v},\boldsymbol{e}_i)-(\boldsymbol{v},\boldsymbol{e}_i)=0$

$\Longrightarrow$ 因此$\boldsymbol{w}\in U^{\perp}$,而$\boldsymbol{v}=\boldsymbol{u}+\boldsymbol{w}$,这就证明了$V=U\oplus U^{\perp}$.

(2)设$\{e_1,e_2,\cdots,e_m\}$是U上任一组标准正交基,$\{e_{m+1},\cdots,e_n\}$是$U^{\perp}$上任一组标准正交基,则显然$\{e_1,e_2,\cdots,e_n\}$是V的一组标准正交

Definition 9.7 (子空间的正交和)

设V是n维内积空间,$V_i(i=1,2,\cdots,k)$是V的子空间.

如果对任意的$\boldsymbol{\alpha}\in V_i$和任意的$\boldsymbol{\beta}\in V_j(j\neq i)$均有$(\boldsymbol{\alpha},\boldsymbol{\beta})=0$,则称子空间$V_i$和$V_j$正交.

若$V=V_1+V_2+\cdots+V_k$且V_i两两正交,则称V是$V_i(i=1,2,\cdots,k)$的正交和,记为$V=V_1\perp V_2\perp\cdots\perp V_k$

Proposition 9.6

正交和必为直和且任一V_i和其余子空间的和正交.

Proof 对任意的$v_i \in V_i$和$\sum\limits_{j\neq i} v_j\,(v_j \in V_j)$, $\left(\boldsymbol{v}_i, \sum\limits_{j\neq i} \boldsymbol{v}_j\right) = \sum\limits_{j\neq i} (\boldsymbol{v}_i, \boldsymbol{v}_j) = 0$因此后一个结论成立.

任取$v \in V_i \cap \left(\sum\limits_{j\neq i} V_j\right)$，则由上述论证可得$(v, v) = 0$，故$\boldsymbol{v} = 0$, 从而$V_i \cap \left(\sum\limits_{j\neq i} V_j\right) = 0$, 即正交和必为直和.因此正交和也称为正交直和。

Definition 9.8 (正交投影)

设$V = V_1 \perp V_2 \perp \cdots \perp V_k$

定义V上的线性变换$\boldsymbol{E}_i (i = 1, 2, \cdots, k)$如下：若$\boldsymbol{v} = \boldsymbol{v}_1 + \cdots + \boldsymbol{v}_i + \cdots + \boldsymbol{v}_k\,(\boldsymbol{v}_i \in V_i)$, 令$\boldsymbol{E}_i(\boldsymbol{v}) = \boldsymbol{v}_i$

容易验证$\boldsymbol{E}_i$是V上的线性变换且$\boldsymbol{E}_i^2 = \boldsymbol{E}_i, \boldsymbol{E}_i\boldsymbol{E}_j = 0 (i \neq j), \boldsymbol{E}_1 + \boldsymbol{E}_2 + \cdots + \boldsymbol{E}_k = \boldsymbol{I}_V$

线性变换$\boldsymbol{E}_i$称为V到V_i的正交投影.

Proposition 9.7

设U是内积空间V的子空间, $V = U \perp U^{\perp}$. $\boldsymbol{E}$是V到U的正交投影, 则对任意的$\boldsymbol{\alpha}, \boldsymbol{\beta} \in V$, 都有$(\boldsymbol{E}(\alpha), \boldsymbol{\beta}) = (\alpha, E(\boldsymbol{\beta}))$

Proof 设$\boldsymbol{\alpha} = \boldsymbol{u}_1 + \boldsymbol{w}_1, \boldsymbol{\beta} = \boldsymbol{u}_2 + \boldsymbol{w}_2$, 其中$\boldsymbol{u}_1, \boldsymbol{u}_2 \in U, \boldsymbol{w}_1, \boldsymbol{w}_2 \in U^{\perp}$, 则$\boldsymbol{E}(\alpha) = \boldsymbol{u}_1, \boldsymbol{E}(\boldsymbol{\beta}) = \boldsymbol{u}_2$, 所以

$(\boldsymbol{E}(\alpha), \boldsymbol{\beta}) = (\boldsymbol{u}_1, \boldsymbol{u}_2 + \boldsymbol{w}_2) = (\boldsymbol{u}_1, \boldsymbol{u}_2) + (\boldsymbol{u}_1, \boldsymbol{w}_2) = (\boldsymbol{u}_1, \boldsymbol{u}_2)$

$(\alpha, \boldsymbol{E}(\boldsymbol{\beta})) = (\boldsymbol{u}_1 + \boldsymbol{w}_1, \boldsymbol{u}_2) = (\boldsymbol{u}_1, \boldsymbol{u}_2) + (\boldsymbol{w}_1, \boldsymbol{u}_2) = (\boldsymbol{u}_1, \boldsymbol{u}_2)$

由此即得结论.

9.3 伴随

Definition 9.9 (伴随算子)

设φ是内积空间V上的线性算子

若存在V上的线性算子φ^*, 使等式 $(\boldsymbol{\varphi}(\boldsymbol{\alpha}),\boldsymbol{\beta})=(\boldsymbol{\alpha},\varphi^*(\boldsymbol{\beta}))$ 对一切$\alpha,\beta\in V$成立, 则称φ^*是φ的伴随算子, 简称为φ的伴随.

Theorem 9.3.1 (有限维内积空间上的伴随算子存在唯一性)

设V是n维内积空间, φ是V上的线性变换, 则存在V上唯一的线性变换φ^*使对一切$\alpha,\beta\in V$, 成立 $(\boldsymbol{\varphi}(\boldsymbol{\alpha}),\boldsymbol{\beta})=(\boldsymbol{\alpha},\boldsymbol{\varphi}^*(\boldsymbol{\beta}))$

Proof 证明唯一性.

若φ是V上的线性变换且 $(\boldsymbol{\varphi}(\boldsymbol{\alpha}),\boldsymbol{\beta})=\left(\boldsymbol{\alpha},\boldsymbol{\varphi}^\sharp(\boldsymbol{\beta})\right)$ 对一切$\boldsymbol{\alpha},\boldsymbol{\beta}$成立 同理也有 $(\boldsymbol{\varphi}(\boldsymbol{\alpha}),\boldsymbol{\beta})=(\boldsymbol{\alpha},\boldsymbol{\varphi}^*(\boldsymbol{\beta}))$ 对一切$\boldsymbol{\alpha},\boldsymbol{\beta}$成立

则 $\left(\boldsymbol{\alpha},\boldsymbol{\varphi}^\sharp(\boldsymbol{\beta})\right)=(\boldsymbol{\alpha},\boldsymbol{\varphi}^*(\boldsymbol{\beta}))$ 对一切$\boldsymbol{\alpha}\in V$成立即 $\left(\boldsymbol{\alpha},\boldsymbol{\varphi}^\sharp(\boldsymbol{\beta})-\ \boldsymbol{\varphi}^*(\boldsymbol{\beta})\right)=0$对一切$\alpha$成立

特别, 对$\alpha=\varphi^\sharp(\beta)-\varphi^*(\beta)$也成立。由内积正定性即知$\varphi^\sharp(\boldsymbol{\beta})-\varphi^*(\boldsymbol{\beta})=0$, 即$\boldsymbol{\varphi}^\sharp(\boldsymbol{\beta})=\varphi^*(\boldsymbol{\beta})$.而$\boldsymbol{\beta}$是任意的, 故有$\varphi^\sharp=\varphi^*$.

证明存在性

现设V是n维内积空间 (不妨设之为酉空间), 取V的一组标准正交基 $\{e_1,e_2,\cdots,e_n\}$

假定φ是V上的线性变换且它在这组基下的表示矩阵是$\boldsymbol{A}=\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&&\vdots\\a_{n1}&a_{n2}&\cdots&a_{nn}\end{pmatrix}$

设$\boldsymbol{\alpha}=a_1\boldsymbol{e}_1+a_2\boldsymbol{e}_2+\cdots+a_n\boldsymbol{e}_n$ $\quad\boldsymbol{\beta}=b_1\boldsymbol{e}_1+b_2\boldsymbol{e}_2+\cdots+b_n\boldsymbol{e}_n$.

记$\boldsymbol{x}=(a_1,a_2,\cdots,a_n)'$, $\boldsymbol{y}=(b_1,b_2,\cdots,b_n)'$ 分别是$\boldsymbol{\alpha},\boldsymbol{\beta}$的坐标向量

则 $(\boldsymbol{\varphi}(\boldsymbol{\alpha}),\boldsymbol{\beta})=(\boldsymbol{A}\boldsymbol{x})'\overline{\boldsymbol{y}}=\boldsymbol{x}'\boldsymbol{A}'\overline{\boldsymbol{y}}=\boldsymbol{x}'\left(\overline{\boldsymbol{A}}'\boldsymbol{y}\right)$

记矩阵$\overline{\boldsymbol{A}}'$在$V$上定义的线性变换为$\boldsymbol{\psi}$

即对任意的$\boldsymbol{\beta}=b_1\boldsymbol{e}_1+b_2\boldsymbol{e}_2+\cdots+b_n\boldsymbol{e}_n$.定义$\boldsymbol{\psi}(\boldsymbol{\beta})=c_1\boldsymbol{e}_1+c_2\boldsymbol{e}_2+\cdots+c_n\boldsymbol{e}_n$, 其中$c_i=\bar{a}_{1i}b_1+\bar{a}_{2i}b_2+\cdots+\bar{a}_{ni}b_n(i=1,2,\cdots,n)$

即$\boldsymbol{\psi}(\boldsymbol{\beta})$的坐标向量$\begin{pmatrix}c_1\\c_2\\\vdots\\c_n\end{pmatrix}=\overline{\boldsymbol{A}}'\boldsymbol{y}=\begin{pmatrix}\bar{a}_{11}&\bar{a}_{21}&\cdots&\bar{a}_{n1}\\\bar{a}_{12}&\bar{a}_{22}&\cdots&\bar{a}_{n2}\\\vdots&\vdots&&\vdots\\\bar{a}_{1n}&\bar{a}_{2n}&\cdots&\bar{a}_{nn}\end{pmatrix}\begin{pmatrix}b_1\\b_2\\\vdots\\b_n\end{pmatrix}$

于是表明 $(\boldsymbol{\varphi}(\boldsymbol{\alpha}),\boldsymbol{\beta})=(\boldsymbol{\alpha},\boldsymbol{\psi}(\boldsymbol{\beta}))$对一切$\alpha,\beta\in V$成立.

Theorem 9.3.2 (线性算子与伴随算子在基下表示矩阵的联系)

设V是n维内积空间, $\{e_1,e_2,\cdots,e_n\}$ 是V的一组标准正交基 若V上的线性算子$\boldsymbol{\varphi}$在这组基下的表示矩阵为$\boldsymbol{A}=(a_{ij})$

则如果V是西空间, 那么$\boldsymbol{\varphi}^*$在同一组基下的表示矩阵为$\overline{\boldsymbol{A}}'=(\bar{a}_{ij})'$, 即$\boldsymbol{A}$的共轭转置

如果V是欧氏空间, 那么$\boldsymbol{\varphi}^*$的表示矩阵为$\boldsymbol{A}'$, 即$\boldsymbol{A}$的转置。

Proposition 9.8

设V是有限维内积空间, 若φ及ψ是V上的线性变换, c为常数, 则

(1)$(\boldsymbol{\varphi}+\boldsymbol{\psi})^*=\boldsymbol{\varphi}^*+\boldsymbol{\psi}^*$

(2)$(c\boldsymbol{\varphi})^*=\bar{c}\varphi^*$

(3)$(\boldsymbol{\varphi}\boldsymbol{\psi})^*=\boldsymbol{\psi}^*\boldsymbol{\varphi}^*$

(4) $(\boldsymbol{\varphi}^*)^*=\varphi$.

Proof 证明由矩阵和线性变换的对应关系及矩阵共轭转置的性质即得.

Proposition 9.9

设V是n维内积空间,φ是V上的线性算子.

(1)若U是φ的不变子空间，则$U^{\perp}$是φ^*的不变子空间

(2)若φ的全体特征值为$\lambda_1,\lambda_2,\cdots,\lambda_n$,则$\varphi^*$的全体特征值为$\bar{\lambda}_1,\bar{\lambda}_2,\cdots,\bar{\lambda}_n$.

Proof (1) 设$\boldsymbol{\alpha}\in U,\boldsymbol{\beta}\in U^{\perp}$，因为$(\boldsymbol{\alpha},\boldsymbol{\varphi}^*(\boldsymbol{\beta}))=(\boldsymbol{\varphi}(\boldsymbol{\alpha}),\boldsymbol{\beta})=0$所以$U^{\perp}$是$\varphi^*$的不变子空间

(2) 取V的一组标准正交基，记$\boldsymbol{A}$是$\boldsymbol{\varphi}$的表示矩阵,则无论V是西空间还是欧氏空间,φ^*的表示矩阵总可写为$\overline{\boldsymbol{A}}'$
由假设$|\lambda\boldsymbol{I}_n-\boldsymbol{A}|=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n)$则容易验证$\left|\lambda I_n-\overline{\boldsymbol{A}}'\right|=\left(\lambda-\bar{\lambda}_1\right)\left(\lambda-\bar{\lambda}_2\right)\cdots\left(\lambda-\bar{\lambda}_n\right)$故结论成立.

9.4 正交变换与矩阵与酉变换与酉矩阵

Definition 9.10 (保积同构 (正交性矩阵的的几何意义))

设V与U是域$\mathbb{K}$上的内积空间,$\mathbb{K}$是实数域或复数域,φ是$V \to U$的线性映射.
若对任意的$\boldsymbol{x}, \boldsymbol{y} \in V$, 有 $(\varphi(x), \varphi(y)) = (x, y)$则称$\varphi$是$V \to U$的保持内积的线性映射.
又若φ作为线性映射是同构, 则称φ是内积空间V到U上的保积同构.

注 注 (1)在不引起误解的情况下，我们常把内积空间的保积同构就称为同构.
(2)保持内积的线性映射一定是单映射, 这是因为 $\| \varphi(x) \| = \| x \|$, 从 $\| \boldsymbol{\varphi}(\boldsymbol{x}) \| = 0$得到 $\| x \| = 0$, 故$x = 0$
(3)容易证明保持内积的同构关系是一个等价关系, 读者可自己证明之.

Proposition 9.10

若φ是内积空间V到内积空间U的保持范数的线性映射, 则φ保持内积

Proof 向量的范数可以用内积表示, 反过来内积也可以用范数来表示.
设$\boldsymbol{x}, \boldsymbol{y}$是$V$中的任意两个向量, 则
$\| \boldsymbol{x}+\boldsymbol{y} \|^2 = (\boldsymbol{x}+\boldsymbol{y}, \boldsymbol{x}+\boldsymbol{y}) = (\boldsymbol{x}, \boldsymbol{x}) + (\boldsymbol{y}, \boldsymbol{y}) + (\boldsymbol{x}, \boldsymbol{y}) + (\boldsymbol{y}, \boldsymbol{x})$
$\| \boldsymbol{x}-\boldsymbol{y} \|^2 = (\boldsymbol{x}-\boldsymbol{y}, \boldsymbol{x}-\boldsymbol{y}) = (\boldsymbol{x}, \boldsymbol{x}) + (\boldsymbol{y}, \boldsymbol{y}) - (\boldsymbol{x}, \boldsymbol{y}) - (\boldsymbol{y}, \boldsymbol{x})$
故 $\| \boldsymbol{x}+\boldsymbol{y} \|^2 - \| \boldsymbol{x}-\boldsymbol{y} \|^2 = 2(\boldsymbol{x}, \boldsymbol{y}) + 2\overline{(\boldsymbol{x}, \boldsymbol{y})}$
另一方面
$\| \boldsymbol{x}+\mathrm{i}\boldsymbol{y} \|^2 = (\boldsymbol{x}+\mathrm{i}\boldsymbol{y}, \boldsymbol{x}+\mathrm{i}\boldsymbol{y}) = (\boldsymbol{x}, \boldsymbol{x}) + (\boldsymbol{y}, \boldsymbol{y}) + \mathrm{i}(\boldsymbol{y}, \boldsymbol{x}) - \mathrm{i}(\boldsymbol{x}, \boldsymbol{y})$
$\| \boldsymbol{x}-\mathrm{i}\boldsymbol{y} \|^2 = (\boldsymbol{x}-\mathrm{i}\boldsymbol{y}, \boldsymbol{x}-\mathrm{i}\boldsymbol{y}) = (\boldsymbol{x}, \boldsymbol{x}) + (\boldsymbol{y}, \boldsymbol{y}) - \mathrm{i}(\boldsymbol{y}, \boldsymbol{x}) + \mathrm{i}(\boldsymbol{x}, \boldsymbol{y})$
故 $\| \boldsymbol{x}+\mathrm{i}\boldsymbol{y} \|^2 - \| \boldsymbol{x}-\mathrm{i}\boldsymbol{y} \|^2 = -2\mathrm{i}(\boldsymbol{x}, \boldsymbol{y}) + 2\mathrm{i}\overline{(\boldsymbol{x}, \boldsymbol{y})}$
得 $(\boldsymbol{x}, \boldsymbol{y}) = \frac{1}{4} \| \boldsymbol{x}+\boldsymbol{y} \|^2 - \frac{1}{4} \| \boldsymbol{x}-\boldsymbol{y} \|^2 + \frac{\mathrm{i}}{4} \| \boldsymbol{x}+i\boldsymbol{y} \|^2 - \frac{\mathrm{i}}{4} \| \boldsymbol{x}-i\boldsymbol{y} \|^2$ 由此即可得到结论.
注 注我们仅对复空间进行了讨论，事实上对实空间，可得下列等式：$(\boldsymbol{x}, \boldsymbol{y}) = \frac{1}{4} \| \boldsymbol{x}+\boldsymbol{y} \|^2 - \frac{1}{4} \| \boldsymbol{x}-\boldsymbol{y} \|^2$
因此对实空间结论也正确.由于保持内积与保持范数的等价性, 保持内积的同构也称为保范同构或保距同构.

Theorem 9.4.1 (保积同构的判定准则与几何性质)

设V与U都是n维内积空间 (同为实空间或同为复空间), 若φ是$V \to U$的线性映射, 则下列命题等价 :
(1)φ保持内积;
(2)φ是保积同构;
(3) φ将V的任一组标准正交基变成U的一组标准正交基
(4) φ将V的某一组标准正交基变成U的一组标准正交基.

Proof (1) $\Rightarrow$ (2) : φ保持内积, 因此φ为单映射.由维数公式可得 $\dim\mathrm{Im}\boldsymbol{\varphi} = \dim V = n$因此 $\mathrm{Im}\varphi = U$, 即$\boldsymbol{\varphi}$是映上的, 故为同构.

(2) $\Rightarrow$ (3) : 设 $\{e_1, e_2, \cdots, e_n\}$ 是V的任意一组标准正交基.由于φ保持内积, 故对$i \neq j$, 有
$(\boldsymbol{\varphi}(\boldsymbol{e}_i), \boldsymbol{\varphi}(\boldsymbol{e}_j)) = (\boldsymbol{e}_i, \boldsymbol{e}_j) = 0$又 $(\varphi(e_i), \varphi(e_i)) = (e_i, e_i) = 1$
$\Longrightarrow \{\varphi(e_1) \sim \varphi(e_n)\}$ 为正交单位向量组 $\Longrightarrow \{\varphi(e_1), \varphi(e_2), \cdots, \varphi(e_n)\}$ 是U的标准正交基.

(3) $\Longrightarrow$ (4) 显然

(4) $\Rightarrow$ (1)：设 $\{e_1, e_2, \cdots, e_n\}$ 是V的标准正交基且 $\{\boldsymbol{\varphi}(e_1), \varphi(e_2), \cdots,, \boldsymbol{\varphi}(\boldsymbol{e}_n)\}$ 是U的标准正交基.

假定$\boldsymbol{u}=\sum_{i=1}^{n}a_i\boldsymbol{e}_i,\boldsymbol{v}=\sum_{i=1}^{n}b_i\boldsymbol{e}_i$则$\boldsymbol{\varphi}(\boldsymbol{u})=\sum_{i=1}^{n}a_i\boldsymbol{\varphi}(\boldsymbol{e}_i),\boldsymbol{\varphi}(\boldsymbol{v})=\sum_{i=1}^{n}b_i\boldsymbol{\varphi}(\boldsymbol{e}_i)$

$$\Longrightarrow \begin{aligned}(\boldsymbol{\varphi}(\boldsymbol{u}),\boldsymbol{\varphi}(\boldsymbol{v}))&=\left(\sum_{i=1}^{n}a_i\boldsymbol{\varphi}(\boldsymbol{e}_i),\sum_{i=1}^{n}b_i\boldsymbol{\varphi}(\boldsymbol{e}_i)\right)\\&=a_1\bar{b}_1+a_2\bar{b}_2+\cdots+a_n\bar{b}_n\\&=(\boldsymbol{u},\boldsymbol{v})\end{aligned}$$

Corollary 9.1

两个有限维内积空间V与U(同为实空间或同为复空间)同构的充分必要条件是它们有相同的维数.

Proof 只需证明充分性.设$\{e_1,e_2,\cdots,e_n\}$是V的标准正交基,$\{f_1,f_2,\cdots,f_n\}$是U的标准正交基
令φ是$V\to U$的线性映射:$\varphi(e_i)=f_i$,则φ将标准正交基变为标准正交基,由定理可知φ是保积同构.

Definition 9.11 (正交算子与酉算子与正交矩阵与酉矩阵)

设V是欧氏空间,若φ是V上保持内积的线性变换,则称φ为V上的正交变换或正交算子
若U是酉空间,则U上保持内积的线性变换称为酉变换或酉算子
设$\boldsymbol{A}$是n阶实方阵,若$\boldsymbol{A}'=\boldsymbol{A}^{-1}$,则称$\boldsymbol{A}$是正交矩阵.又若$\boldsymbol{C}$是$n$阶复方阵且$\overline{\boldsymbol{C}}'=\boldsymbol{C}^{-1}$,则称$\boldsymbol{C}$是酉矩阵.

注 显然正交变换及酉变换都是可逆线性变换(因为把基变为基)
正交变换可定义为把欧氏空间中一组标准正交基变成标准正交基的线性变换.酉变换也类似.

Theorem 9.4.2 (正交算子与酉算子的等价几何刻画)

设φ是欧氏空间或酉空间V上的线性变换
则φ是正交变换或酉变换$\Longleftrightarrow\varphi$非异,且$\varphi^{-1}=\varphi^*$即$\varphi\varphi^*=\varphi^*\varphi=I_V$

Proof 设φ是欧氏空间V上的正交变换,则对V中的任意向量$\boldsymbol{\alpha},\boldsymbol{\beta}$,有$(\varphi(\boldsymbol{\alpha}),\boldsymbol{\beta})=\left(\varphi(\boldsymbol{\alpha}),\varphi\left(\varphi^{-1}(\boldsymbol{\beta})\right)\right)=\left(\boldsymbol{\alpha},\varphi^{-1}(\boldsymbol{\beta})\right)$
此即$\boldsymbol{\varphi}^*=\boldsymbol{\varphi}^{-1}$
反过来,若$\varphi^*=\varphi^{-1}$,则$(\varphi(\alpha),\varphi(\beta))=(\alpha,\varphi^*\varphi(\beta))=(\alpha,\beta)\varphi$保持内积,故$\varphi$是正交变换。对酉变换可类似证明.

Theorem 9.4.3 (正交算子与酉算子的等价代数刻画)

设φ是欧氏空间(酉空间)V上的正交变换(酉变换)$\Longleftrightarrow$在V的任一组(某一组)标准正交基下,φ的表示矩阵是正交阵(酉矩阵)

Proof 当φ是正交变换时,若φ在V的任一组标准正交基下的表示矩阵为$\boldsymbol{A}$,则$\boldsymbol{\varphi}^*$在同一组基下的表示矩阵为$\boldsymbol{A}'$
由$\boldsymbol{\varphi}^*=\boldsymbol{\varphi}^{-1}$得$\boldsymbol{A}'=\boldsymbol{A}^{-1}\Longrightarrow\boldsymbol{A}$是正交矩阵
同理,当φ是酉变换时,$\boldsymbol{\varphi}$在标准正交基下的表示矩阵$\boldsymbol{A}$应适合$\overline{\boldsymbol{A}}'=\boldsymbol{A}^{-1}$,即$\boldsymbol{A}$是酉矩阵

此外若线性变换φ在某一组标准正交基下的表示矩阵为正交(酉)矩阵,则φ是正交(酉)变换.这由线性变换与其表示矩阵的关系即得.

Corollary 9.2

$\boldsymbol{A}$为正交矩阵的充分必要条件是它的n个行向量是n维实行向量空间组成的欧氏空间(取标准内积)的标准正交基
或它的n个列向量是n维实列向量空间组成的欧氏空间(取标准内积)的标准正交基.

$\boldsymbol{A}$为酉矩阵的充分必要条件是它的n个行向量是n维复行向量空间组成的酉空间(取标准内积)的标准正交基
或它的n个列向量是n维复列向量空间组成的酉空间(取标准内积)的标准正交基.

Proposition 9.11 (正交矩阵与酉矩阵的性质)

若n阶实矩阵$\boldsymbol{A}$是正交矩阵(或者酉矩阵),则

(1)$\boldsymbol{A}$的行列式值等于1或-1

(2)$\boldsymbol{A}$的特征值的绝对值（模长）等于1.

Proof (1) 由$\boldsymbol{A}\boldsymbol{A}'=\boldsymbol{I}_n$,取行列式即得结论.

(2)设λ是$\boldsymbol{A}$的特征值,$\boldsymbol{x}\neq 0$是属于λ的特征向量,则$\boldsymbol{A}\boldsymbol{x}=\lambda\boldsymbol{x}$,于是$\overline{x}'\overline{\boldsymbol{A}}'=\bar{\lambda}\overline{\boldsymbol{x}}'$.所以$\overline{\boldsymbol{x}}'\overline{\boldsymbol{A}}'\boldsymbol{A}\boldsymbol{x}=\bar{\lambda}\overline{\boldsymbol{x}}'\lambda\boldsymbol{x}$即$\overline{\boldsymbol{x}}'\boldsymbol{x}=\bar{\lambda}\lambda\,(\overline{\boldsymbol{x}}'\boldsymbol{x})$

因此$\bar{\lambda}\lambda=1$,即$|\lambda|=1$.

Corollary 9.3

对角阵是正交矩阵的充分必要条件是主对角线上的元素为1或-1.

Theorem 9.4.4 (矩阵 QR 分解)

设$\boldsymbol{A}$是n阶实(复)矩阵,则$\boldsymbol{A}$可分解为$\boldsymbol{A}=\boldsymbol{Q}\boldsymbol{R}$其中$\boldsymbol{Q}$是正交(酉)矩阵,$\boldsymbol{R}$是一个上三角阵且主对角线上的元素均大于等于零并且若$\boldsymbol{A}$是非异阵,则这样的分解必唯一.

Proof 设$\boldsymbol{A}$是n阶实矩阵,$\boldsymbol{A}=(\boldsymbol{u}_1,\boldsymbol{u}_2,\cdots,\boldsymbol{u}_n)$是$\boldsymbol{A}$的列分块.

考虑n维实列向量空间$\mathbb{R}_n$,并取其标准内积,我们先通过类似于$Gram-Schmidt$方法的正交化过程

把$\{\boldsymbol{u}_1,\boldsymbol{u}_2,\cdots,\boldsymbol{u}_n\}$变成一组两两正交的向量$\{\boldsymbol{w}_1,\boldsymbol{w}_2,\cdots,\boldsymbol{w}_n\}$，并且$\boldsymbol{w}_k$或者是零向量或者是单位向量.

我们用数学归纳法来定义上述向量$\boldsymbol{w}_k(k=1,2,\cdots,n)$.

假设$\boldsymbol{w}_1,\cdots,\boldsymbol{w}_{k-1}$已经定义好,现来定义$\boldsymbol{w}_k$，令$\boldsymbol{v}_k=\boldsymbol{u}_k-\sum\limits_{j=1}^{k-1}(\boldsymbol{u}_k,\boldsymbol{w}_j)\boldsymbol{w}_j$

若$\boldsymbol{v}_k=0$,则令$\boldsymbol{w}_k=0$

若$\boldsymbol{v}_k\neq 0$,则令$\boldsymbol{w}_k=\dfrac{\boldsymbol{v}_k}{\|\boldsymbol{v}_k\|}$.

容易验证$\{\boldsymbol{w}_1,\boldsymbol{w}_2,\cdots,\boldsymbol{w}_n\}$是一组两两正交的向量,$w_k$或者是零向量或者是单位向量,并且满足：

$$\boldsymbol{u}_k=\sum_{j=1}^{k-1}(\boldsymbol{u}_k,\boldsymbol{w}_j)\boldsymbol{w}_j+\|\boldsymbol{v}_k\|\,\boldsymbol{w}_k,\ k=1,2,\cdots,n$$

$$\Longrightarrow \boldsymbol{A}=(\boldsymbol{u}_1\cdots\boldsymbol{u}_n)=(\boldsymbol{w}_1\cdots\boldsymbol{w}_n)\underbrace{\begin{pmatrix}\|\boldsymbol{v}_1\| & * & * & * \\ & \|\boldsymbol{v}_2\| & * & * \\ & & \ddots & * \\ & & & \|\boldsymbol{v}_n\|\end{pmatrix}}_{\text{即为题干中所求的}\boldsymbol{R}}$$

但此时$(\boldsymbol{w}_1\cdots\boldsymbol{w}_n)$未必为正交阵因为可能有些$\boldsymbol{w}_i=0$

但我们注意到若$\boldsymbol{w}_i=0$意味着$\boldsymbol{v}_k=0$则$\boldsymbol{R}$矩阵中第i行为0$\left(\boldsymbol{u}_k=\sum\limits_{j=1}^{k-1}(\boldsymbol{u}_k,\boldsymbol{w}_j)\boldsymbol{w}_j+\|\boldsymbol{v}_k\|\,\boldsymbol{w}_k,k=1,2,\cdots,n\text{可以看出}\right)$

此时假设$\boldsymbol{w}_{i_1},\boldsymbol{w}_{i_2},\cdots,\boldsymbol{w}_{i_r}$是其中的非零向量全体

可将它们扩张为$\mathbb{R}_n$的一组标准正交基$\{\widetilde{\boldsymbol{w}}_1,\widetilde{\boldsymbol{w}}_2,\cdots,\widetilde{\boldsymbol{w}}_n\}$,其中$\widetilde{\boldsymbol{w}}_j=\boldsymbol{w}_j,j=i_1,i_2,\cdots,i_r$

令$\boldsymbol{Q}=(\widetilde{\boldsymbol{w}}_1,\widetilde{\boldsymbol{w}}_2,\cdots,\widetilde{\boldsymbol{w}}_n)$,知$\boldsymbol{Q}$是正交矩阵

注意到若$\boldsymbol{w}_k=0$,则$\boldsymbol{R}$的第k行元素全为零,此时用$\widetilde{\boldsymbol{w}_k}$代替$\boldsymbol{w}_k$仍然可使

$$A = (u_1 \cdots u_n) = (w_1 \cdots w_n)\underbrace{\begin{pmatrix} \|v_1\| & * & * & * \\ & \|v_2\| & * & * \\ & & \ddots & * \\ & & & \|v_n\| \end{pmatrix}}_{\text{即为题干中所求的}\boldsymbol{R}}\text{成立}$$

因此$\boldsymbol{A} = (\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_n) = (\widetilde{\boldsymbol{w}}_1, \widetilde{\boldsymbol{w}}_2, \cdots, \widetilde{\boldsymbol{w}}_n)\,\boldsymbol{R} = \boldsymbol{QR}$从而得到了$\boldsymbol{A}$的$QR$分解.

再证明唯一性

若$A = Q_1R_1 = Q_2R_2$由A可逆知道Q_iR_i均可逆 $\Longrightarrow Q_2^{-1}Q_1 = R_2R_1^{-1}$

此时$Q_2^{-1}Q_1$为正交阵的乘积仍然为正交阵，此时$R_2R_1^{-1}$为主对角为正数的上三角阵乘积 $\Longrightarrow$ 仍然为主对角为正数的上三角阵

此时我们来证明一个性质：断言一个主对角线为正数的上三角阵$\boldsymbol{P}$也为正交阵那么$\boldsymbol{P}$必定为单位阵

因为$P = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix} \Longrightarrow P$的最后一行作为单位行向量就有$a_{nn}^2 = 1$又主对角全为正 $\Longrightarrow a_{nn} = 1$

但$\boldsymbol{P}$得最后一列作为单位列向量又$a_{nn} = 1$迫使$a_{1n} = a_{2n} = \cdots = a_{n-1,n} = 0$

以此类推考察第$n-1$行得到$P = I$

那么在唯一性证明中运用该性质得到$Q_2^{-1}Q_1 = R_2R_1^{-1} = I \Longrightarrow Q_1 = Q_2$且$R_2 = R_1$

Corollary 9.4

设$\boldsymbol{A}$是实数域上的$m \times n$矩阵, 其中$m > n$

证明：如果$\boldsymbol{A}$的列向量组$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n$线性无关, 那么$\boldsymbol{A}$可以唯一分解成$\boldsymbol{A} = \boldsymbol{QR}$

其中$\boldsymbol{Q}$是列向量组为正交单位向量组的$m \times n$矩阵, $\boldsymbol{R}$是主对角元都为正数的n级上三角矩阵, 这称为QR分解

Proof 先证可分解性

由于$\boldsymbol{A}$的列向量组$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n$线性无关, 因此经过施密特正交化过程, 可得到与它等价的正交向量组$\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_n$

$\boldsymbol{\alpha}_1 = \boldsymbol{\beta}_1$

....

$$\boldsymbol{\alpha}_n = \sum_{j=1}^{n-1} \frac{\left(\boldsymbol{\alpha}_n, \boldsymbol{\beta}_j\right)}{\left(\boldsymbol{\beta}_j, \boldsymbol{\beta}_j\right)} \boldsymbol{\beta}_j + \boldsymbol{\beta}_n$$

记$b_{ji} = \dfrac{\left(\boldsymbol{\alpha}_i, \boldsymbol{\beta}_j\right)}{\left(\boldsymbol{\beta}_j, \boldsymbol{\beta}_j\right)}, \quad i = 2, 3, \cdots, n; \quad j = 1, 2, \cdots, i-1$

再单位化可得到正交单位向量组$\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \cdots, \boldsymbol{\eta}_n$

此时

$$\begin{aligned}
\boldsymbol{A} &= (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n) \\
&= (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_n)\begin{pmatrix} 1 & b_{12} & \cdots & b_{1n} \\ 0 & 1 & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \\
&= (\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \cdots, \boldsymbol{\eta}_n)\begin{pmatrix} |\boldsymbol{\beta}_1| & b_{12}|\boldsymbol{\beta}_1| & \cdots & b_{1n}|\boldsymbol{\beta}_1| \\ 0 & |\boldsymbol{\beta}_2| & \cdots & b_{2n}|\boldsymbol{\beta}_2| \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & |\boldsymbol{\beta}_n| \end{pmatrix} \\
&= \boldsymbol{QR},
\end{aligned}$$

其中,$\boldsymbol{Q} = (\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \cdots, \boldsymbol{\eta}_n)$ 是列向量组为正交单位向量组的$m \times n$矩阵,$\boldsymbol{R}$是主对角元都为正数的n级上三角矩阵

再证唯一性

假如$\boldsymbol{A}$还有一种分解式,即$\boldsymbol{A} = \boldsymbol{Q}_1\boldsymbol{R}_1$符合所要求的条件,则$\boldsymbol{QR} = \boldsymbol{Q}_1\boldsymbol{R}_1$

由于$\boldsymbol{R}$是可逆的上三角矩阵,因此$\boldsymbol{Q} = \boldsymbol{Q}_1\boldsymbol{R}_1\boldsymbol{R}^{-1} = \boldsymbol{Q}_1\boldsymbol{C}$,其中$\boldsymbol{C} = \boldsymbol{R}_1\boldsymbol{R}^{-1}$是主对角元都为正数的上三角矩阵,设$\boldsymbol{C} = (\boldsymbol{C}_{ij})$

记$\boldsymbol{Q}_1$的列向量组为$\boldsymbol{\delta}_1, \boldsymbol{\delta}_2, \cdots, \boldsymbol{\delta}_n$,则

$$\begin{aligned}
(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \cdots, \boldsymbol{\eta}_n) &= (\boldsymbol{\delta}_1, \boldsymbol{\delta}_2, \cdots, \boldsymbol{\delta}_n)\begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ 0 & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & c_{nn} \end{pmatrix} \\
&= (c_{11}\boldsymbol{\delta}_1, c_{12}\boldsymbol{\delta}_1 + c_{22}\boldsymbol{\delta}_2, \cdots, c_{1n}\boldsymbol{\delta}_1 + c_{2n}\boldsymbol{\delta}_2 + \cdots + c_{nn}\boldsymbol{\delta}_n)
\end{aligned}$$

由于 $(\boldsymbol{\eta}_1, \boldsymbol{\eta}_1) = 1, (c_{11}\boldsymbol{\delta}_1, c_{11}\boldsymbol{\delta}_1) = c_{11}^2(\boldsymbol{\delta}_1, \boldsymbol{\delta}_1) = c_{11}^2$,因此$c_{11}^2 = 1$,由此得出$c_{11} = 1$

由于 $(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2) = 0, (\boldsymbol{\eta}_2, \boldsymbol{\eta}_2) = 1$,且

$(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2) = (c_{11}\boldsymbol{\delta}_1, c_{12}\boldsymbol{\delta}_1 + c_{22}\boldsymbol{\delta}_2) = c_{11}c_{12}(\boldsymbol{\delta}_1, \boldsymbol{\delta}_1) + c_{11}c_{22}(\boldsymbol{\delta}_1, \boldsymbol{\delta}_2) = c_{11}c_{12}$,

$(\boldsymbol{\eta}_2, \boldsymbol{\eta}_2) = (c_{12}\boldsymbol{\delta}_1 + c_{22}\boldsymbol{\delta}_2, c_{12}\boldsymbol{\delta}_1 + c_{22}\boldsymbol{\delta}_2) = c_{12}^2 + c_{22}^2$,

因此$c_{11}c_{12} = 0, c_{12}^2 + c_{22}^2 = 1$.由此得出,$c_{12} = 0, c_{22} = 1$

依次类推可得出,$c_{1k} = c_{2k} = \cdots = c_{k-1,k} = 0, c_{kk} = 1, k = 3, \cdots, n$,因此$\boldsymbol{C} = \boldsymbol{I}$,从而$\boldsymbol{Q} = \boldsymbol{Q}_1, \boldsymbol{R} = \boldsymbol{R}_1$

9.5 欧几里得空间上的正交矩阵

Definition 9.12 (正交矩阵定义)

实数域上的n级矩阵$\boldsymbol{A}$如果满足$\boldsymbol{A}'\boldsymbol{A}=\boldsymbol{I}$, 那么称$\boldsymbol{A}$是正交矩阵

Proposition 9.12

1.实数域上n级矩阵$\boldsymbol{A}$是正交矩阵

$\Longleftrightarrow \boldsymbol{A}'\boldsymbol{A}=\boldsymbol{I} \Longleftrightarrow \boldsymbol{A}$可逆, 且$\boldsymbol{A}^{-1}=\boldsymbol{A}' \Longleftrightarrow \boldsymbol{A}\boldsymbol{A}'=\boldsymbol{I}$

2.$\boldsymbol{I}$是正交矩阵

3.若$\boldsymbol{A}$和$\boldsymbol{B}$都是n级正交矩阵, 则$\boldsymbol{A}\boldsymbol{B}$也是正交矩阵

4.若$\boldsymbol{A}$是正交矩阵, 则$\boldsymbol{A}^{-1}$(即$\boldsymbol{A}'$)也是正交矩阵

5.若$\boldsymbol{A}$是正交矩阵, 则$|\boldsymbol{A}|=1$或-1

6.设实数域上n级矩阵$\boldsymbol{A}$的行向量组为$\boldsymbol{\gamma}_1,\boldsymbol{\gamma}_2,\cdots,\boldsymbol{\gamma}_n$; 列向量组为$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_i,\cdots,\boldsymbol{\alpha}_n$, 则

(1)$\boldsymbol{A}$为正交矩阵当且仅当$\boldsymbol{A}$的行向量组满足$\boldsymbol{\gamma}_i\boldsymbol{\gamma}_j'=\begin{cases}1, & 当i=j,\\ 0, & 当i\neq j;\end{cases}$

(2)$\boldsymbol{A}$为正交矩阵当且仅当$\boldsymbol{A}$的列向量组满足$\boldsymbol{\alpha}_i'\boldsymbol{\alpha}_j=\begin{cases}1, & 当i=j,\\ 0, & 当i\neq j\end{cases}$

Proof $\boldsymbol{A}$为正交矩阵 $\Longleftrightarrow \boldsymbol{A}\boldsymbol{A}'=\boldsymbol{I} \Longleftrightarrow \begin{pmatrix}\boldsymbol{\gamma}_1\\ \boldsymbol{\gamma}_2\\ \vdots\\ \boldsymbol{\gamma}_n\end{pmatrix}(\boldsymbol{\gamma}_1',\boldsymbol{\gamma}_2',\cdots,\boldsymbol{\gamma}_n')=\begin{pmatrix}1&0&\cdots&0\\ 0&1&\cdots&0\\ \vdots&\vdots&&\vdots\\ 0&0&\cdots&1\end{pmatrix} \Longleftrightarrow \gamma_i\boldsymbol{\gamma}_j'=\begin{cases}1, & 当i=j,\\ 0, & 当i\neq j。\end{cases}$

类似于 (1)的方法, 利用$\boldsymbol{A}$为正交矩阵 $\Longleftrightarrow \boldsymbol{A}'\boldsymbol{A}=\boldsymbol{I}$可证得结论

引用$Kronecker$记号δ_{ij}, 它的含义是$\delta_{ij}=\begin{cases}1, & 当i=j,\\ 0, & 当i\neq j,\end{cases}$

Definition 9.13 (酉矩阵定义)

设A是复数域上的矩阵, 用A^*表示$\bar{A}'$, 即把A的每个元素取共轭复数得到的矩阵$\bar{A}$再转置

(注意从上下文区别A^*是表示$\bar{A}'$还是表示A的伴随矩阵)

如果n级复矩阵A满足$A^*A=I$, 那么称A是酉矩阵

Proposition 9.13

证明：下列每一个条件都是n级复矩阵$A=(a_{ij})$为酉矩阵的充分必要条件：

(1)A可逆, 且$A^{-1}=A^*$

(2)$AA^*=I$

(3) $\sum\limits_{k=1}^{n}a_{ik}\bar{a}_{jk}=\delta_{ij}, 1\leqslant i,j\leqslant n$

(4) $\sum\limits_{k=1}^{n}\bar{a}_{ki}\cdot a_{kj}=\delta_{ij}, 1\leqslant i,j\leqslant n$

Proof (1)和 (2)n级复矩阵A是酉矩阵 $\Leftrightarrow A^*A=I \Leftrightarrow A$可逆, 且$A^{-1}=A^* \Leftrightarrow AA^*=I$。

(3)n级复矩阵$A=(a_{ij})$是酉矩阵 $\Leftrightarrow AA^*=I \Leftrightarrow AA^*(i;j)=I(i;j), 1\leqslant i,j\leqslant n \Leftrightarrow \sum\limits_{k=1}^{n}a_{ik}\bar{a}_{jk}=\delta_{ij}, 1\leqslant i,j\leqslant n$

(4)n级复矩阵$A=(a_{ij})$ 是西矩阵 $\Leftrightarrow A^*A=I \Leftrightarrow A^*A(i;j)=I(i;j), 1\leqslant i,j\leqslant n \Leftrightarrow \sum_{k=1}^{n}\bar{a}_{ki}a_{kj}=\delta_{ij}, 1\leqslant i,j\leqslant n$

Lemma 9.1

(1) $\left(\overline{AB}\right)^T=\left(\overline{B}\right)^T\left(\overline{A}\right)^T$

(2) $\left(\overline{A^{-1}}\right)^T=\left(\left(\overline{A}\right)^T\right)^{-1}$

Proof (1) 在等式两边同时取 (i,j) 元观察即可

(2) 在等式两边同时取 (i,j) 元观察即可

Proposition 9.14

1.证明：两个n级西矩阵的乘积是西矩阵;西矩阵的逆矩阵是酉矩阵

2.证明：西矩阵的行列式的模等于1

Proof 1.

利用上述引理很容易得到

2.设A是n级西矩阵,则$AA^*=I$。从而$1=|I|=|AA^*|=|A|\,|A^*|$.

n级复矩阵$A=(a_{ij})$ 的行列式为$|A|=\sum\limits_{j_1j_2\cdots j_n}(-1)^{r(j_1j_2\cdots j_n'}a_{1j_1}a_{2j_2}\cdots a_{nj_n}$.

n级复矩阵$\bar{A}=(\bar{a}_{ij})$ 的行列式为

$$|\bar{A}|=\sum_{j_1j_2\cdots j_n}(-1)^{\tau(j_1j_2\cdots j_n\bar{a}_{1j_1}\bar{a}_{2j_2}\cdots\bar{a}_{nj_n}}=\overline{\left(\sum_{j_1j_2\cdots j_n}(-1)^{(j_1j_2\cdots j_n)}a_{1j_1}a_{2j_2}\cdots a_{nj_n}\right)}=\overline{A\,|}$$

因此$1=|A|\,|A^*|=|A|\left|\bar{A}'\right|=|A||\bar{A}|=|A|\overline{|A|}$从而$|A|$的模等于1。

Proposition 9.15

决定所有的2级正交矩阵

Proof 设$\boldsymbol{A}=(a_{ij})$是2级正交矩阵,则$\boldsymbol{A}^{-1}=\boldsymbol{A}'$,即$\frac{1}{|\boldsymbol{A}|}\begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}=\begin{pmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{pmatrix}$

由于$|\boldsymbol{A}|=1$或-1.因此分两种情形:

情形1:$|\boldsymbol{A}|=1$,此时有$a_{22}=a_{11},\quad a_{21}=-a_{12}$.由于$a_{11}^2+a_{21}^2=1$

可以设$a_{11}=\cos\theta,\quad a_{21}=\sin\theta,\quad 0\leqslant\theta<2\pi$,于是$\boldsymbol{A}=\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix},\quad 0\leqslant\theta<2\pi$,容易直接验证,所求出的$\boldsymbol{A}$是正交矩阵

情形2:$|\boldsymbol{A}|=-1$,此时有$a_{22}=-a_{11},\quad a_{21}=a_{12}$

由于$a_{11}^2+a_{21}^2=1$,因此同情形1的理由得$a_{11}=\cos\theta,\quad a_{21}=\sin\theta$,于是$\boldsymbol{A}=\begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix},\quad 0\leqslant\theta<2\pi$

容易直接验证,所求出的$\boldsymbol{A}$是正交矩阵

综上所述,2级正交矩阵有且只有下列两种类型:$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix},\begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix},\quad 0\leqslant\theta<2\pi$

Proposition 9.16

设$\boldsymbol{A}$是n级正交矩阵

证明:任意取定$\boldsymbol{A}$的两行(或两列),由这两行(或两列)的元素组成的所有2阶子式的平方和等于1

(上述结论可以推广到任意k行k列)

Proof 取定$\boldsymbol{A}$的第i_1,i_2行$(i_1<i_2)$。由于$\boldsymbol{A}\boldsymbol{A}'=\boldsymbol{I}$,因此得

$$\sum_{1\leqslant v_1<v_2\leqslant n}\left[\boldsymbol{A}\begin{pmatrix} i_1, & i_2 \\ v_1, & v_2 \end{pmatrix}\right]^2=\sum_{1\leqslant v_1<v_2\leqslant n}\boldsymbol{A}\begin{pmatrix} i_1, & i_2 \\ v_1, & v_2 \end{pmatrix}\boldsymbol{A}'\begin{pmatrix} v_1, & v_2 \\ i_1, & i_2 \end{pmatrix}=\boldsymbol{A}\boldsymbol{A}'\begin{pmatrix} i_1, & i_2 \\ i_1, & i_2 \end{pmatrix}=\boldsymbol{I}\begin{pmatrix} i_1, & i_2 \\ i_1, & i_2 \end{pmatrix}=1。$$

Proposition 9.17

证明:实数域上的一个n级矩阵如果具有下列三个性质中的任意两个性质,那么必有第三个性质:

正交矩阵,对称矩阵,对合矩阵

Proof 设n级实矩阵$\boldsymbol{A}$是正交矩阵,且是对称矩阵,则$\boldsymbol{A}^2=\boldsymbol{A}\boldsymbol{A}=\boldsymbol{A}\boldsymbol{A}'=\boldsymbol{I}$,因此$\boldsymbol{A}$是对合矩阵

设$\boldsymbol{A}$是正交矩阵和对合矩阵,则$\boldsymbol{A}'=\boldsymbol{A}^{-1}=\boldsymbol{A}$,因此$\boldsymbol{A}$是对称矩阵

设实矩阵$\boldsymbol{A}$是对称矩阵和对合矩阵,则$\boldsymbol{A}\boldsymbol{A}'=\boldsymbol{A}\boldsymbol{A}=\boldsymbol{A}^2=\boldsymbol{I}$,因此$\boldsymbol{A}$是正交矩阵。

Proposition 9.18

设$\boldsymbol{A}$是n级正交矩阵,证明:对于欧几里得空间$\mathbb{R}^n$中任一列向量$\boldsymbol{\alpha}$,有$|\boldsymbol{A}\boldsymbol{\alpha}|=|\boldsymbol{\alpha}|$

Proof $|\boldsymbol{A}\boldsymbol{\alpha}|^2=(\boldsymbol{A}\boldsymbol{\alpha},\boldsymbol{A}\boldsymbol{\alpha})=(\boldsymbol{A}\boldsymbol{\alpha})'(\boldsymbol{A}\boldsymbol{\alpha})=\boldsymbol{\alpha}'\boldsymbol{A}'\boldsymbol{A}\boldsymbol{\alpha}=\boldsymbol{\alpha}'\boldsymbol{\alpha}=(\boldsymbol{\alpha},\boldsymbol{\alpha})=|\boldsymbol{\alpha}|^2$。因此$|\boldsymbol{A}\boldsymbol{\alpha}|=|\boldsymbol{\alpha}|$。

9.6 QR 分解与方程组的最小二乘解与正交投影

笔记 在实际问题中常常遇到方程个数m大于未知量个数n的线性方程组$\boldsymbol{Ax}=\boldsymbol{\beta}$, 它可能无解
这时要设法找一个n维列向量$\boldsymbol{x}_0$, 使得 $|\boldsymbol{Ax}_0-\boldsymbol{\beta}|^2$ 最小, $\boldsymbol{x}_0$称为线性方程组$\boldsymbol{Ax}=\boldsymbol{\beta}$的最小二乘解
接下来我们将逐步证明：
$\boldsymbol{x}_0$是$\boldsymbol{Ax}=\boldsymbol{\beta}$的最小二乘解 $\Longleftrightarrow$ $\boldsymbol{x}_0$是线性方程组$\boldsymbol{A}'\boldsymbol{Ax}=\boldsymbol{A}'\boldsymbol{\beta}$的解
当$\boldsymbol{A}$是$m\times n$列满秩矩阵时, 线性方程组$\boldsymbol{Ax}=\boldsymbol{\beta}$的最小二乘解只有一个, 它是$\boldsymbol{R}^{-1}Q'\boldsymbol{\beta}$, 其中$Q\boldsymbol{R}=\boldsymbol{A}$
且Q是列向量组为正交单位向量组的$m\times n$矩阵, $\boldsymbol{R}$是主对角元都为正数的上三角矩阵

Definition 9.14 (正交补空间定义)

设U是欧几里得空间R^n的一个子空间, 如果向量α与U中每一个向量正交, 那么称α与U正交, 记作$\alpha\perp U$
令$U^\perp \overset{\text{def}}{=\!=} \{\alpha\in\mathrm{R}^n \mid \alpha\perp U\}$, 称$U^\perp$是$U$的正交补
证明：$U^\perp$是R^n的一个子空间

Proof 由于$0\perp U$, 因此$U^\perp$是R^n的一个非空子集
任取$\boldsymbol{\alpha},\boldsymbol{\beta}\in U^\perp$, 则对一切$\boldsymbol{\gamma}\in U$, 有
$(\boldsymbol{\alpha}+\boldsymbol{\beta},\boldsymbol{\gamma})=(\boldsymbol{\alpha},\boldsymbol{\gamma})+(\boldsymbol{\beta},\boldsymbol{\gamma})=0+0=0,$
$(k\boldsymbol{\alpha},\boldsymbol{\gamma})=k(\boldsymbol{\alpha},\boldsymbol{\gamma})=0, k\in K,$
因此$\boldsymbol{\alpha}+\boldsymbol{\beta}\in U^\perp, k\boldsymbol{\alpha}\in U^\perp$, 从而$U^\perp$是$\mathrm{R}^n$的一个空间

Definition 9.15 (正交投影)

设U是欧几里得空间R^n的一个子空间
令 $\begin{aligned}\boldsymbol{P}_U:\mathrm{R}^n&\longrightarrow\mathrm{R}^n\\ \boldsymbol{\alpha}&\longrightarrow\boldsymbol{\alpha}_1,\end{aligned}$ 其中$\boldsymbol{\alpha}_1\in U$, 并且$\boldsymbol{\alpha}-\boldsymbol{\alpha}_1\in U^\perp$
则称$\boldsymbol{P}_U$是R^n在U上的正交投影, 把$\boldsymbol{\alpha}_1$称为向量$\boldsymbol{\alpha}$在U上的正交投影

Proposition 9.19

证明：对于$\boldsymbol{\alpha}\in\mathrm{R}^n, \boldsymbol{\alpha}_1\in U$是$\boldsymbol{\alpha}$在$U$上的正交投影当且仅当 $|\boldsymbol{\alpha}-\boldsymbol{\alpha}_1|\leqslant|\boldsymbol{\alpha}-\boldsymbol{\gamma}|, \forall\boldsymbol{\gamma}\in U$

Proof 必要性
设$\boldsymbol{\alpha}_1\in U$是$\boldsymbol{\alpha}$在$U$上的正交投影, 则$\boldsymbol{\alpha}-\boldsymbol{\alpha}_1\in U^\perp$, 从而$\forall\boldsymbol{\gamma}\in U$, 有 $(\boldsymbol{\alpha}-\boldsymbol{\alpha}_1)\perp(\boldsymbol{\alpha}_1-\boldsymbol{\gamma})$, 于是

$$\begin{aligned}|\boldsymbol{\alpha}-\boldsymbol{\gamma}|^2&=|\boldsymbol{\alpha}-\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_1-\boldsymbol{\gamma}|^2=(\boldsymbol{\alpha}-\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_1-\boldsymbol{\gamma},\boldsymbol{\alpha}-\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_1-\boldsymbol{\gamma})\\&=(\boldsymbol{\alpha}-\boldsymbol{\alpha}_1,\boldsymbol{\alpha}-\boldsymbol{\alpha}_1)-2(\boldsymbol{\alpha}-\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_1-\boldsymbol{\gamma})+(\boldsymbol{\alpha}_1-\boldsymbol{\gamma},\boldsymbol{\alpha}_1-\boldsymbol{\gamma})\\&=|\boldsymbol{\alpha}-\boldsymbol{\alpha}_1|^2+|\boldsymbol{\alpha}_1-\boldsymbol{\gamma}|^2\geqslant|\boldsymbol{\alpha}-\boldsymbol{\alpha}_1|^2,\end{aligned}$$

从而$|\boldsymbol{\alpha}-\boldsymbol{\gamma}|\geqslant|\boldsymbol{\alpha}-\boldsymbol{\alpha}_1|$

充分性
设 $|\boldsymbol{\alpha}-\boldsymbol{\alpha}_1|\leqslant|\boldsymbol{\alpha}-\boldsymbol{\gamma}|, \forall\boldsymbol{\gamma}\in U$
假设$\boldsymbol{\delta}$是$\boldsymbol{\alpha}$在U上的正交投影
一方面 $|\boldsymbol{\alpha}-\boldsymbol{\alpha}_1|\leqslant|\boldsymbol{\alpha}-\boldsymbol{\gamma}|, \forall\boldsymbol{\gamma}\in U$取$\gamma=\delta \Longrightarrow |\boldsymbol{\alpha}-\boldsymbol{\delta}|\geqslant|\boldsymbol{\alpha}-\boldsymbol{\alpha}_1|$
另一方面根据证得的必要性得, $|\boldsymbol{\alpha}-\boldsymbol{\delta}|\leqslant|\boldsymbol{\alpha}-\boldsymbol{\alpha}_1|$
从而$|\boldsymbol{\alpha}-\boldsymbol{\delta}|=|\boldsymbol{\alpha}-\boldsymbol{\alpha}_1|$
由于$\boldsymbol{\alpha}-\boldsymbol{\delta}\in U^\perp, \boldsymbol{\delta}-\boldsymbol{\alpha}_1\in U$, 因此 $(\boldsymbol{\alpha}-\boldsymbol{\delta})\perp(\boldsymbol{\delta}-\boldsymbol{\alpha}_1)$

同上理, 得 $|\alpha-\alpha_1|^2=|\alpha-\delta+\delta-\alpha_1|^2=|\alpha-\delta|^2+|\delta-\alpha_1|^2=|\alpha-\alpha_1|^2+|\delta-\alpha_1|^2$ (勾股定理)
由此得出, $|\delta-\alpha_1|^2=0$, 因此 $\delta=\alpha_1$, 即 α_1 是 α 在 U 上的正交投影

Lemma 9.2

设 $\boldsymbol{A}$ 是实数域上的一个 $m\times n$ 矩阵, $m>n,\boldsymbol{\beta}\in\mathrm{R}^m$
如果 $\boldsymbol{x}_0\in\mathrm{R}^n$ 使得 $|\boldsymbol{\beta}-\boldsymbol{A}\boldsymbol{x}_0|^2\leqslant|\boldsymbol{\beta}-\boldsymbol{A}\boldsymbol{x}|^2,\forall\boldsymbol{x}\in\mathrm{R}^n$, 那么称 $\boldsymbol{x}_0$ 是线性方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{\beta}$ 的最小二乘解
证明 : $\boldsymbol{x}_0$ 是 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{\beta}$ 的最小二乘解 $\Longleftrightarrow$ $\boldsymbol{x}_0$ 是线性方程组 $\boldsymbol{A}'\boldsymbol{A}\boldsymbol{x}=\boldsymbol{A}'\boldsymbol{\beta}$ 的解

Proof 用 U 表示矩阵 $\boldsymbol{A}$ 的列空间, $U=\langle\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n\rangle$, 则
$\boldsymbol{x}_0$ 是 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{\beta}$ 的最小二乘解
$\Longleftrightarrow|\boldsymbol{\beta}-\boldsymbol{A}\boldsymbol{x}_0|^2\leqslant|\boldsymbol{\beta}-\boldsymbol{A}\boldsymbol{x}|^2,\forall\boldsymbol{x}\in\mathrm{R}^n$
$\Longleftrightarrow|\boldsymbol{\beta}-\boldsymbol{A}\boldsymbol{x}_0|\leqslant|\boldsymbol{\beta}-\boldsymbol{\gamma}|,\forall\boldsymbol{\gamma}\in U$
$\Longleftrightarrow\boldsymbol{A}\boldsymbol{x}_0$ 是 $\boldsymbol{\beta}$ 在 U 上的正交投影
$\Longleftrightarrow\boldsymbol{\beta}-\boldsymbol{A}\boldsymbol{x}_0\in U^{\perp}\Longleftrightarrow(\boldsymbol{\beta}-\boldsymbol{A}\boldsymbol{x}_0,\boldsymbol{\alpha}_i)=0,i=1,2,\cdots,n$
$\Longleftrightarrow\boldsymbol{\alpha}_i'(\boldsymbol{\beta}-\boldsymbol{A}\boldsymbol{x}_0)=0,i=1,2,\cdots,n$
$\Longleftrightarrow\boldsymbol{A}'(\boldsymbol{\beta}-\boldsymbol{A}\boldsymbol{x}_0)=0$
$\Longleftrightarrow\boldsymbol{A}'\boldsymbol{A}\boldsymbol{x}_0=\boldsymbol{A}'\boldsymbol{\beta}$
$\Longleftrightarrow\boldsymbol{x}_0$ 是 $\boldsymbol{A}'\boldsymbol{A}\boldsymbol{x}=\boldsymbol{A}'\boldsymbol{\beta}$ 的解

Lemma 9.3

设 $\boldsymbol{A}$ 是实数域上的 $m\times n$ 列满秩矩阵
它可分解成 $\boldsymbol{A}=Q\boldsymbol{R}$, 其中 $\boldsymbol{Q}$ 是列向量组为正交单位向量组的 $m\times n$ 矩阵, $\boldsymbol{R}$ 为主对角元都为正数的上三角矩阵
证明 : 对于任意 $\boldsymbol{\beta}\in\mathrm{R}^m$, $\boldsymbol{R}^{-1}Q'\boldsymbol{\beta}$ 是线性方程组 $\boldsymbol{A}'\boldsymbol{A}\boldsymbol{x}=\boldsymbol{A}'\boldsymbol{\beta}$ 的唯一解

Proof 设 Q 的列向量为 $\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\cdots,\boldsymbol{\eta}_n$, 则 $Q'Q=\begin{pmatrix}\boldsymbol{\eta}_1'\\ \boldsymbol{\eta}_2'\\ \vdots\\ \boldsymbol{\eta}_n'\end{pmatrix}(\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\cdots,\boldsymbol{\eta}_n)=\begin{pmatrix}\boldsymbol{\eta}_1'\boldsymbol{\eta}_1 & \boldsymbol{\eta}_1'\boldsymbol{\eta}_2 & \cdots & \boldsymbol{\eta}_1'\boldsymbol{\eta}_n\\ \boldsymbol{\eta}_2'\boldsymbol{\eta}_1 & \boldsymbol{\eta}_2'\boldsymbol{\eta}_2 & \cdots & \boldsymbol{\eta}_2'\boldsymbol{\eta}_n\\ \vdots & \vdots & & \vdots\\ \boldsymbol{\eta}_n'\boldsymbol{\eta}_1 & \boldsymbol{\eta}_n'\boldsymbol{\eta}_2 & \cdots & \boldsymbol{\eta}_n'\boldsymbol{\eta}_n\end{pmatrix}=\boldsymbol{I}$
从而 $\boldsymbol{A}'\boldsymbol{A}(\boldsymbol{R}^{-1}\boldsymbol{Q}'\boldsymbol{\beta})=(\boldsymbol{Q}\boldsymbol{R})'(\boldsymbol{Q}\boldsymbol{R})(\boldsymbol{R}^{-1}\boldsymbol{Q}'\boldsymbol{\beta})=\boldsymbol{R}'\boldsymbol{Q}'\boldsymbol{Q}\boldsymbol{R}\boldsymbol{R}^{-1}\boldsymbol{Q}'\boldsymbol{\beta}=\boldsymbol{R}'\boldsymbol{Q}'\boldsymbol{\beta}=(\boldsymbol{Q}\boldsymbol{R})'\boldsymbol{\beta}=\boldsymbol{A}'\boldsymbol{\beta}$,
因此 $\boldsymbol{R}^{-1}\boldsymbol{Q}'\boldsymbol{\beta}$ 是线性方程组 $\boldsymbol{A}'\boldsymbol{A}\boldsymbol{x}=\boldsymbol{A}'\boldsymbol{\beta}$ 的一个解
由于 $\operatorname{rank}(\boldsymbol{A}'\boldsymbol{A})=\operatorname{rank}(\boldsymbol{A})=n$, 因此 $|\boldsymbol{A}'\boldsymbol{A}|\neq0$, 从而线性方程组 $\boldsymbol{A}'\boldsymbol{A}\boldsymbol{x}=\boldsymbol{A}'\boldsymbol{\beta}$ 有唯一解, 它就是 $\boldsymbol{R}^{-1}Q'\boldsymbol{\beta}$

Proposition 9.20

设 $\boldsymbol{A}$ 是实数域上的 $m\times n$ 列满秩矩阵, $m>n$
证明 : $(\boldsymbol{A}'\boldsymbol{A})^{-1}\boldsymbol{A}'\boldsymbol{\beta}$ 是线性方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{\beta}$ 的唯一的最小二乘解

Proof 一方面 $r(\boldsymbol{A}'\boldsymbol{A})_{n\times n}=r(A)=n$
由上面引理知道此时 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{\beta}$ 的存在唯一的最小二乘解.
而此时 $\boldsymbol{A}'\boldsymbol{A}\boldsymbol{y}=\boldsymbol{A}'\boldsymbol{\beta}$ 的确有解 $(\boldsymbol{A}'\boldsymbol{A})^{-1}\boldsymbol{A}'\boldsymbol{\beta}$ 故 $(\boldsymbol{A}'\boldsymbol{A})^{-1}\boldsymbol{A}'\boldsymbol{\beta}$ 是一个最小二乘解
又有唯一性保证即可 (唯一性保证也可以不用引理而改用线性方程组的唯一解定理)