

线性代数笔记

Linear Algebra Done Wrong?

林晓烁

<https://xiaoshuo-lin.github.io>

2023 年春季-秋季

导言

1. 所有带括号的编号均表示来自 2023 年春季版本讲义中对应部分. 习题部分带有“旧”标记的是 2021 年 10 月第 1 版印刷的《线性代数讲义》相应题号的习题, 它们在后续版本中已经被修改或删除, 仅供玩赏.
2. 王新茂教授线性代数课程主页 (需在校园网下访问): <http://210.45.79.25/xxds>.



目录

第一章 拾遗补阙篇	1
1.1 分块矩阵	1
1.2 矩阵在图论中的应用	2
1.3 行列式的计算	2
1.3.1 反对称阵的行列式	2
1.3.2 Laplace 展开	5
1.4 相抵标准形	9
1.4.1 基本理论	9
1.4.2 线性方程组	9
1.4.3 广义逆矩阵	11
1.4.4 用相抵研究特殊矩阵	13
1.5 秩不等式与等式	16
1.6 Smith 标准形	19
1.7 矩阵的相似	22
1.8 根子空间	26
第二章 讲义习题选做	28
第三章 考前自测清单	154
3.1 矩阵的相似	154
3.2 正交方阵与酉方阵	156
3.3 内积空间	157

第一章 拾遗补阙篇

1.1 分块矩阵

下面引进一种方法, 把矩阵转变成一个与 Kronecker 乘积运算相容的向量.

定义 1.1.1 设 $X = (\mathbf{x}_1 \ \mathbf{x}_2 \ \cdots \ \mathbf{x}_n) \in \mathbb{F}^{m \times n}$. 算子 $\text{vec} : \mathbb{F}^{m \times n} \rightarrow \mathbb{F}^{mn}$ 由

$$\text{vec } X = \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_n \end{pmatrix} \in \mathbb{F}^{mn}$$

定义, 也就是说, vec 把 X 的列竖直地叠放在一起.

定理 1.1.2 设 $A \in \mathbb{F}^{m \times n}$, $X \in \mathbb{F}^{n \times p}$, $B \in \mathbb{F}^{p \times q}$. 那么 $\text{vec}(AXB) = (B^\top \otimes A) \text{vec } X$.

证明 设 $B = (b_{ij}) = (\mathbf{b}_1 \ \mathbf{b}_2 \ \cdots \ \mathbf{b}_q)$, $X = (\mathbf{x}_1 \ \mathbf{x}_2 \ \cdots \ \mathbf{x}_p)$. 则 AXB 的第 k 列是

$$AX\mathbf{b}_k = A \sum_{i=1}^p b_{ik} \mathbf{x}_i = \sum_{i=1}^p (b_{ik} A) \mathbf{x}_i = \begin{pmatrix} b_{1k} A & b_{2k} A & \cdots & b_{pk} A \end{pmatrix} \text{vec } X = (\mathbf{b}_k^\top \otimes A) \text{vec } X.$$

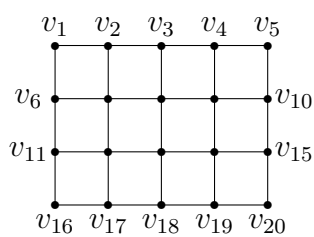
把这些向量竖直地叠放在一起, 就得到

$$\text{vec}(AXB) = \begin{pmatrix} \mathbf{b}_1^\top \otimes A \\ \mathbf{b}_2^\top \otimes A \\ \vdots \\ \mathbf{b}_q^\top \otimes A \end{pmatrix} \text{vec } X = (B^\top \otimes A) \text{vec } X.$$

□

1.2 矩阵在图论中的应用

例 1.2.1 考虑如下的网格图 G 的邻接矩阵 $A = (a_{ij})_{20 \times 20}$. 把 A 的 400 个元素都写出来显然很费事, 逐个指明 A 中所有 1 的位置也很麻烦, 而把 A 表示为分块矩阵的形式则相对容易一些.



$$A = \begin{pmatrix} P_5 & I_5 & O & O \\ I_5 & P_5 & I_5 & O \\ O & I_5 & P_5 & I_5 \\ O & O & I_5 & P_5 \end{pmatrix}, \quad P_5 = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

运用 Kronecker 乘积可写成 $A = I_4 \otimes P_5 + P_4 \otimes I_5$ (P_k 的具体定义见习题 2.2.4 证明).

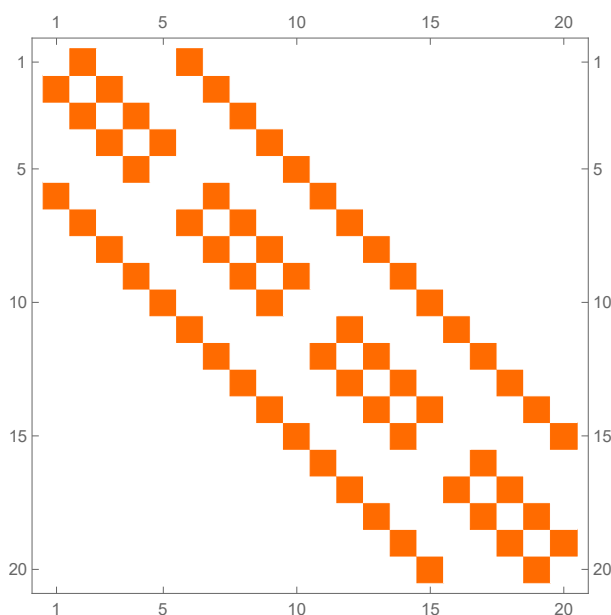


图 1.1: 网格图 $\text{GridGraph}[\{5,4\}]$ 对应的矩阵图

1.3 行列式的计算

1.3.1 反对称阵的行列式

命题 1.3.1 设 A 是奇数阶反对称阵, 则 $|A| = 0$.

证明 由于 A 的主对角元全为 0, 故只需考虑单项

$$T = \{a_{k_1,1}a_{k_2,2} \cdots a_{k_n,n} \mid \{k_1, \cdots, k_n\} \text{ 是 } \{1, \cdots, n\} \text{ 的一个排列, 且 } k_i \neq i \ (i = 1, \cdots, n)\}.$$

定义映射

$$\varphi: T \rightarrow T, \quad a_{k_1,1} a_{k_2,2} \cdots a_{k_n,n} \mapsto a_{1,k_1} a_{2,k_2} \cdots a_{n,k_n}.$$

由 A 是反对称阵可得 $\varphi^2 = \text{Id}_T$, 因此 φ 是双射. 断言 $a_{k_1,1} a_{k_2,2} \cdots a_{k_n,n}$ 和 $a_{1,k_1} a_{2,k_2} \cdots a_{n,k_n}$ 作为 $|A|$ 的单项¹不相同, 否则对第 1 个单项中任一元素 $a_{k_i,i}$, 必能在第 2 个单项中找到与之相等的元素 a_{j,k_j} , 也即 $k_i = j, k_j = i$. 因为 $k_i \neq i$, 所以 $i \neq j$, 由此可将 $\{1, \cdots, n\}$ 划分成两两一组的若干组, 但这与 n 为奇数矛盾. 将上述两个单项看成一组, 则易知它们在 $|A|$ 中符号相同. 因为 A 反对称, 所以 $a_{1,k_1} a_{2,k_2} \cdots a_{n,k_n} = (-1)^n a_{k_1,1} a_{k_2,2} \cdots a_{k_n,n} = -a_{k_1,1} a_{k_2,2} \cdots a_{k_n,n}$. 从而每组和为 0, 于是 $|A| = 0$. \square

注 1.3.2 注意不能由

$$\det(A) = \det(A^T) = \det(-A) = (-1)^n \det(A)$$

及 n 为奇数得到 $\det(A) = 0$. 这是因为 A 的元素所在数域 \mathbb{F} 未知, 可能有 $\text{char } \mathbb{F} = 2$, 此时对任意 n 均有 $(-1)^n = 1$.

引理 1.3.3 设 $A = (a_{ij}) \in \mathbb{F}^{2n \times 2n}$ 为反对称阵, 则 $|A| = f^2$, 其中 f 是关于 a_{ij} ($i < j$) 的某个多项式.

证明 对阶数 $2n$ 进行归纳.

① 当 $n = 1$ 时, $A = \begin{pmatrix} 0 & a_{12} \\ -a_{12} & 0 \end{pmatrix}$, $|A| = a_{12}^2$, 结论成立.

② 设结论对 $2(n-1)$ 阶反对称阵成立, 现证 $2n$ 阶反对称阵的情形. 若 $|A| = 0$, 结论显然成立. 下面不妨设 $|A| \neq 0$, 即 A 是 $2n$ 阶非异反对称阵.

设 C_{ij} 是去掉 A 的第 i 行和第 j 列得到的 $2n-1$ 阶子矩阵, $C_{ij,ij}$ 是去掉 A 的第 i, j 行和第 i, j 列得到的 $2(n-1)$ 阶子矩阵. 显然, 对任意的 i, j , C_{ij} 和 $C_{ij,ij}$ 都是反对称阵. 设 A^* 是 A 的伴随方阵, 则 $AA^* = A^*A = |A|I_{2n}$. 取行列式可得 $|A||A^*| = |A|^{2n}$, 从而 $|A^*| = |A|^{2n-1}$. 考虑如下矩阵乘积:

$$\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3,2n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{2n,1} & a_{2n,2} & a_{2n,3} & \cdots & a_{2n,2n} \end{pmatrix} \begin{pmatrix} A_{11} & A_{21} & A_{31} & \cdots & A_{2n,1} \\ A_{12} & A_{22} & A_{32} & \cdots & A_{2n,2} \\ A_{13} & A_{23} & A_{33} & \cdots & A_{2n,3} \\ \vdots & \vdots & \vdots & & \vdots \\ A_{1,2n} & A_{2,2n} & A_{3,2n} & \cdots & A_{2n,2n} \end{pmatrix}$$

¹这里把一般的方阵 $A = (a_{ij}) \in \mathbb{F}^{n \times n}$ 的元素 a_{ij} 看成是 n^2 个未定元, 并把 $|A|$ 看成是关于未定元 a_{ij} 的多元多项式.

$$= \begin{pmatrix} A_{12} & A_{22} & A_{32} & \cdots & A_{2n,2} \\ A_{11} & A_{21} & A_{31} & \cdots & A_{2n,1} \\ 0 & 0 & |A| & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & |A| \end{pmatrix},$$

取行列式即得

$$-|C_{12,12}| \cdot |A|^{2n-1} = (A_{12}A_{21} - A_{11}A_{22}) \cdot |A|^{2n-2}.$$

由 C_{11}, C_{22} 是奇数阶反对称阵, 由命题 1.3.1 可知它们的行列式均为 0, 于是 $A_{11} = A_{22} = 0$. 又 $C_{12} = -(C^T)_{21}$, 故 $A_{12} = -|C_{12}| = (-1)^{-1+2n-1} |(C^T)_{21}| = -A_{21}$. 从而 $A_{12}^2 = |C_{12,12}| \cdot |A|$, 即 $A_{12} = \sqrt{|C_{12,12}| \cdot |A|}$ (此处平方根为多值函数, 根据实际情况取正负号). 对任意的 $i < j$, 同理可证 $A_{ij} = \sqrt{|C_{ij,ij}| \cdot |A|}$. 按 A 的第一行进行展开可得

$$|A| = \sum_{j=2}^{2n} a_{1j} A_{1j} = \sum_{j=2}^{2n} a_{1j} \sqrt{|C_{1j,1j}| \cdot |A|},$$

于是

$$\sqrt{|A|} = \sum_{j=2}^{2n} a_{1j} \sqrt{|C_{1j,1j}|}.$$

因为 $C_{1j,1j}$ 是 $2(n-1)$ 阶反对称阵, 由归纳假设存在关于 a_{ij} ($i < j$) 的多项式 f_{1j} , 使得 $|C_{1j,1j}| = f_{1j}^2$, 故由上式可得

$$|A| = \left(\sum_{j=2}^{2n} a_{1j} f_{1j} \right)^2.$$

□

定义 1.3.4 设

$$A_{2n} = \{(i_1, j_1, i_2, j_2, \dots, i_n, j_n) \in S_{2n} \mid i_1 < j_1, i_2 < j_2, \dots, i_n < j_n, i_1 < i_2 < \dots < i_n\},$$

我们可以证明 $\sharp A_{2n} = (2n-1)!!$. 事实上, i_1 只能取 1, 于是 j_1 有 $2n-1$ 种取法; i_1, j_1 确定之后, i_2 只能取剩下元素中最小者, 于是 j_2 有 $2n-3$ 种取法; 依此类推即得上述结论. 显然, 常序排列 $(1, 2, 3, 4, \dots, 2n-1, 2n) \in A_{2n}$. 设 $A = (a_{ij})$ 为 $2n$ 阶反对称阵, 则

$$\text{Pf}(A) = \sum_{(i_1, j_1, i_2, j_2, \dots, i_n, j_n) \in A_{2n}} (-1)^{\tau(i_1, j_1, i_2, j_2, \dots, i_n, j_n)} a_{i_1 j_1} a_{i_2 j_2} \cdots a_{i_n j_n}$$

称为 A 的 Pfaffian 多项式, 这是一个关于 a_{ij} ($i < j$) 的整系数多项式. 等价地, 我们也可以把 Pfaffian 多项式写成如下的形式:

$$\text{Pf}(A) = \frac{1}{2^n n!} \sum_{(i_1, i_2, i_3, i_4, \dots, i_{2n-1}, i_{2n}) \in S_{2n}} (-1)^{\tau(i_1, i_2, i_3, i_4, \dots, i_{2n-1}, i_{2n})} a_{i_1 i_2} a_{i_3 i_4} \cdots a_{i_{2n-1} i_{2n}}.$$

定理 1.3.5 设 $A = (a_{ij})$ 为 $2n$ 阶反对称阵, 则 $|A| = \text{Pf}(A)^2$.

证明 我们称由常序排列 $(1, 2, 3, 4, \dots, 2n-1, 2n)$ 决定的单项 $a_{12}a_{34} \cdots a_{2n-1, 2n}$ 为常序单项. 由引理 1.3.3 的证明过程不难发现, $\sqrt{|A|}$ 的表达式中含有常序单项, 其符号为 1 或 -1 . 为方便起见, 我们记包含符号等于 1 的常序单项的 $|A|$ 的那个平方根为 $\text{Qr}(A)$, 则由引理 1.3.3 的证明过程可得

$$\text{Qr}(A) = \sum_{j=2}^{2n} (-1)^{\tau_j} a_{1j} \text{Qr}(C_{1j, 1j}). \quad (1)$$

下面我们将用数学归纳法同时证明以下三个结论:

$$t_j \equiv j \pmod{2}, \quad \forall j \geq 2, \quad (2)$$

$$\text{Qr}(A) = \text{Pf}(A), \quad (3)$$

$$\text{Qr}(S_{i, i+1} A S_{i, i+1}) = -\text{Qr}(A), \quad \forall 1 \leq i \leq 2n-1, \quad (4)$$

其中 $S_{i, i+1}$ 表示将第 i 行和第 $i+1$ 行对换的初等方阵. □

1.3.2 Laplace 展开

例 (3.16) 计算 n 阶方阵 $A = (a_{ij}) = \begin{pmatrix} x & 1 & & \\ 1 & x & \ddots & \\ & \ddots & \ddots & 1 \\ & & 1 & x \end{pmatrix}$ 的逆矩阵, 其中空白处元素都是 0.

解 1 由 Laplace 展开, $\Delta_n = \det(A) = x\Delta_{n-1} - \Delta_{n-2}$, 于是 $\Delta_n = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} x & -1 \\ 1 & 0 \end{pmatrix}^n \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

$A^{-1} = \frac{1}{\Delta_n} A^* = \left(\frac{A_{ji}}{\Delta_n} \right)$. 下面求 a_{ij} 的代数余子式 A_{ij} .

① 当 $i = j$ 时, 显然有 $A_{ij} = (-1)^{i+i} \Delta_{i-1} \Delta_{n-i} = \Delta_{i-1} \Delta_{n-i}$.

② 当 $i < j$ 时,

$$A_{ij} = (-1)^{i+j} \left| \begin{array}{c|ccc} \begin{array}{ccc} x & 1 & \ddots \\ 1 & x & \ddots \\ & \ddots & 1 \end{array} & & & \\ \hline & \begin{array}{ccc} 1 & x & 1 \\ & 1 & \ddots \\ & & 1 \end{array} & & \\ \hline & & \begin{array}{ccc} x & 1 & \ddots \\ 1 & \ddots & 1 \\ & & 1 \end{array} & \\ \hline \end{array} \right| = (-1)^{i+j} \Delta_{i-1} \Delta_{n-j}.$$

其中 3 个红色方块依次是 $i-1$ 阶、 $j-i$ 阶、 $n-j$ 阶方阵.

③ 当 $i > j$ 时, 由 A 的对称性可得 $A_{ij} = (-1)^{i+j} \Delta_{j-1} \Delta_{n-i}$.

□

解 2 ① 先证明可以对三对角矩阵 A 进行如下分解:

$$A = \begin{pmatrix} x & 1 & & \\ 1 & x & \ddots & \\ & \ddots & \ddots & 1 \\ & & 1 & x \end{pmatrix} = \begin{pmatrix} 1 & & & \\ a_1 & 1 & & \\ & \ddots & \ddots & \\ & & a_{n-1} & 1 \end{pmatrix} \begin{pmatrix} d_1 & & & \\ & \ddots & & \\ & & d_n & \end{pmatrix} \begin{pmatrix} 1 & b_1 & & \\ & 1 & \ddots & \\ & & \ddots & b_{n-1} \\ & & & 1 \end{pmatrix}.$$

注意到可以对 A 进行 $n-1$ 次行变换化成上三角方阵:

$$\begin{aligned}
 & \underbrace{\begin{pmatrix} 1 & & & & \\ & 1 & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \\ & & & & & -a_{n-1} & 1 \end{pmatrix} \cdots \begin{pmatrix} 1 & & & & \\ & 1 & & & \\ & & -a_2 & 1 & \\ & & & \ddots & \\ & & & & 1 \\ & & & & & 1 \end{pmatrix} \begin{pmatrix} 1 & & & & \\ -a_1 & 1 & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \\ & & & & & 1 \end{pmatrix}}_P A \\
 &= \underbrace{\begin{pmatrix} * & 1 & & & \\ & * & 1 & & \\ & & * & \ddots & \\ & & & \ddots & * & 1 \\ & & & & & * \end{pmatrix}}_U.
 \end{aligned}$$

由初等方阵意义可知

$$\begin{aligned}
 A = P^{-1}U &= \begin{pmatrix} 1 & & & & \\ a_1 & 1 & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \\ & & & & & 1 \end{pmatrix} \begin{pmatrix} 1 & & & & \\ & 1 & & & \\ & a_2 & 1 & & \\ & & & \ddots & \\ & & & & 1 \\ & & & & & 1 \end{pmatrix} \cdots \begin{pmatrix} 1 & & & & \\ & 1 & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \\ & & & & & a_{n-1} & 1 \end{pmatrix} U \\
 &\stackrel{\text{行(列)变换视角}}{=} \begin{pmatrix} 1 & & & & \\ a_1 & 1 & & & \\ & a_2 & 1 & & \\ & & \ddots & \ddots & \\ & & & a_{n-1} & 1 \end{pmatrix} U.
 \end{aligned}$$

再对 U 作类似的列变换即可得所要的分解形式.

② 由 ① 中初等变换的过程可知, 若对 $A [1, \dots, n-1]$ 进行同样的分解, 所得 3 个方阵恰为 ① 中对应方阵的第 $n-1$ 个顺序主子矩阵. 于是 $\Delta_n = d_1 \cdots d_n$, 即 $d_k = \frac{\Delta_k}{\Delta_{k-1}}$. 下面通过待定

系数法求 $a_1, \dots, a_{n-1}, b_1, \dots, b_{n-1}$. 因为

$$a_{ij} = \begin{pmatrix} 0 & \cdots & 0 & a_{i-1}d_{i-1} & d_i & 0 & \cdots & 0 \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ b_{j-1} \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{cases} a_{i-1}b_{i-1}d_{i-1} + d_i, & i = j, \\ b_id_i, & j = i + 1, \\ a_{i-1}d_{i-1}, & i = j + 1, \\ 0, & \text{其他}, \end{cases}.$$

所以

$$a_id_i = b_id_i = 1 \implies a_i = \frac{1}{d_i}, \quad b_i = \frac{1}{d_i} \quad (i = 1, \dots, n-1).$$

故

$$A^{-1} = \begin{pmatrix} 1 & b_1 & & & \\ & 1 & \ddots & & \\ & & \ddots & b_{n-1} & \\ & & & & 1 \end{pmatrix}^{-1} \begin{pmatrix} \frac{1}{d_1} & & & & \\ & \ddots & & & \\ & & \frac{1}{d_n} & & \end{pmatrix} \begin{pmatrix} 1 & & & & \\ a_1 & 1 & & & \\ & \ddots & \ddots & & \\ & & & a_{n-1} & 1 \end{pmatrix}^{-1}.$$

$$\text{记 } C = \begin{pmatrix} 0 & & & & \\ a_1 & 0 & & & \\ & \ddots & \ddots & & \\ & & & a_{n-1} & 0 \end{pmatrix}, \quad D = \begin{pmatrix} 0 & b_1 & & & \\ & 0 & \ddots & & \\ & & \ddots & b_{n-1} & \\ & & & & 0 \end{pmatrix}, \text{ 则 } C^n = D^n = O. \text{ 因此}$$

$$\begin{aligned} \begin{pmatrix} 1 & & & & \\ a_1 & 1 & & & \\ & \ddots & \ddots & & \\ & & & a_{n-1} & 1 \end{pmatrix}^{-1} &= (I + C)^{-1} = \sum_{k=0}^{n-1} (-1)^k C^k, \\ \begin{pmatrix} 1 & b_1 & & & \\ & 1 & \ddots & & \\ & & \ddots & b_{n-1} & \\ & & & & 1 \end{pmatrix}^{-1} &= (I + D)^{-1} = \sum_{k=0}^{n-1} (-1)^k D^k. \end{aligned}$$

□

1.4 相抵标准形

1.4.1 基本理论

定理 1.4.1 (1) 行满秩矩阵只需经过列变换即可化成标准形.

(2) 列满秩矩阵只需经过行变换即可化成标准形.

证明 设 $A \in \mathbb{F}^{m \times n}$.

① 若 A 行满秩, 则 $A = P \begin{pmatrix} I_m & O \end{pmatrix} Q$, 其中 $P \in \mathbb{F}^{m \times m}, Q \in \mathbb{F}^{n \times n}$ 是可逆方阵. 我们有

$$A = P \begin{pmatrix} I_m & O \end{pmatrix} Q = \begin{pmatrix} I_m & O \end{pmatrix} \begin{pmatrix} P & \\ & I_{n-m} \end{pmatrix} Q.$$

② 若 A 列满秩, 则 $A = P \begin{pmatrix} I_n \\ O \end{pmatrix} Q$, 其中 $P \in \mathbb{F}^{m \times m}, Q \in \mathbb{F}^{n \times n}$ 是可逆方阵. 我们有

$$A = P \begin{pmatrix} I_n \\ O \end{pmatrix} Q = P \begin{pmatrix} Q & \\ & I_{m-n} \end{pmatrix} \begin{pmatrix} I_n \\ O \end{pmatrix}.$$

□

1.4.2 线性方程组

定理 1.4.2 设 $A \in \mathbb{F}^{m \times n}, \mathbf{b} \in \mathbb{F}^{m \times 1}$.

(1) $A\mathbf{x} = \mathbf{b}$ 有解 $\mathbf{x} \in \mathbb{F}^{n \times 1} \iff \text{rank} \begin{pmatrix} A & \mathbf{b} \end{pmatrix} = \text{rank}(A)$.

(2) 设 $A\mathbf{x} = \mathbf{b}$ 有特解 α_0 . 则 $A\mathbf{x} = \mathbf{b}$ 的解集为

$$\{\alpha_0 + \lambda_1 \alpha_1 + \lambda_2 \alpha_2 + \cdots + \lambda_k \alpha_k \mid \lambda_1, \lambda_2, \cdots, \lambda_k \in \mathbb{F}\}.$$

其中 $\alpha_1, \alpha_2, \cdots, \alpha_k$ 是 $A\mathbf{x} = \mathbf{0}$ 的基础解系, $k = n - \text{rank}(A)$.

(3) 特别地², 当 $\text{rank}(A) = n$ (即 A 列满秩) 时, $A\mathbf{x} = \mathbf{b}$ 的解是唯一的.

下面从线性空间理论与矩阵论两种方法证明 (1).

²即在 $\text{rank} \begin{pmatrix} A & \mathbf{b} \end{pmatrix} = \text{rank}(A)$ 的前提下.

证明 1 设 A 的行向量为 $\alpha_1, \alpha_2, \dots, \alpha_n$, 则

$$\begin{aligned}
 Ax = b & \text{有解} \\
 & \Downarrow \\
 b & \in \langle \alpha_1, \alpha_2, \dots, \alpha_n \rangle \\
 & \Downarrow \\
 \langle \alpha_1, \alpha_2, \dots, \alpha_n, b \rangle & \subseteq \langle \alpha_1, \alpha_2, \dots, \alpha_n \rangle \\
 & \Downarrow \\
 \langle \alpha_1, \alpha_2, \dots, \alpha_n, b \rangle & = \langle \alpha_1, \alpha_2, \dots, \alpha_n \rangle \\
 & \Downarrow \\
 \dim \langle \alpha_1, \alpha_2, \dots, \alpha_n, b \rangle & = \dim \langle \alpha_1, \alpha_2, \dots, \alpha_n \rangle \\
 & \Downarrow \\
 \text{rank} \begin{pmatrix} A & b \end{pmatrix} & = \text{rank}(A).
 \end{aligned}$$

□

证明 2 \Rightarrow : 若 x_0 是 $Ax = b$ 的一个解, 则

$$\text{rank} \begin{pmatrix} A & b \end{pmatrix} = \text{rank} \begin{pmatrix} A & Ax_0 \end{pmatrix} \xrightarrow{\text{初等变换}} \text{rank} \begin{pmatrix} A & \mathbf{0} \end{pmatrix} = \text{rank}(A).$$

\Leftarrow : 设 $A = P \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} Q$. 作换元 $y = Qx$, 则 $Ax = b$ 同解变形为

$$\begin{pmatrix} I_r & O \\ O & O \end{pmatrix} y = P^{-1}b =: \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}.$$

【上述换元也可由矩阵运算直接得到:

$$\text{rank} \begin{pmatrix} A & b \end{pmatrix} = \text{rank} \left(P^{-1} \begin{pmatrix} A & b \end{pmatrix} \begin{pmatrix} Q^{-1} & \\ & 1 \end{pmatrix} \right) = \text{rank} \begin{pmatrix} I_r & O & \beta_1 \\ O & O & \beta_2 \end{pmatrix}.$$

其中 $P^{-1}b =: \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}$. 而 $\begin{pmatrix} I_r & O & \beta_1 \\ O & O & \beta_2 \end{pmatrix}$ 相抵于 $\begin{pmatrix} I_r & O & O \\ O & O & \beta_2 \end{pmatrix}$, 结合 $\text{rank} \begin{pmatrix} A & b \end{pmatrix} = \text{rank}(A)$ 可得 $\beta_2 = \mathbf{0}$.】

因为 $y = \begin{pmatrix} \beta_1 \\ \mathbf{0} \end{pmatrix}$ 满足方程 $\begin{pmatrix} I_r & O \\ O & O \end{pmatrix} y = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}$, 所以 $x = Q^{-1} \begin{pmatrix} \beta_1 \\ \mathbf{0} \end{pmatrix}$ 是 $Ax = b$ 的一个解.

□

例 (4.5) 对于任意 $A \in \mathbb{C}^{m \times n}$, $\text{rank}(A^H A) = \text{rank}(A)$.

证明 1 由 $Ax = 0 \implies A^H Ax = 0 \implies x^H A^H Ax = 0 \implies \|Ax\|^2 = 0 \implies Ax = 0$ 可知线性方程组 $Ax = 0$ 与 $A^H Ax = 0$ 同解, 因此 $\text{rank}(A^H A) = \text{rank}(A)$. \square

证明 2 由常用秩不等式可知 $\text{rank}(A^H A) \leq \text{rank}(A)$, 下面只需再证 $\text{rank}(A^H A) \geq \text{rank}(A)$. 设 A 有 r 阶可逆子矩阵 $A \begin{bmatrix} i_1, \dots, i_r \\ j_1, \dots, j_r \end{bmatrix}$, 我们断言 $A^H A$ 也有 r 阶可逆子矩阵: 由 Binet–Cauchy 公式,

$$\begin{aligned} \det((A^H A) \begin{bmatrix} j_1, \dots, j_r \\ j_1, \dots, j_r \end{bmatrix}) &= \det(A^H \begin{bmatrix} j_1, \dots, j_r \\ 1, \dots, m \end{bmatrix} A \begin{bmatrix} 1, \dots, m \\ j_1, \dots, j_r \end{bmatrix}) \\ &= \sum_{1 \leq k_1 < \dots < k_r \leq n} \det(A^H \begin{bmatrix} j_1, \dots, j_r \\ k_1, \dots, k_r \end{bmatrix}) \det(A \begin{bmatrix} k_1, \dots, k_r \\ j_1, \dots, j_r \end{bmatrix}) \\ &= \sum_{1 \leq k_1 < \dots < k_r \leq n} \left| \det(A \begin{bmatrix} k_1, \dots, k_r \\ j_1, \dots, j_r \end{bmatrix}) \right|^2 > 0. \end{aligned}$$

\square

推论 1.4.3 (1) 若 A 行满秩, 则 AA^H 可逆; 若 A 列满秩, 则 $A^H A$ 可逆.

(2) 更常用的结论: 对于任意 $A \in \mathbb{R}^{m \times n}$, $\text{rank}(A^T A) = \text{rank}(A)$.

1.4.3 广义逆矩阵

定义 1.4.4 设 $m \times n$ 矩阵 $A = P \begin{pmatrix} I & O \\ O & O \end{pmatrix} Q$, 其中 P, Q 都是可逆方阵, 则 $n \times m$ 矩阵

$$B = Q^{-1} \begin{pmatrix} I & O \\ O & O \end{pmatrix} P^{-1}$$

称为 A 的一个广义逆矩阵.

定理 1.4.5 矩阵 B 是 A 的一个广义逆矩阵 $\iff ABA = A$ 且 $BAB = B$.

证明 \Rightarrow : 由定义直接代入即可验证.

\Leftarrow : 设 $A = P \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} Q$, 则

$$\begin{aligned} \begin{cases} ABA = A \\ BAB = B \end{cases} &\iff \begin{cases} P \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} QBP \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} Q = P \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} Q \\ BP \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} QB = B \end{cases} \\ &\iff \begin{cases} \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} QBP \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} \\ BP \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} QB = B \end{cases} \end{aligned}$$

这启发我们设 $B = Q^{-1} \begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix} P^{-1}$. 于是

$$\begin{aligned} \begin{cases} ABA = A \\ BAB = B \end{cases} &\iff \begin{cases} \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} \begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix} \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} \\ \begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix} \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} \begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix} = \begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix} \end{cases} \\ &\xrightarrow{\text{矩阵乘法}} \begin{cases} B_1 = I_r \\ B_4 = B_3 B_2 \end{cases} \\ &\iff B = Q^{-1} \begin{pmatrix} I_r & B_2 \\ B_3 & B_3 B_2 \end{pmatrix} P^{-1} = Q^{-1} \begin{pmatrix} I_r & O \\ B_3 & I \end{pmatrix} \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} \begin{pmatrix} I_r & B_2 \\ O & I \end{pmatrix} P^{-1} \\ &\iff B = Q^{-1} \begin{pmatrix} I_r & O \\ -B_3 & I \end{pmatrix}^{-1} \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} \begin{pmatrix} I_r & -B_2 \\ O & I \end{pmatrix}^{-1} P^{-1} \end{aligned}$$

再由

$$\begin{pmatrix} I_r & -B_2 \\ O & I \end{pmatrix} \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} \begin{pmatrix} I_r & O \\ -B_3 & I \end{pmatrix} = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$$

可得

$$A = P \begin{pmatrix} I_r & -B_2 \\ O & I \end{pmatrix} \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} \begin{pmatrix} I_r & O \\ -B_3 & I \end{pmatrix} Q.$$

故取 $\tilde{P} = P \begin{pmatrix} I_r & -B_2 \\ O & I \end{pmatrix}$, $\tilde{Q} = \begin{pmatrix} I_r & O \\ -B_3 & I \end{pmatrix} Q$, 就有

$$A = \tilde{P} \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} \tilde{Q}, \quad B = \tilde{Q}^{-1} \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} \tilde{P}^{-1}.$$

□

1.4.4 用相抵研究特殊矩阵

¶ 幂等方阵

例 1.4.6 求所有满足 $A^k = A$ 的 $A \in \mathbb{F}^{n \times n}$, 其中 $k \in \mathbb{N}^*$.

解 设 $A = P \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} Q$, 其中 $P, Q \in \mathbb{F}^{n \times n}$ 是可逆方阵. 由此可将 A 重新写成

$$A = P \left(\begin{pmatrix} I_r & O \\ O & O \end{pmatrix} QP \right) P^{-1} = P \begin{pmatrix} R_1 & R_2 \\ O & O \end{pmatrix} P^{-1},$$

其中 $R_1 \in \mathbb{F}^{r \times r}$, $R_2 \in \mathbb{F}^{r \times (n-r)}$. 于是

$$A^k = P \begin{pmatrix} R_1 & R_2 \\ O & O \end{pmatrix}^k P^{-1}.$$

归纳易得 $\begin{pmatrix} R_1 & R_2 \\ O & O \end{pmatrix}^k = \begin{pmatrix} R_1^k & R_1^{k-1}R_2 \\ O & O \end{pmatrix}$. 因此

$$\begin{aligned} A^k = A &\iff P \begin{pmatrix} R_1^k & R_1^{k-1}R_2 \\ O & O \end{pmatrix} P^{-1} = P \begin{pmatrix} R_1 & R_2 \\ O & O \end{pmatrix} P^{-1} \\ &\iff \begin{pmatrix} R_1^k & R_1^{k-1}R_2 \\ O & O \end{pmatrix} = \begin{pmatrix} R_1 & R_2 \\ O & O \end{pmatrix} \\ &\iff R_1^{k-1} \begin{pmatrix} R_1 & R_2 \end{pmatrix} = \begin{pmatrix} R_1 & R_2 \end{pmatrix}. \end{aligned}$$

由最后一个等式及 $\begin{pmatrix} R_1 & R_2 \end{pmatrix}$ 行满秩可得 $\text{rank}(R_1^{k-1}) = r$, 因此 R_1^{k-1} 可逆, 从而 R_1 可逆, 将 $R_1^k = R_1$ 两边同乘 R_1^{-1} 即得 $R_1^{k-1} = I_r$. 【也可以由行满秩矩阵有右逆, 等式两边同乘 $\begin{pmatrix} R_1 & R_2 \end{pmatrix}$ 的一个右逆直接得到 $R^{k-1} = I_r$.】故

$$\begin{aligned} A &= P \begin{pmatrix} R_1 & R_2 \\ O & O \end{pmatrix} P^{-1} = P \underbrace{\begin{pmatrix} I & -R_1^{-1}R_2 \\ O & I \end{pmatrix}}_{\tilde{P}} \begin{pmatrix} R_1 & O \\ O & O \end{pmatrix} \underbrace{\begin{pmatrix} I & R_1^{-1}R_2 \\ O & I \end{pmatrix}}_{\tilde{P}^{-1}} P^{-1} \\ &= \tilde{P} \begin{pmatrix} R_1 & O \\ O & O \end{pmatrix} \tilde{P}^{-1}, \quad \text{其中 } R_1^{k-1} = I_r. \end{aligned}$$

反过来, 对于任意满足 $R^{k-1} = I_r$ ($0 \leq r \leq n$) 的方阵 R 与任意可逆方阵 $P \in \mathbb{F}^{n \times n}$, $A = P \begin{pmatrix} R & O \\ O & O \end{pmatrix} P^{-1}$ 都满足 $A^k = A$, 因此这样的 A 即是所求. \square

¶ 幂幺方阵

例 1.4.7 求所有满足 $A^2 = I$ 的 $A \in \mathbb{F}^{n \times n}$.

解 1 $A^2 = I \iff \left(\frac{A+I}{2}\right)^2 = \frac{A+I}{2}$, 由例 1.4.6 可知 $\frac{A+I}{2} = P \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} P^{-1}$, 也即 $A = P \begin{pmatrix} 2I_r & O \\ O & O \end{pmatrix} P^{-1} - I$, 其中 $P \in \mathbb{F}^{n \times n}$ 是可逆方阵, $0 \leq r \leq n$. 进一步化简可得

$$A = P \begin{pmatrix} 2I_r & O \\ O & O \end{pmatrix} P^{-1} - PIP^{-1} = P \begin{pmatrix} I_r & \\ & -I_{n-r} \end{pmatrix} P^{-1}, \quad r = \text{rank}(I + A).$$

\square

解 2 设 λ 是 A 的特征值, α 是与之对应的特征向量, 则由 $A^2 = I$ 得

$$A^2\alpha = \lambda^2\alpha = \alpha \implies \lambda = \pm 1.$$

-1 的几何重数

$$m_1 = n - \text{rank}(I + A),$$

1 的几何重数

$$m_2 = n - \text{rank}(I - A).$$

而 $A^2 = I \iff (I + A)(I - A) = O$, 由 Sylvester 秩不等式有

$$\text{rank}(I + A) + \text{rank}(I - A) \leq \text{rank}((I + A)(I - A)) + n = n,$$

故

$$m_1 + m_2 \geq 2n - n = n \implies m_1 + m_2 = n = A \text{ 的代数重数}.$$

因此 A 可相似对角化为 $\begin{pmatrix} I_r & \\ & -I_{n-r} \end{pmatrix}$, 其中 $r = \text{rank}(I + A)$. □

¶ 幂零方阵

例 1.4.8 求所有满足 $A^2 = O$ 的 $A \in \mathbb{F}^{n \times n}$.

证明 设 $A = P \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} Q$, 其中 $P, Q \in \mathbb{F}^{n \times n}$ 是可逆方阵. 由此可将 A 重新写成

$$A = P \left(\begin{pmatrix} I_r & O \\ O & O \end{pmatrix} QP \right) P^{-1} = P \begin{pmatrix} R_1 & R_2 \\ O & O \end{pmatrix} P^{-1},$$

其中 $R_1 \in \mathbb{F}^{r \times r}$, $R_2 \in \mathbb{F}^{r \times (n-r)}$. 于是

$$A^2 = P \begin{pmatrix} R_1^2 & R_1 R_2 \\ O & O \end{pmatrix} P^{-1}.$$

因此

$$A^2 = O \iff P \begin{pmatrix} R_1^2 & R_1 R_2 \\ O & O \end{pmatrix} P^{-1} = P \begin{pmatrix} R_1 & R_2 \\ O & O \end{pmatrix} P^{-1} \iff R_1 \begin{pmatrix} R_1 & R_2 \end{pmatrix} = O.$$

因为 $\begin{pmatrix} R_1 & R_2 \end{pmatrix}$ 行满秩, 可在最后一个等式两边同乘它的一个右逆, 得到 $R_1 = O$, 进而 R_2 行满秩, $r \leq n - r$. 设 $R_2 = \begin{pmatrix} O & I_r \end{pmatrix} \tilde{Q}$, 其中 $\tilde{Q} \in \mathbb{F}^{(n-r) \times (n-r)}$ 是可逆方阵. 故

$$\begin{aligned} A &= P \begin{pmatrix} O & R_2 \\ O & O \end{pmatrix} P^{-1} = P \begin{pmatrix} O & (O \ I_r) \tilde{Q} \\ O & O \end{pmatrix} P^{-1} \\ &= \underbrace{P \begin{pmatrix} I & O \\ O & \tilde{Q}^{-1} \end{pmatrix}}_{\tilde{P}} \begin{pmatrix} O & I_r \\ O & O \end{pmatrix} \underbrace{\begin{pmatrix} I & O \\ O & \tilde{Q} \end{pmatrix} P^{-1}}_{\tilde{P}^{-1}} = \tilde{P} \begin{pmatrix} O & I_r \\ O & O \end{pmatrix} \tilde{P}^{-1}. \end{aligned}$$

反过来, 对于任意可逆方阵 $P \in \mathbb{F}^{n \times n}$ 和 $0 \leq r \leq \frac{n}{2}$, $A = P \begin{pmatrix} O & I_r \\ O & O \end{pmatrix} P^{-1}$ 都满足 $A^2 = O$, 因此这样的 A 即是所求. □

注 1.4.9 在 Frobenius 秩不等式中代入 $B = I, C = A$, 可得 $2\text{rank}(A) \leq n$, 也能得到上面证明中的 $r \leq n - r$.

1.5 秩不等式与等式

例 (4.6) $\text{rank} \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \text{rank}(A) \iff$ 存在矩阵 X, Y , 使得

$$AX = B, \quad YA = C, \quad YAX = D.$$

证明 \Leftarrow : 由初等变换可得

$$\text{rank} \begin{pmatrix} A & AX \\ YA & YAX \end{pmatrix} = \text{rank} \begin{pmatrix} A & O \\ YA & O \end{pmatrix} = \text{rank} \begin{pmatrix} A & O \\ O & O \end{pmatrix} = \text{rank}(A).$$

\Rightarrow : 设 $A = P \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} Q$, 其中 P, Q 是可逆方阵. 则可设

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} P & \\ & I \end{pmatrix} \begin{pmatrix} I_r & O & B_1 & B_2 \\ O & O & B_3 & B_4 \\ C_1 & C_2 & D_1 & D_2 \\ C_3 & C_4 & D_3 & D_4 \end{pmatrix} \begin{pmatrix} Q & \\ & I \end{pmatrix}.$$

注意到左上角分块 I_r 可逆, 运用 Schur 公式 (本质上即初等变换) 可得 Schur 补

$$\begin{pmatrix} O & B_3 & B_4 \\ C_2 & D_1 & D_2 \\ C_4 & D_3 & D_1 \end{pmatrix} - \begin{pmatrix} O \\ C_1 \\ C_3 \end{pmatrix} I_r^{-1} \begin{pmatrix} O & B_1 & B_2 \end{pmatrix} \xlongequal[\text{rank}(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}) = \text{rank}(A)]{} O,$$

也即

$$\begin{pmatrix} O & B_3 & B_4 \\ C_2 & D_1 & D_2 \\ C_4 & D_3 & D_1 \end{pmatrix} = \begin{pmatrix} O \\ C_1 \\ C_3 \end{pmatrix} \begin{pmatrix} O & B_1 & B_2 \end{pmatrix} = \begin{pmatrix} O & O & O \\ O & C_1 B_1 & C_1 B_2 \\ O & C_3 B_1 & C_3 B_2 \end{pmatrix}.$$

故

$$B = P \begin{pmatrix} B_1 & B_2 \\ O & O \end{pmatrix}, \quad C = \begin{pmatrix} C_1 & O \\ C_3 & O \end{pmatrix} Q, \quad D = \begin{pmatrix} C_1 \\ C_3 \end{pmatrix} \begin{pmatrix} B_1 & B_2 \end{pmatrix}.$$

下面通过解矩阵方程求所需的 X, Y .

$$AX = B \iff P \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} QX = P \begin{pmatrix} B_1 & B_2 \\ O & O \end{pmatrix} \iff \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} QX = \begin{pmatrix} B_1 & B_2 \\ O & O \end{pmatrix},$$

$$YA = C \iff YP \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} Q = \begin{pmatrix} C_1 & O \\ C_3 & O \end{pmatrix} Q \iff YP \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} = \begin{pmatrix} C_1 & O \\ C_3 & O \end{pmatrix}.$$

由此可见 $QX = \begin{pmatrix} B_1 & B_2 \\ O & O \end{pmatrix}$, $YP = \begin{pmatrix} C_1 & O \\ C_3 & O \end{pmatrix}$ 即 $X = Q^{-1} \begin{pmatrix} B_1 & B_2 \\ O & O \end{pmatrix}$, $Y = \begin{pmatrix} C_1 & O \\ C_3 & O \end{pmatrix} P^{-1}$ 满足上面两个方程. 此外,

$$YAX = \begin{pmatrix} C_1 & O \\ C_3 & O \end{pmatrix} \begin{pmatrix} I_r \\ O \end{pmatrix} \begin{pmatrix} I_r & O \end{pmatrix} \begin{pmatrix} B_1 & B_2 \\ O & O \end{pmatrix} = \begin{pmatrix} C_1 \\ C_3 \end{pmatrix} \begin{pmatrix} B_1 & B_2 \end{pmatrix} = D.$$

这就证明了所构造的 X, Y 满足要求. □

例 (4.7) $\text{rank}(AB) = \text{rank}(A) \iff$ 存在矩阵 X , 使得 $ABX = A$.

证明 \Leftarrow : 设 X 满足 $ABX = A$. 由常用的秩不等式可知

$$\text{rank}(ABX) \leq \text{rank}(AB) \leq \text{rank}(A) = \text{rank}(ABX),$$

这就迫使 $\text{rank}(AB) = \text{rank}(A)$.

\Rightarrow : 设 $A = P \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} Q$, 其中 P, Q 是可逆方阵. 记 $QB = \begin{pmatrix} R_1 & R_2 \\ * & * \end{pmatrix}$, 则

$$AB = P \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} QB = P \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} \begin{pmatrix} R_1 & R_2 \\ * & * \end{pmatrix} = P \begin{pmatrix} R_1 & R_2 \\ O & O \end{pmatrix}.$$

由 $\text{rank}(AB) = r$ 可推出 $\begin{pmatrix} R_1 & R_2 \end{pmatrix}$ 行满秩, 故其必有右逆. 设 Y 满足 $\begin{pmatrix} R_1 & R_2 \end{pmatrix} Y = I_r$, 观察到 $ABY = P \begin{pmatrix} R_1 & R_2 \\ O & O \end{pmatrix} Y = P \begin{pmatrix} I_r \\ O \end{pmatrix}$. 为了构造满足 $ABX = A$ 的 X , 只需取 $X = \begin{pmatrix} Y & O \end{pmatrix} Q$. □

推论 1.5.1 由此可得

$$\begin{aligned} \text{rank}(AB) = \text{rank}(B) &\iff \text{rank}(B^T A^T) = \text{rank}(B^T) \iff \text{存在 } X \text{ 使得 } B^T A^T X = B^T \\ &\xLeftrightarrow{Y=X^T} \text{存在 } Y \text{ 使得 } YAB = B. \end{aligned}$$

例 (4.8) $\text{rank} \begin{pmatrix} A & B \\ O & D \end{pmatrix} = \text{rank}(A) + \text{rank}(D) \iff$ 存在矩阵 X, Y , 使得 $AX + YD = B$.

证明 \Leftarrow : 由初等变换可得

$$\text{rank} \begin{pmatrix} A & AX + YD \\ O & D \end{pmatrix} = \text{rank} \begin{pmatrix} A & YD \\ O & D \end{pmatrix} = \text{rank} \begin{pmatrix} A & O \\ O & D \end{pmatrix} = \text{rank}(A) + \text{rank}(D).$$

\Rightarrow : 设 $A = P_1 \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} Q_1$, $D = P_2 \begin{pmatrix} I_s & O \\ O & O \end{pmatrix} Q_2$, 其中 P_1, Q_1, P_2, Q_2 均为可逆方阵. 则

可设

$$\begin{pmatrix} A & B \\ O & D \end{pmatrix} = \begin{pmatrix} P_1 & \\ & P_2 \end{pmatrix} \begin{pmatrix} I_r & B_1 & B_2 \\ O & B_3 & B_4 \\ \hline & I_s & \\ & & O \end{pmatrix} \begin{pmatrix} Q_1 & \\ & Q_2 \end{pmatrix},$$

其中 $B = P_1 \begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix} Q_2$. 由初等变换可知

$$\text{rank} \begin{pmatrix} I_r & B_1 & B_2 \\ O & B_3 & B_4 \\ \hline & I_s & \\ & & O \end{pmatrix} = \text{rank} \begin{pmatrix} I_r & & \\ O & B_4 & \\ \hline & I_s & \\ & & O \end{pmatrix} = \text{rank} \begin{pmatrix} I_r & & \\ & I_s & \\ \hline & O & B_4 \\ & & O \end{pmatrix},$$

从 $\text{rank} \begin{pmatrix} A & B \\ O & D \end{pmatrix} = r + s$ 就推出 $B_4 = O$. 接下来只需构造如下矩阵方程:

$$\underbrace{\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}}_{\text{选前 } r \text{ 行}} \begin{pmatrix} * & * \\ * & * \end{pmatrix} + \begin{pmatrix} * & * \\ * & * \end{pmatrix} \underbrace{\begin{pmatrix} I_s & O \\ O & O \end{pmatrix}}_{\text{选前 } s \text{ 列}} = \begin{pmatrix} B_1 & B_2 \\ B_3 & O \end{pmatrix}.$$

易知可按如下方式构造:

$$\begin{pmatrix} I_r & O \\ O & O \end{pmatrix} \begin{pmatrix} B_1 & B_2 \\ O & O \end{pmatrix} + \begin{pmatrix} O & O \\ B_3 & O \end{pmatrix} \begin{pmatrix} I_s & O \\ O & O \end{pmatrix} = \begin{pmatrix} B_1 & B_2 \\ B_3 & O \end{pmatrix}.$$

进而

$$P_1 \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} \begin{pmatrix} B_1 & B_2 \\ O & O \end{pmatrix} Q_2 + P_1 \begin{pmatrix} O & O \\ B_3 & O \end{pmatrix} \begin{pmatrix} I_s & O \\ O & O \end{pmatrix} Q_2 = P_1 \begin{pmatrix} B_1 & B_2 \\ B_3 & O \end{pmatrix} Q_2 = B.$$

也即

$$B = \underbrace{A Q^{-1} \begin{pmatrix} B_1 & B_2 \\ O & O \end{pmatrix} Q_2}_{X} + \underbrace{P_1 \begin{pmatrix} O & O \\ B_3 & O \end{pmatrix} P_2^{-1} D}_{Y}.$$

□

推论 1.5.2 此结论可理解成已知矩阵 A, B, D , 问何时线性方程组 $AX + YD = B$ 有解 X, Y .

例 (4.9) 对于任意 $A \in \mathbb{F}^{m \times n}$, $B \in \mathbb{F}^{n \times p}$, $C \in \mathbb{F}^{p \times q}$, 有 Frobenius 秩不等式

$$\text{rank}(AB) + \text{rank}(BC) \leq \text{rank}(ABC) + \text{rank}(B).$$

特别地, 当 $B = I_n$ 时, 上式成为 Sylvester 秩不等式

$$\text{rank}(A) + \text{rank}(C) \leq \text{rank}(AC) + n.$$

证明 由初等变换可得

$$\begin{aligned} \text{rank}(AB) + \text{rank}(BC) &= \text{rank} \begin{pmatrix} BC & \\ & AB \end{pmatrix} \leq \text{rank} \begin{pmatrix} BC & B \\ & AB \end{pmatrix} \\ &= \text{rank} \begin{pmatrix} & B \\ -ABC & AB \end{pmatrix} = \text{rank} \begin{pmatrix} & B \\ -ABC & \end{pmatrix} \\ &= \text{rank}(ABC) + \text{rank}(B). \end{aligned}$$

等号成立当且仅当 $\text{rank} \begin{pmatrix} BC & \\ & AB \end{pmatrix} = \text{rank} \begin{pmatrix} BC & B \\ & AB \end{pmatrix}$. 由例 (4.8) 结论, 这等价于存在矩阵 X, Y , 使得 $B = BCX + YAB$. □

1.6 Smith 标准形

定义 1.6.1 设 P 是整数方阵. 若 P 是可逆方阵, 并且 P^{-1} 也是整数方阵, 则 P 称为 \mathbb{Z} 上的模方阵. 设 $A, B \in \mathbb{Z}^{m \times n}$. 若存在 \mathbb{Z} 上的模方阵 P, Q , 使得 $A = PBQ$, 则称 A 与 B 在 \mathbb{Z} 上模相抵.

定理 1.6.2 设 P 是整数方阵. P 是 \mathbb{Z} 上的模方阵当且仅当 $\det(P) = \pm 1$.

定义 1.6.3 设 $P \in \mathbb{F}[x]^{n \times n}$. 当且仅当 $0 \neq \det(P) \in \mathbb{F}$ 时, $P^{-1} \in \mathbb{F}[x]^{n \times n}$, P 称为 $\mathbb{F}[x]$ 上的模方阵. 设 $A, B \in \mathbb{F}[x]^{m \times n}$. 若存在 $\mathbb{F}[x]$ 上的模方阵 P, Q , 使得 $A = PBQ$, 则称 A 与 B 在 $\mathbb{F}[x]$ 上模相抵.

定义 1.6.4 设 $P \in R^{n \times n}$, 其中 R 是整环. 当且仅当 $\det(P)$ 是 R 中的单位 (即可逆元) 时, $P^{-1} \in R^{n \times n}$, P 称为 R 上的模方阵. 设 $A, B \in R^{m \times n}$. 若存在 R 上的模方阵 P, Q , 使得 $A = PBQ$, 则称 A 与 B 在 R 上模相抵.

例 1.6.5 $\mathbb{Z}^{2 \times 2}$ 中的模方阵全体在矩阵乘法运算下构成群 G , G 可由 $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ 和 $\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$ 生成.

证明 ① 先证明所有 2 阶初等模方阵都在 G 中: 对任意 $n \in \mathbb{N}$,

$$\begin{aligned} \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} &= \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^n, & \begin{pmatrix} 1 & -n \\ 0 & 1 \end{pmatrix} &= \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}^n, \\ \begin{pmatrix} 1 & 0 \\ n & 1 \end{pmatrix} &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \\ \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} &= \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}, \\ \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \end{aligned}$$

② 对任意 $A = \begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix} \in G$, 总可以通过初等模变换将其变成单位方阵, 从而 $A \in G$:

- 用更相减损术对 A 进行初等模变换, 得到 $B = \begin{pmatrix} b_1 & b_2 \\ b_3 & b_4 \end{pmatrix}$, 其中 $b_1 > 0$, $|b_3| < b_1$.
- 不断重复上一步骤, 最终得到 $C = \begin{pmatrix} c_1 & c_2 \\ 0 & c_4 \end{pmatrix}$. 因为 $\det(C) = c_1 c_4 = \pm 1$, 所以 $c_1, c_4 \in \{1, -1\}$.
- 继续作初等模变换, 得到 $D = \begin{pmatrix} 1 & d_2 \\ 0 & 1 \end{pmatrix}$, 进而可得 I . □

例 1.6.6 $\mathbb{F}[x]^{2 \times 2}$ 中的模方阵全体在矩阵乘法运算下构成群 G , G 可由 $\begin{pmatrix} 1 & \lambda(x) \\ 0 & 1 \end{pmatrix}$, $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ 和 $\begin{pmatrix} a_1 & 0 \\ 0 & a_2 \end{pmatrix}$ 这三类矩阵生成, 其中 $\lambda(x) \in \mathbb{F}[x]$, $a_1, a_2 \in \mathbb{F} \setminus \{0\}$.

证明 ① 所有 2 阶初等模方阵都在 G 中: 第一类和第二类初等模变换方阵已经被包含; 第三类初等模变换方阵可由前两个生成元得到.

② 对任意 $A = \begin{pmatrix} a_1(x) & a_2(x) \\ a_3(x) & a_4(x) \end{pmatrix} \in G$, 总可以通过初等模变换将其变成单位方阵, 从而 $A \in G$:

• 用初等模方阵对 A 进行多项式带余除法, 可以得到 $B = \begin{pmatrix} b_1(x) & b_2(x) \\ 0 & b_4(x) \end{pmatrix}$. 这时有 $\det(B) = b_1(x)b_4(x) \in \mathbb{F} \setminus \{0\}$. \square

定理 1.6.7 对于任意整数矩阵 $A \in \mathbb{Z}^{m \times n}$, 存在 \mathbb{Z} 上的模方阵 P, Q , 使得

$$PAQ = \text{diag}(d_1, d_2, \dots, d_r, O),$$

其中 d_1, d_2, \dots, d_r 是正整数, 并且每个 d_k 整除 d_{k+1} , $1 \leq k \leq r-1$.

证明 对矩阵的维数 (m, n) 使用数学归纳法. 不妨设 $A \neq O$ 并且 $mn \geq 2$. 当 $A = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$ 时,

由辗转相除法, 存在模方阵 P , 使得 $PA = \begin{pmatrix} d \\ 0 \end{pmatrix}$, 其中 $d = \gcd(a_1, a_2)$. 当 $A = \begin{pmatrix} a_1 & a_2 \end{pmatrix}$ 时,

同样存在模方阵 Q , 使得 $AQ = \begin{pmatrix} d & 0 \end{pmatrix}$. 当 $m \geq 2$ 或 $n \geq 2$ 时, 设 $B = (b_{ij})$ 与 A 模相抵, $b_{11} > 0$, 并且使得 b_{11} 尽可能小. 我们断言 b_{11} 整除所有 b_{ij} . 分三种情形来反证.

(1) 存在 b_{i1} 不能被 b_{11} 整除. 存在模方阵 P , 使得 $PB = (c_{ij})$, 其中 $c_{11} = \gcd(b_{11}, b_{i1}) < b_{11}$, 与 b_{11} 的最小性矛盾.

(2) 存在 b_{1j} 不能被 b_{11} 整除. 存在模方阵 Q , 使得 $BQ = (c_{ij})$, 其中 $c_{11} = \gcd(b_{11}, b_{1j}) < b_{11}$, 与 b_{11} 的最小性矛盾.

(3) 存在 $i \neq 1$ 和 $j \neq 1$, 使得 b_{ij} 不能被 b_{11} 整除. 设 $BT_{1j}(-\frac{b_{1j}}{b_{11}})T_{j1}(1) = (c_{ij})$, 则 $c_{11} = b_{11}$, $c_{i1} = (1 - \frac{b_{1j}}{b_{11}})b_{i1} + b_{ij}$ 不能被 c_{11} 整除. 化为情形 (1), 矛盾. 断言证毕.

由于 b_{11} 整除所有 b_{ij} , 故存在初等模方阵 P, Q , 使得 $PBQ = \begin{pmatrix} b_{11} & \\ & b_{11}C \end{pmatrix}$. 对整数方阵 C 应用归纳假设, 即可完成证明. \square

定义 1.6.8 设 A 是整数矩阵. A 的所有 k 阶子式的最大公约数 D_k 称为 A 的 k 阶行列式因子. 特别规定: $D_0 = 1$; 当 $k > \text{rank}(A)$ 时, $D_k = 0$.

定理 1.6.9 $D_k = d_1 d_2 \cdots d_k$, $1 \leq k \leq r$, 其中 d_1, d_2, \dots, d_r 如定理 1.6.7 所述.

证明 设 P, Q 如定理 1.6.7 所述, 则

$$\text{diag}(d_1, d_2, \dots, d_k) = P \begin{bmatrix} 1 & 2 & \cdots & k \\ 1 & 2 & \cdots & m \end{bmatrix} A Q \begin{bmatrix} 1 & 2 & \cdots & n \\ 1 & 2 & \cdots & k \end{bmatrix}.$$

根据 Binet–Cauchy 公式,

$$d_1 d_2 \cdots d_k = \sum_{\substack{1 \leq i_1 < i_2 < \cdots < i_k \leq m \\ 1 \leq j_1 < j_2 < \cdots < j_k \leq n}} \det \left(P \begin{bmatrix} 1 & 2 & \cdots & k \\ i_1 & i_2 & \cdots & i_k \end{bmatrix} \right) \det \left(A \begin{bmatrix} i_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{bmatrix} \right) \det \left(Q \begin{bmatrix} j_1 & j_2 & \cdots & j_k \\ 1 & 2 & \cdots & k \end{bmatrix} \right).$$

可被 D_k 整除. 另一方面, 对于任意 $1 \leq i_1 < i_2 < \cdots < i_k \leq m$, $1 \leq j_1 < j_2 < \cdots < j_k \leq n$,

$$A \begin{bmatrix} i_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{bmatrix} = P^{-1} \begin{bmatrix} i_1 & i_2 & \cdots & i_k \\ 1 & 2 & \cdots & r \end{bmatrix} \text{diag}(d_1, d_2, \dots, d_r) Q^{-1} \begin{bmatrix} 1 & 2 & \cdots & r \\ j_1 & j_2 & \cdots & j_k \end{bmatrix}.$$

再根据 Binet–Cauchy 公式,

$$\det \left(A \begin{bmatrix} i_1 & i_2 & \cdots & i_k \\ j_1 & j_2 & \cdots & j_k \end{bmatrix} \right) = \sum_{1 \leq t_1 < t_2 < \cdots < t_k \leq r} d_{t_1} d_{t_2} \cdots d_{t_k} \det \left(P^{-1} \begin{bmatrix} i_1 & i_2 & \cdots & i_k \\ t_1 & t_2 & \cdots & t_k \end{bmatrix} \right) \det \left(Q^{-1} \begin{bmatrix} t_1 & t_2 & \cdots & t_k \\ j_1 & j_2 & \cdots & j_k \end{bmatrix} \right)$$

可被 $d_1 d_2 \cdots d_k$ 整除. 综上, $D_k = d_1 d_2 \cdots d_k$. □

定义 1.6.10 根据定理 1.6.9, $d_k = \frac{D_k}{D_{k-1}}$ 由 A 唯一确定, 称为 A 的第 k 个不变因子, $1 \leq k \leq r$. $m \times n$ 矩阵 $\text{diag}(d_1, d_2, \dots, d_r, O)$ 称为 A 的模相抵标准形或 Smith 标准形.

注 1.6.11 模方阵在任意环上都可以定义, 而 Smith 标准形一般需要在主理想整环上定义.

1.7 矩阵的相似

例 1.7.1 设 $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{F}^{2 \times 2}$, $P = \begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix} \in \mathbb{F}^{2 \times 2}$ 是初等方阵, 则

$$B = (b_{ij}) := P^{-1} A P = \begin{pmatrix} a - \lambda c & b + \lambda(a - d) - \lambda^2 c \\ c & d + \lambda c \end{pmatrix}.$$

注意到: ① B 与 A 的主元之和相等. ② 若 $c \neq 0$, 总能找到 λ , 使 $b_{11} = 0$ 或 $b_{22} = 0$. ③ 若 $\mathbb{F} = \mathbb{C}$, 总能找到 λ , 使 $b_{12} = 0$.

例 1.7.2 举出符合以下要求的同阶方阵 A 与 B :

① $\operatorname{tr}(A) = \operatorname{tr}(B)$ 但 A 与 B 不相似: $A = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$

② $\det(A) = \det(B)$ 但 A 与 B 不相似: 同上.

③ $\operatorname{rank}(A) = \operatorname{rank}(B)$ 但 A 与 B 不相似: $A = \begin{pmatrix} 1 \end{pmatrix}, B = \begin{pmatrix} 2 \end{pmatrix}.$

例 1.7.3 若 A_i 与 B_i 相似 ($i = 1, 2$), 则 $A_1 \otimes A_2$ 与 $B_1 \otimes B_2$ 相似.

证明 设 $A_i = P_i B_i P_i^{-1}$, $P_i \in \mathbb{F}^{n \times n}$ 是可逆方阵 ($i = 1, 2$), 则

$$\begin{aligned} (P_1 \otimes P_2)(B_1 \otimes B_2)(P_1 \otimes P_2)^{-1} &= (P_1 \otimes P_2)(B_1 \otimes B_2)(P_1^{-1} \otimes P_2^{-1}) \\ &= (P_1 B_1 P_1^{-1}) \otimes (P_2 B_2 P_2^{-1}) = A_1 \otimes A_2. \end{aligned}$$

由习题 3.3.8 可知 $\det(P_1 \otimes P_2) = (\det(P_1))^n (\det(P_2))^n \neq 0$. □

例 1.7.4 设 $A = (a_{ij}) \in \mathbb{R}^{n \times n}$, 其中 $a_{ij} = i + j$. 求 A 的所有特征值.

解 先作分解

$$A = \begin{pmatrix} 1 & 1 \\ 2 & 1 \\ \vdots & \vdots \\ n & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & 2 & \cdots & n \end{pmatrix}.$$

则由例 (3.12) 知

$$\begin{aligned} \varphi_A(x) &= \det(xI_n - A) = x^{n-2} \det \left(xI_2 - \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & 2 & \cdots & n \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 2 & 1 \\ \vdots & \vdots \\ n & 1 \end{pmatrix} \right) \\ &= x^{n-2} \det \begin{pmatrix} x - \frac{n(n+1)}{2} & -n \\ -\frac{n(n+1)(2n+1)}{6} & x - \frac{n(n+1)}{2} \end{pmatrix} \\ &= x^{n-2} \left[\left(x - \frac{n(n+1)}{2} \right)^2 - \frac{n^2(n+1)(2n+1)}{6} \right]. \end{aligned}$$

所以 A 的所有特征值为 $n-2$ 个 0 加上 $\frac{n(n+1)}{2} \pm n\sqrt{\frac{(n+1)(2n+1)}{6}}$. □

定理 1.7.5 设 $A \in \mathbb{F}^{n \times n}$ 的所有特征值都属于 \mathbb{F} . 则 A 可以在 \mathbb{F} 上相似于对角方阵 \iff 每个特征值的几何重数等于代数重数.

证明 \Rightarrow : 设 $\varphi_A(x) = \prod_{i=1}^k (x - \lambda_i)^{n_i}$, $\lambda_1, \dots, \lambda_k$ 两两不同, 则 $A = P^{-1} \begin{pmatrix} \lambda_1 I_{n_1} & & \\ & \ddots & \\ & & \lambda_k I_{n_k} \end{pmatrix} P$,

其中 P 是可逆方阵. 根据定义, λ_i 的几何重数 (即 λ_i 对应的特征子空间的维数)

$$\begin{aligned} m_i &= n - \text{rank}(\lambda_i I - A) = n - \text{rank} \left(P^{-1} \left(\lambda_i I - \begin{pmatrix} \lambda_1 I_{n_1} & & \\ & \ddots & \\ & & \lambda_k I_{n_k} \end{pmatrix} \right) P \right) \\ &= n - \text{rank} \left(\lambda_i I - \begin{pmatrix} \lambda_1 I_{n_1} & & \\ & \ddots & \\ & & \lambda_k I_{n_k} \end{pmatrix} \right) = n_i. \end{aligned}$$

\Leftarrow : 设 $\varphi_A(x) = \prod_{i=1}^k (x - \lambda_i)^{n_i}$, 其中 $\lambda_1, \lambda_2, \dots, \lambda_k$ 两两不同. 设 $\{\alpha_j \mid s_{i-1} + 1 \leq j \leq s_i\}$ 是线性方程组 $(\lambda_i I - A)\mathbf{x} = \mathbf{0}$ 的一个基础解系, $s_0 = 0$, $s_i = n_1 + \dots + n_i$, $1 \leq i \leq k$. 易知 $P = (\alpha_1 \ \alpha_2 \ \dots \ \alpha_n)$ 满足 $AP = P \text{diag}(\lambda_1 I_{n_1}, \lambda_2 I_{n_2}, \dots, \lambda_k I_{n_k})$. 假设 $P\mathbf{x} = \mathbf{0}$, 即

$$A^T \sum_{i=1}^k \sum_{j=s_{i-1}+1}^{s_i} x_j \alpha_j = \sum_{i=1}^k \lambda_i^T \sum_{j=s_{i-1}+1}^{s_i} x_j \alpha_j = \sum_{i=1}^k \lambda_i^T \beta_i = \mathbf{0}, \quad \forall t.$$

其中 $\beta_i = \sum_{j=s_{i-1}+1}^{s_i} x_j \alpha_j$. 则

$$(\beta_1 \ \dots \ \beta_k) \begin{pmatrix} 1 & \lambda_1 & \dots & \lambda_1^{k-1} \\ 1 & \lambda_2 & \dots & \lambda_2^{k-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_k & \dots & \lambda_k^{k-1} \end{pmatrix} = \mathbf{0}.$$

由于 $\lambda_1, \lambda_2, \dots, \lambda_k$ 两两不同, k 阶 Vandermonde 方阵 (λ_i^{j-1}) 可逆. 故 $\beta_i = \sum_{j=s_{i-1}+1}^{s_i} x_j \alpha_j = \mathbf{0}$, $\forall i$. 故 $\mathbf{x} = \mathbf{0}$. 因此, P 是可逆方阵, A 与 $\text{diag}(\lambda_1 I_{n_1}, \lambda_2 I_{n_2}, \dots, \lambda_k I_{n_k})$ 相似. \square

例 1.7.6 任给 n 次首一多项式 $f \in \mathbb{F}[x]$, 求 $A \in \mathbb{F}^{n \times n}$, 使得 $\varphi_A(x) = f(x)$.

解 记 $f(x) = a_0 + a_1x + \cdots + a_{n-1}x^{n-1} + x^n$, 只需找 f 的友方阵

$$A = \begin{pmatrix} 0 & & & -a_0 \\ 1 & \ddots & & -a_1 \\ & \ddots & 0 & \vdots \\ & & 1 & -a_{n-1} \end{pmatrix}.$$

□

定理 1.7.7 对于任意 $A \in \mathbb{F}^{n \times n}$, 存在 $\alpha \in \mathbb{F}^{n \times 1}$ 使得 $d_{A,\alpha} = d_A$.

证明 设 $d_A = \prod_{i=1}^k p_i^{n_i}$, 其中 $p_1, \dots, p_k \in \mathbb{F}[x]$ 是两两互素的首一不可约多项式, $\deg(p_i)$ 和 n_i 都是正整数. 设 $q_i = \frac{d_A}{p_i}$. 由于 $q_i(A) \neq O$, 故存在 $\alpha_i \in \mathbb{F}^{n \times 1}$ 使得 $q_i(A)\alpha_i \neq \mathbf{0}$. 设 $f_i = \frac{d_A}{p_i^{n_i}}$, $\beta_i = f_i(A)\alpha_i$. 由 $p_i^{n_i}(A)\beta_i = d_A(A)\alpha_i = \mathbf{0}$ 和 $p_i^{n_i-1}(A)\beta_i = q_i(A)\alpha_i \neq \mathbf{0}$ 得 $d_{A,\beta_i} = p_i^{n_i}$. 设 $\alpha = \sum_{i=1}^k \beta_i$. 若 $d_{A,\alpha} \neq d_A$, 则由 $p_i (i = 1, \dots, k)$ 的不可约性知 $d_{A,\alpha}$ 整除某个 q_i , 而对于 $j \neq i$, $q_i(A)\beta_j = \mathbf{0}$, 因而 $q_i(A)\beta_i = q_i(A)\alpha = \mathbf{0}$, 但这与 $d_{A,\beta_i} = p_i^{n_i} \nmid q_i$ 矛盾. 故 $d_{A,\alpha} = d_A$. □

定理 1.7.8 设 $A \in \mathbb{F}^{n \times n}$ 满足 $d_A = \varphi_A$, $f \in \mathbb{F}[x]$, $g = \gcd(d_A, f)$, 则 $\text{rank}(f(A)) = n - \deg(g)$.

证明 根据定理 5.15, 存在 $\alpha \in \mathbb{F}^{n \times 1}$ 满足 $d_{A,\alpha} = d_A = \varphi_A$, 于是 $\alpha, A\alpha, \dots, A^{n-1}\alpha$ 是 $\mathbb{F}^{n \times 1}$ 的一组基, 进而 $f(A)$ 的零空间中的每个元素可表示为 $x = r(A)\alpha$, 其中 $r \in \mathbb{F}[x]$ 且 $\deg(r) \leq n-1$. 下面求 r 需满足的性质.

$$f(A)x = \mathbf{0} \iff f(A)r(A)\alpha = \mathbf{0} \iff d_{A,\alpha} \mid fr \iff d_A \mid fr \iff \frac{d_A}{g} \mid r.$$

记 $h = \frac{d_A}{g}$ (注意 h 是由已知条件确定的), 则

$$\{x \mid f(A)x = \mathbf{0}\} = \{r(A)x \mid r \text{ 是 } h \text{ 的倍式且 } \deg(r) \leq n-1\}.$$

而

$$\{r(A) \mid r \text{ 是 } h \text{ 的倍式且 } \deg(r) \leq n-1\}$$

是由

$$h(x), xh(x), \dots, x^{n-\deg(h)-1}h(x)$$

生成的 $n - \deg(h)$ 维线性空间, 于是 $\text{rank}(f(A)) = n - (n - \deg(h)) = n - \deg(g)$. □

1.8 根子空间

本节的主线是对 \mathbb{F} -线性空间 V 上线性变换 \mathcal{A} 的化零多项式要求的减弱:

$$\varphi_{\mathcal{A}} \xrightarrow{\varphi_{\mathcal{A}} \text{可能不存在}} d_{\mathcal{A}} \xrightarrow{d_{\mathcal{A}} \text{可能不存在}} d_{\mathcal{A},\alpha} \xrightarrow{d_{\mathcal{A},\alpha} \text{可能不存在}} \dots$$

在引入三大根子空间分解定理前, 先熟悉各种例子.

例 1.8.1 (\mathcal{A} 可能没有特征值) 设 $V = \mathbb{F}[x]$, $\mathcal{A} \in \mathcal{L}(V) : f(x) \mapsto xf(x)$. 对于任意非零多项式 $f(x) \in V$, 不存在常数 $\lambda \in \mathbb{F}$, 使得 $xf(x) = \lambda f(x)$. 因此, \mathcal{A} 没有特征值.

注 1.8.2 不同于矩阵的特征多项式一定可以在 \mathbb{F} 的某个扩域 \mathbb{K} 上找到根, 即使 V 是有限维线性空间, 线性变换 \mathcal{A} 的特征值也要求只能在 \mathbb{F} 中 (因为线性空间 V 是在数域 \mathbb{F} 上的), 即若 $\varphi_{\mathcal{A}}(x)$ 在 \mathbb{F} 中没有根, 则 \mathcal{A} 没有特征值和特征向量. 另外, $\varphi_{\mathcal{A}}(x)$ 仅当 V 是有限维线性空间才有定义.

例 1.8.3 (\mathcal{A} 可能有无穷多个特征值) 设 \mathbb{R} 上的线性空间 $V = \mathcal{C}^\infty(\mathbb{R})$. 对于任意 $\lambda \in \mathbb{R}$, $f'(x) = \lambda f(x) \iff f(x) = ae^{\lambda x}$, 其中 $a \in \mathbb{R}$. 因此, 所有实数 λ 都是微分变换 \mathcal{D} 的特征值, 对应的特征向量 $f(x) = ae^{\lambda x}$, $a \neq 0$.

例 1.8.4 ($d_{\mathcal{A}}(x)$ 可能不存在) \mathbb{R} 上的微分变换 \mathcal{D} 有特征值, 但没有最小多项式.

例 1.8.5 ($d_{\mathcal{A}}(x)$ 可能不存在) \mathbb{R} 上的积分变换 \mathcal{S} 既没有特征值, 也没有最小多项式.

例 1.8.6 ($d_{\mathcal{A}}(x)$ 不存在但 $d_{\mathcal{A},\alpha}$ 存在) 设 $V = \mathbb{F}[x]$, $\mathcal{D} \in \mathcal{L}(V)$ 是微分映射. 则对任意 $f \in V$,

$$d_{\mathcal{D},f}(x) = \begin{cases} x^{\deg(f)+1}, & f \neq 0, \\ 1, & f = 0. \end{cases}$$

例 1.8.7 ($d_{\mathcal{A},\alpha}(x)$ 可能不存在) 设 $\text{char } \mathbb{F} = 0$, 则 $\mathbb{F}[x]$ 上的积分变换 \mathcal{S} 关于任意非零向量 α 都没有最小多项式.

例 1.8.8 ($d_{\mathcal{A},\alpha}(x)$ 可能不存在) 设 $V = \mathbb{F}[x]$, $\mathcal{A} \in \mathcal{L}(V) : f(x) \mapsto xf(x)$. 则

$$a_0\mathcal{I} + a_1\mathcal{A} + a_2\mathcal{A}^2 + \dots + a_k\mathcal{A}^k : f(x) \mapsto (a_0 + a_1x + a_2x^2 + \dots + a_kx^k)f(x).$$

由此可知对任意 $0 \neq f \in V$, $d_{\mathcal{A},f}$ 不存在.

$d_{\mathcal{A}}(x)$ 包含 $\varphi_{\mathcal{A}}(x)$ 的所有根, 而 $d_{\mathcal{A},\alpha}(x)$ 在 \mathbb{F} 中的根一定是 \mathcal{A} 的特征值.

定理 1.8.9 设 V 是 \mathbb{F} 上的有限维线性空间, $\mathcal{A} \in \mathcal{L}(V)$. 若 $\varphi_{\mathcal{A}} = \prod_{i=1}^k f_i$, 其中 $f_1, \dots, f_k \in \mathbb{F}[x]$ 两两互素, 则 $V = \bigoplus_{i=1}^k W_i$, 其中 $W_i = \text{Ker } f_i(\mathcal{A})$, $\dim W_i = \deg(f_i)$. 特别地, 若某个 $f_i = (x - \lambda_i)^{n_i}$, $n_i \geq 1$, 则 $W_i = \text{Ker}(\lambda_i \mathcal{I} - \mathcal{A})^{n_i}$ 是 λ_i 对应的根子空间.

定理 1.8.10 设 V 是 \mathbb{F} 上的线性空间, $\mathcal{A} \in \mathcal{L}(V)$ 有最小多项式 $d_{\mathcal{A}}$. 若 $d_{\mathcal{A}} = \prod_{i=1}^k f_i$, 其中 $f_1, \dots, f_k \in \mathbb{F}[x]$ 两两互素, 则 $V = \bigoplus_{i=1}^k \text{Ker } f_i(\mathcal{A})$. 特别地, 若某个 $f_i = (x - \lambda_i)^{n_i}$, $n_i \geq 1$, 则 λ_i 是 \mathcal{A} 的特征值, $\text{Ker}(\lambda_i \mathcal{I} - \mathcal{A})^{n_i}$ 是 λ_i 对应的根子空间.

定理 1.8.11 (根子空间分解) 设 V 是 \mathbb{F} 上的线性空间, $\mathcal{A} \in \mathcal{L}(V)$, S 是 \mathcal{A} 的所有特征值的集合, W_{λ} 是 λ 对应的根子空间. 若对于任意 $\alpha \in V$, $d_{\mathcal{A}, \alpha}$ 存在并且可以在 $\mathbb{F}[x]$ 中分解为一次因式的乘积, 则 $V = \bigoplus_{\lambda \in S} W_{\lambda}$.

注 1.8.12 注意定理 1.8.9 只适用于有限维线性空间, 而定理 1.8.10 和定理 1.8.11 对空间维数不作限制. 另外, 定理 1.8.11 的逆命题也成立 (参见习题 9.6.9).

第二章 讲义习题选做

习题 1.2.5 求所有满足 $f(1) = f(2) = f(3) = 1$, $f(-1) = f(-2) = f(-3) = -1$ 的多项式 $f(x)$.

解 若 f_1, f_2 均为满足题意的解, 则 $h := f_1 - f_2$ 满足 $h(\pm 1) = h(\pm 2) = h(\pm 3) = 0$, 于是若已得到一个特解 $f_0(x)$, 则满足题意的所有多项式为

$$f(x) = f_0(x) + g(x)(x+3)(x+2)(x+1)(x-1)(x-2)(x-3),$$

其中 $g(x)$ 为任意多项式. 在求特解时, 除了常规的待定系数法外, 还可以设

$$f_0(x) = 1 + (x-1)(x-2)(x-3)(c_0 + c_1x + c_2x^2).$$

当然, 由于本题的特殊性, 也可以直接设 $f_0(x) = c_1x + c_3x^3 + c_5x^5$ 为奇函数, 用此方法可解出一个特解

$$f_0(x) = \frac{1}{60}x^5 - \frac{1}{4}x^3 + \frac{37}{30}x.$$

□

习题 1.2.6 求所有满足 $f(1) = f'(1) = f''(1) = 1$, $f(-1) = f'(-1) = f''(-1) = -1$ 的多项式 $f(x)$.

解 若 f_1, f_2 均为满足题意的解, 则 $h := f_1 - f_2$ 满足 $h(1) = h'(1) = h''(1) = h(-1) = h'(-1) = h''(-1) = 0$. 由多项式知识可知, ± 1 均是 $h(x)$ 的重根, 且重数至少为 3. 于是若已得到一个特解 $f_0(x)$, 则满足题意的所有多项式为

$$f(x) = f_0(x) + g(x)(x-1)^3(x+1)^3,$$

其中 $g(x)$ 为任意多项式. 由待定系数法可解出一个特解

$$f_0(x) = \frac{1}{2}x^5 - \frac{1}{8}x^4 - \frac{3}{2}x^3 + \frac{3}{4}x^2 + 2x - \frac{5}{8}.$$

□

注 2.0.1 本题的背景是结点处具有二阶导数信息的 **Hermite 插值**.

习题 2.1.8 (8) 设 $m \in \mathbb{N}^*$, 求 A^m , 其中

$$A = \begin{pmatrix} a_1b_1 & a_1b_2 & a_1b_3 & a_1b_4 \\ a_2b_1 & a_2b_2 & a_2b_3 & a_2b_4 \\ a_3b_1 & a_3b_2 & a_3b_3 & a_3b_4 \\ a_4b_1 & a_4b_2 & a_4b_3 & a_4b_4 \end{pmatrix}.$$

解 因为 $A = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{pmatrix} \begin{pmatrix} b_1 & b_2 & b_3 & b_4 \end{pmatrix}$, 而 $\begin{pmatrix} b_1 & b_2 & b_3 & b_4 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{pmatrix} = a_1b_1 + a_2b_2 + a_3b_3 + a_4b_4$,

所以

$$\begin{aligned} A^m &= \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{pmatrix} (a_1b_1 + a_2b_2 + a_3b_3 + a_4b_4)^{m-1} \begin{pmatrix} b_1 & b_2 & b_3 & b_4 \end{pmatrix} \\ &= (a_1b_1 + a_2b_2 + a_3b_3 + a_4b_4)^{m-1} A. \end{aligned}$$

□

习题 2.1.16 设多项式列 $Q_0 = 1, Q_1 = x, Q_n = \frac{Q_{n-1}^2 - 1}{Q_{n-2}}, \forall n \geq 2$. 用矩阵乘积表示 Q_n 的通项公式, 并证明 $Q_n \in \mathbb{Z}[x]$.

解 由递推关系, $1 = Q_{n-1}^2 - Q_n Q_{n-2} = Q_n^2 - Q_{n+1} Q_{n-1}$, 移项整理得 $\frac{Q_{n+1} + Q_{n-1}}{Q_n} = \frac{Q_n + Q_{n-2}}{Q_{n-1}} = \frac{Q_2 + Q_0}{Q_1} = x$. 于是 $Q_n = xQ_{n-1} - Q_{n-2}$ (因此 $Q_n \in \mathbb{Z}[x]$), 用矩阵乘积表示即

$$Q_n = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} Q_n \\ Q_{n-1} \end{pmatrix} = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} x & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} Q_{n-1} \\ Q_{n-2} \end{pmatrix} = \cdots = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} x & -1 \\ 1 & 0 \end{pmatrix}^{n-1} \begin{pmatrix} x \\ 1 \end{pmatrix}.$$

□

习题 2.1.16 (2) $x_n = \frac{ax_{n-1} + b}{cx_{n-1} + d}$, 其中 a, b, c, d 是常数. 利用矩阵乘积求递推数列 $\{x_n\}_{n \in \mathbb{N}}$ 的通项公式.

解 令 $x_n = \frac{y_n}{z_n}$, 则 $\frac{y_n}{z_n} = \frac{ay_{n-1} + bz_{n-1}}{cy_{n-1} + dz_{n-1}}$. 不妨设 $y_n = ay_{n-1} + bz_{n-1}$, $z_n = cy_{n-1} + dz_{n-1}$, 则

$$\begin{pmatrix} y_n \\ z_n \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} y_{n-1} \\ z_{n-1} \end{pmatrix} = \cdots = \begin{pmatrix} a & b \\ c & d \end{pmatrix}^n \begin{pmatrix} y_0 \\ z_0 \end{pmatrix}.$$

若记 $\begin{pmatrix} a & b \\ c & d \end{pmatrix}^n = \begin{pmatrix} a_n & b_n \\ c_n & d_n \end{pmatrix}$, 则 $x_n = \frac{a_n x_0 + b_n}{c_n x_0 + d_n}$. □

习题 2.1.17 设 n 是正整数, a_1, a_2, \dots, a_{n-1} 是任意实数, 定义递推数列 $u_0 = u_1 = v_0 = v_1 = 1$,

$$u_{k+1} = u_k + a_k u_{k-1}, \quad v_{k+1} = v_k + a_{n-k} v_{k-1}, \quad \forall k = 1, \dots, n-1.$$

证明: $u_n = v_n$.

证明 类似例 2.7, 有

$$\begin{aligned} u_n &= \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & a_{n-1} \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & a_{n-2} \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} 1 & a_1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \\ v_n &= \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & a_1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & a_2 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} 1 & a_{n-1} \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ a_{n-1} & 0 \end{pmatrix} \cdots \begin{pmatrix} 1 & 1 \\ a_2 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ a_1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \end{aligned}$$

利用

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

即可得 $u_n = v_n$. □

习题 2.1.18 设 a_0, a_1, a_2, \dots 是整数, b_0, b_1, b_2, \dots 是正整数, 满足 $a_0 = 0, a_1 = 1$,

$$a_{n+1} = \begin{cases} a_n b_n + a_{n-1}, & b_{n-1} = 1 \\ a_n b_n - a_{n-1}, & b_{n-1} > 1 \end{cases}, \quad n = 1, 2, \dots.$$

证明: $\max\{a_n, a_{n+1}\} \geq n$.

证明 设

$$c_n = \begin{cases} 1, & b_n = 1, \\ -1, & b_n > 1, \end{cases}$$

则递推关系可表示为

$$\begin{pmatrix} a_{n+1} \\ a_n \end{pmatrix} = \begin{pmatrix} b_n & c_{n-1} \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_n \\ a_{n-1} \end{pmatrix}.$$

利用

$$\begin{pmatrix} b_k & c_{k-1} \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} b_k & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & c_{k-1} \end{pmatrix} \quad \text{与} \quad \begin{pmatrix} 1 & 0 \\ 0 & c_{k-1} \end{pmatrix} \begin{pmatrix} b_{k-1} & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} b_{k-1} & 1 \\ c_{k-1} & 0 \end{pmatrix}$$

可得

$$\begin{aligned} \begin{pmatrix} a_{n+1} \\ a_n \end{pmatrix} &= \begin{pmatrix} b_n & c_{n-1} \\ 1 & 0 \end{pmatrix} \begin{pmatrix} b_{n-1} & c_{n-2} \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} b_1 & c_0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} b_n & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} b_{n-1} & 1 \\ c_{n-1} & 0 \end{pmatrix} \begin{pmatrix} b_{n-2} & 1 \\ c_{n-2} & 0 \end{pmatrix} \cdots \begin{pmatrix} b_1 & 1 \\ c_1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \end{aligned}$$

设

$$\begin{pmatrix} u_n \\ v_n \end{pmatrix} = \begin{pmatrix} b_{n-1} & 1 \\ c_{n-1} & 0 \end{pmatrix} \begin{pmatrix} b_{n-2} & 1 \\ c_{n-2} & 0 \end{pmatrix} \cdots \begin{pmatrix} b_1 & 1 \\ c_1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

下面用归纳法证明当 $n \geq 2$ 时有 $\begin{cases} u_n \geq 1, \\ u_n + v_n \geq n \end{cases}$ (情形 1) 或 $\begin{cases} u_n \geq n, \\ u_n + v_n \geq 1 \end{cases}$ (情形 2):

① 当 $n = 2$ 时 $\begin{pmatrix} u_2 \\ v_2 \end{pmatrix} = \begin{pmatrix} b_1 \\ c_1 \end{pmatrix}$. 若 $b_1 = 1$, 则 $c_1 = 1$, 从而 $u_2 = v_2 = 1$, 满足情形 1; 若 $b_1 > 1$ (即 $b_1 \geq 2$), 则 $c_1 = -1$, 从而 $u_2 \geq 2$, $u_2 + v_2 = u_2 - 1 \geq 1$, 满足情形 2. 因此当 $n = 2$ 时命题成立.

② 假设命题对任意小于 n ($n \geq 3$) 的正整数成立, 且注意到 $\begin{pmatrix} u_n \\ v_n \end{pmatrix} = \begin{pmatrix} b_{n-1}u_{n-1} + v_{n-1} \\ c_{n-1}u_{n-1} \end{pmatrix}$.

• 若 $\begin{pmatrix} u_{n-1} \\ v_{n-1} \end{pmatrix}$ 对应情形 1: 若 $b_{n-1} = 1$, 则 $c_{n-1} = 1$, 从而 $u_n = u_{n-1} + v_{n-1} \geq n - 1 \geq 1$, $u_n + v_n = u_{n-1} + (u_{n-1} + v_{n-1}) \geq 1 + (n - 1) = n$, 满足情形 1; 若 $b_{n-1} > 1$, 则 $c_{n-1} = -1$, 从而 $u_n = b_{n-1}u_{n-1} + v_{n-1} \geq 2u_{n-1} + v_{n-1} \geq 1 + (n - 1) = n$, $u_n + v_n = (b_{n-1} - 1)u_{n-1} + v_{n-1} \geq u_{n-1} + v_{n-1} \geq n - 1 \geq 1$, 满足情形 2.

• 若 $\begin{pmatrix} u_{n-1} \\ v_{n-1} \end{pmatrix}$ 对应情形 2: 若 $b_{n-1} = 1$, 则 $c_{n-1} = 1$, 从而 $u_n = u_{n-1} + v_{n-1} \geq 1$, $u_n + v_n = u_{n-1} + (u_{n-1} + v_{n-1}) \geq (n - 1) + 1 = n$, 满足情形 1; 若 $b_{n-1} > 1$, 则 $c_{n-1} = -1$, 从而 $u_n = b_{n-1}u_{n-1} + v_{n-1} \geq 2u_{n-1} + v_{n-1} \geq n$, $u_n + v_n = (b_{n-1} - 1)u_{n-1} + v_{n-1} \geq u_{n-1} + v_{n-1} \geq 1$, 满足情形 2.

由 ① ② 可知命题对 $n \geq 2$ 成立. 于是由

$$\begin{pmatrix} a_{n+1} \\ a_n \end{pmatrix} = \begin{pmatrix} b_n & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} u_n \\ v_n \end{pmatrix} = \begin{pmatrix} b_n u_n + v_n \\ u_n \end{pmatrix}$$

可知, 若 $a_n = u_n < n$, 则必有 $a_{n+1} = b_n u_n + v_n \geq u_n + v_n \geq n$, 也即 $\max\{a_n, a_{n+1}\} \geq n$.

□

习题 2.2.3 设方阵 $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$, $B = \begin{pmatrix} O & A & O \\ O & O & A \\ A & O & O \end{pmatrix}$. 求所有与 B 乘积可交换的方阵.

解 设与 B 乘积可交换的方阵为 $\begin{pmatrix} X_{11} & X_{12} & X_{13} \\ X_{21} & X_{22} & X_{23} \\ X_{31} & X_{32} & X_{33} \end{pmatrix}$, 其中每个 X_{ij} 均为 3 阶方阵. 则有

$$\begin{pmatrix} X_{13}A & X_{11}A & X_{12}A \\ X_{23}A & X_{21}A & X_{22}A \\ X_{33}A & X_{31}A & X_{32}A \end{pmatrix} = \begin{pmatrix} AX_{21} & AX_{22} & AX_{23} \\ AX_{31} & AX_{32} & AX_{33} \\ AX_{11} & AX_{12} & AX_{13} \end{pmatrix}.$$

若记 $X := X_{11}$, $Y := X_{12}$, $Z := X_{13}$, 则每个 X_{ij} 都可由 A, A^{-1}, X, Y, Z 表示. 计算可得结果

$$\begin{pmatrix} X & Y & Z \\ A^{-1}ZA & A^{-1}XA & A^{-1}YA \\ AYA^{-1} & AZA^{-1} & AXA^{-1} \end{pmatrix},$$

其中 X, Y, Z 为任意 3 阶方阵.

□

习题 2.2.4 设 $m \times n$ 网格图 G 的顶点集 $V = \{(x, y) \mid x = 1, 2, \dots, m; y = 1, 2, \dots, n\}$, 两个顶点 (x_1, y_1) 与 (x_2, y_2) 相邻当且仅当 $|x_1 - x_2| + |y_1 - y_2| = 1$. 把 V 按照字典顺序排列. 用分块矩阵表示 G 的邻接矩阵.

解 题目描述的网格图 G 即 $\text{Grid}[\{m, n\}]$, 其对应的邻接矩阵 $A = I_m \otimes P_n + P_m \otimes I_n$, 其中

$$P_k = \begin{pmatrix} 0 & 1 & & & \\ 1 & 0 & 1 & & \\ & 1 & 0 & \ddots & \\ & & \ddots & \ddots & 1 \\ & & & 1 & 0 \end{pmatrix}_{k \times k}.$$

□

习题 2.2.5 定义两个图 G_1, G_2 的直积 $G = G_1 \times G_2$ 如下: G 的顶点集 $V = \{(v_1, v_2) \mid v_i \text{ 是 } G_i \text{ 的顶点}\}$, $(u_1, u_2) \rightarrow (v_1, v_2)$ 是 G 的边当且仅当 $(u_1 = v_1 \text{ 且 } u_2 \rightarrow v_2 \text{ 是 } G_2 \text{ 的边})$ 或者 $(u_2 = v_2 \text{ 且 } u_1 \rightarrow v_1 \text{ 是 } G_1 \text{ 的边})$. 用 G_1, G_2 的邻接矩阵表示 G 的邻接矩阵.

解 分别记 G_1, G_2, G 的邻接矩阵为 $\tilde{G}_1, \tilde{G}_2, \tilde{G}$. 设 G_1, G_2 分别有 m, n 个顶点, 并按

$$(1, 1), \cdots, (1, n), (2, 1), \cdots, (2, n), \cdots, (m, 1), \cdots, (m, n)$$

的顺序标记 G 的顶点. 将 \tilde{G} 的邻接矩阵按顶点的第一个分量划分成 $m \times m$ 分块, 分析可知, 按 “ $u_1 = v_1$ 且 $u_2 \rightarrow v_2$ 是 G_2 的边” 给出的直积的邻接矩阵为 $I_m \otimes \tilde{G}_2$, 按 “ $u_2 = v_2$ 且 $u_1 \rightarrow v_1$ 是 G_1 的边” 给出的直积的邻接矩阵为 $\tilde{G}_1 \otimes I_n$. 于是通过 “按矩阵元素取或” 可得

$$\tilde{G} = f(I_m \otimes \tilde{G}_2 + \tilde{G}_1 \otimes I_n),$$

其中

$$f: \{0, 1, 2\}^{mn \times mn} \rightarrow \{0, 1\}^{mn \times mn}, \quad (f(M))_{ij} = \begin{cases} 1, & M_{ij} = 2, \\ M_{ij}, & \text{其他}. \end{cases}$$

□

习题 2.2.6 设 n 维超立方体图 Q_n 的顶点集 $V = \{0, 1\}^n$, 两个顶点 (x_1, x_2, \cdots, x_n) 与 (y_1, y_2, \cdots, y_n) 相邻当且仅当 $|x_1 - y_1| + |x_2 - y_2| + \cdots + |x_n - y_n| = 1$. 把 V 按照字典顺序排列. 思考 Q_n 与 Q_{n-1} 的关系, 并用分块矩阵递推表示 Q_n 的邻接矩阵.

解 用 \tilde{Q}_i 表示 Q_i 的邻接矩阵. 则 $\tilde{Q}_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, 且当 $n \geq 2$ 时, $\tilde{Q}_n = \begin{pmatrix} \tilde{Q}_{n-1} & I_{2^{n-1}} \\ I_{2^{n-1}} & \tilde{Q}_{n-1} \end{pmatrix}$. □

习题 2.2.7 设矩阵乘积 AC, BD 有意义. 证明: $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$.

证明 记 $A = (a_{ij}), C = (c_{ij})$, 那么 $A \otimes B = (a_{ij}B), C \otimes D = (c_{ij}D)$. $(A \otimes B)(C \otimes D)$ 在 (i, j) 位置处的分块是

$$\sum_k (a_{ik}B)(c_{kj}D) = \left(\sum_k a_{ik}c_{kj} \right) BD = (AC)_{ij}BD,$$

其中 $(AC)_{ij}$ 表示 AC 在 (i, j) 位置处的元素. 这表明 $(A \otimes B)(C \otimes D)$ 在 (i, j) 位置处的分块就是 $(AC) \otimes (BD)$ 在 (i, j) 位置处的分块. □

推论 2.0.2 ① 如果 $A \in \mathbb{F}^{n \times n}$ 与 $B \in \mathbb{F}^{m \times m}$ 可逆, 那么 $A \otimes B$ 可逆, 且 $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$.

② 归纳可得

$$(A_1 A_2 \cdots A_n) \otimes (B_1 B_2 \cdots B_n) = (A_1 \otimes B_1)(A_2 \otimes B_2) \cdots (A_n \otimes B_n).$$

习题 2.3.2 证明: 对于三类初等方阵 P , 均存在列向量 α, β , 使得 $P = I + \alpha\beta^T$.

证明 利用基础矩阵的分解 $E_{ij} = e_i e_j^T$ 可得

$$\begin{aligned} S_{ij} &= I - E_{ii} - E_{jj} + E_{ij} + E_{ji} = I - e_i e_i^T - e_j e_j^T + e_i e_j^T + e_j e_i^T \\ &= I + (e_i - e_j)(e_j - e_i)^T, \\ D_i(\lambda) &= I + (\lambda - 1)E_{ii} = I + (\lambda - 1)e_i e_i^T, \\ T_{ij}(\lambda) &= I + \lambda E_{ij} = I + \lambda e_i e_j^T. \end{aligned}$$

□

习题 2.3.3 若 n 维列向量 α, β 满足 $\alpha^T \beta = 0$, 则 n 阶方阵 $P = I + \alpha\beta^T$ 称为平延. 证明: 对于任意两个不平行的列向量 u, v , 存在平延 P , 使得 $Pu = v$.

证明 $(I + \alpha\beta^T)u = u + \alpha\beta^T u = u + (\beta^T u)\alpha = v \iff (\beta^T u)\alpha = v - u$. 尝试取 $\alpha = v - u$, 只需证明存在 β 使得 $\beta^T u = 1$ 且 $\beta^T \alpha = 0$. 由 u, v 不平行知 u, α 不平行, 从而 β 至少是二维向量. 由此可知所需方程组必有解. □

习题 2.3.5 设 A 是 $m \times n$ 矩阵, i_1, i_2, \dots, i_m 是 $1, 2, \dots, m$ 的排列, j_1, j_2, \dots, j_n 是 $1, 2, \dots, n$ 的排列. 证明: $P = I_m \begin{bmatrix} i_1 & i_2 & \dots & i_m \\ 1 & 2 & \dots & m \end{bmatrix}$ 和 $Q = I_n \begin{bmatrix} 1 & 2 & \dots & n \\ j_1 & j_2 & \dots & j_n \end{bmatrix}$ 都是置换方阵, 使得

$$PAQ = A \begin{bmatrix} i_1 & i_2 & \dots & i_m \\ j_1 & j_2 & \dots & j_n \end{bmatrix}.$$

特别地, 当 $m = n$ 且 $(i_1, i_2, \dots, i_m) = (j_1, j_2, \dots, j_n)$ 时, $Q = P^T$.

证明 因为 P, Q 分别是由单位方阵 I_m, I_n 作行、列重排得到的, 由定义知它们都是置换方阵. 若 $m = n$ 且 $(i_1, i_2, \dots, i_m) = (j_1, j_2, \dots, j_n)$, 记置换

$$\sigma: \{1, 2, \dots, n\} \rightarrow \{1, 2, \dots, n\}, \quad k \mapsto i_k.$$

则 $P_{ij} = I_{\sigma^{-1}(i)j} = \delta_{\sigma^{-1}(i)j} = \delta_{j\sigma^{-1}(i)} = I_{j\sigma^{-1}(i)} = Q_{ji}$. 于是 $Q = P^T$. □

习题 2.3.6 证明: 任意三阶以上方阵 $T_{ij}(\lambda)$ 都可以表示为

$$T_{pq}(a)T_{uv}(b)T_{pq}(-a)T_{uv}(-b)$$

的形式, 其中 $i \neq j, p \neq q, u \neq v$.

证明 设 $\alpha_i, \alpha_j, \alpha_k$ 为行向量, 则以下初等变换满足要求:

$$\begin{pmatrix} \vdots \\ \alpha_i \\ \vdots \\ \alpha_j \\ \vdots \\ \alpha_k \\ \vdots \end{pmatrix} \xrightarrow{k-=j} \begin{pmatrix} \vdots \\ \alpha_i \\ \vdots \\ \alpha_j \\ \vdots \\ \alpha_k - \alpha_j \\ \vdots \end{pmatrix} \xrightarrow{i-=(\lambda k)} \begin{pmatrix} \vdots \\ \alpha_i - \lambda \alpha_k + \lambda \alpha_j \\ \vdots \\ \alpha_j \\ \vdots \\ \alpha_k - \alpha_j \\ \vdots \end{pmatrix} \\
 \xrightarrow{k+=j} \begin{pmatrix} \vdots \\ \alpha_i - \lambda \alpha_k + \lambda \alpha_j \\ \vdots \\ \alpha_j \\ \vdots \\ \alpha_k \\ \vdots \end{pmatrix} \xrightarrow{i+=(\lambda k)} \begin{pmatrix} \vdots \\ \alpha_i + \lambda \alpha_j \\ \vdots \\ \alpha_j \\ \vdots \\ \alpha_k \\ \vdots \end{pmatrix}.$$

此即 $T_{ij}(\lambda) = T_{ik}(\lambda)T_{kj}(1)T_{ik}(-\lambda)T_{kj}(-1)$, 只要使 k 异于 i, j 即可. \square

习题 2.3.7 证明: 对于任意 $m \times n$ 矩阵 A , 存在一系列 m 阶初等方阵 P_1, P_2, \dots, P_s 和 n 阶置换方阵 Q , 使得 $P_s \cdots P_2 P_1 A Q$ 形如 $\begin{pmatrix} I_r & * \\ O & O \end{pmatrix}$, 其中 $0 \leq r \leq \min\{m, n\}$.

证明 易知任意多个相同阶数的初等方阵 S_{ij} 的乘积为置换方阵, 这允许我们在下面过程中在 $A = (a_{ij})$ 的右边 (先后) 乘上不同的 S_{ij} 作为所需的置换方阵 Q . 下面取定任意的 n , 对 m 使用数学归纳法:

① 当 $m = 1$ 时, 若所有元素全为 0 则已符合所要形式; 若有某个 $a_{1j} \neq 0$, 则先左乘 $D_1 \begin{pmatrix} 1 \\ a_{1j} \end{pmatrix}$, 再右乘 S_{1j} 可化为所要形式.

② 假设当 $m = k$ 时欲证命题成立, 当 $m = k + 1$ 时, 由假设可先作变换将前 k 行化为所要形式:

$$\begin{pmatrix} I_{r-1} & * \\ O & O \\ * & * \end{pmatrix}.$$

利用已有的 I_{r-1} , 可以通过左乘 $T_{k+1,1}(-a_{k+1,1})T_{k+1,2}(-a_{k+1,2})\cdots T_{k+1,r-1}(-a_{k+1,r-1})$ 将第 $k+1$ 行的前 $r-1$ 列元素化为 0. 此时若第 $k+1$ 行所有元素全为 0 则已符合所要形式; 若有某个 $a_{k+1,t} \neq 0$ ($t > r-1$), 先右乘 $S_{r,t}$ 将该元素换位至第 r 列, 再左乘 $D_r\left(\frac{1}{a_{k+1,t}}\right)S_{r,k+1}$ 将第 $k+1$ 行与第 r 行换位并将左上角化为 I_r 的下三角部分. 为了将左上角完全化为 I_r , 还需左乘 $T_{1,r}(-a_{1,r})T_{2,r}(-a_{2,r})\cdots T_{r-1,r}(-a_{r-1,r})$. 故 $m=k$ 时命题也成立.

综上 ① ② 可知命题对任意 $m \times n$ 矩阵成立. \square

习题 2.3.8 对角元素都是 1 的上(下)三角矩阵称为单位上(下)三角矩阵. 证明: 对于任意 $m \times n$ 矩阵 A , 存在一系列 m 阶单位下三角的初等方阵 P_1, P_2, \dots, P_s 和 m 阶置换方阵 Q , 使得 $P_s \cdots P_2 P_1 Q A$ 是上三角矩阵.

证明 (为方便起见, 以下选用不同于题目中的矩阵记号.) 按照线性方程组 Gauss 消元法的步骤, 我们可以对 A 进行 LU 分解, 唯一的问题是当主元为 0 时无法继续进行. 于是改用 Gauss 列主元消去法: 对于第 k 个主元, 我们仅在该主元所在列的下方寻找绝对值最大的元素, 通过左乘置换方阵 P_k 将其所在行与第 k 个主元所在行交换, 接着选取若干个初等方阵 T_{ij} ($i > j$) 使得它们的乘积 L_k (注意到当 $i > j$ 时 T_{ij} 为单位下三角方阵, 从而它们的乘积 L_k 仍是单位下三角方阵) 左乘当前矩阵后, 从第 $k+1$ 行开始都可以利用第 k 个主元进行消元. 这样最终就得到

$$L_m P_m L_{m-1} P_{m-1} \cdots L_2 P_2 L_1 P_1 A = U$$

的形式. 下面只需证明任意 $L_r P_r L_s P_s$ ($r > s$) 都可以合并成 LP 的形式. 事实上,

$$L_r P_r L_s P_s = L_r P_r L_s (P_r^{-1} P_r) P_s = L_r (P_r L_s P_r^{-1}) (P_r P_s),$$

而每个 L_k 都形如

$$\begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & * & 1 & \\ & & \vdots & & \ddots \\ & & * & & & 1 \end{pmatrix},$$

其中除去对角线外非零元素 “*” 位于第 k 列的第 $k+1$ 至 m 行. 可见, 当 $r > s$ 时, P_r 仅对 L_s 含 “*” 行作交换, 从而 $P_r L_s P_r^{-1}$ 仍是单位下三角方阵, 记作 \tilde{L}_s . 于是

$$L_r P_r L_s P_s = (L_r \tilde{L}_s) (P_r P_s) =: LP.$$

重复上述过程最终就能得到 $LPA = U$ 的形式. \square

注 2.0.3 本题的背景是 $LU = PA$ 这种矩阵分解形式, 其中 P 指置换方阵, L 指下三角矩阵, U 指上三角矩阵. 选主元的 PLU 分解可参见 <https://zhuanlan.zhihu.com/p/326665113> 以及 <https://zhuanlan.zhihu.com/p/279323822>. 计算实例:

$$\underbrace{\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & 1 \end{pmatrix}}_L \underbrace{\begin{pmatrix} 2 & 3 & 1 & 0 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 3 & 4 \\ 0 & 0 & 0 & \frac{1}{2} \end{pmatrix}}_U = \underbrace{\begin{pmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix}}_P \underbrace{\begin{pmatrix} 0 & 1 & 2 & 3 \\ 1 & 2 & 0 & 0 \\ 2 & 3 & 1 & 0 \\ 0 & 0 & 3 & 4 \end{pmatrix}}_A.$$

习题 2.3.9 (3) 设 \mathbb{R}^3 中以 x, y, z 轴为转轴的三类旋转变换分别对应矩阵

$$P_1(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix}, \quad P_2(\theta) = \begin{pmatrix} \cos \theta & 0 & \sin \theta \\ 0 & 1 & 0 \\ -\sin \theta & 0 & \cos \theta \end{pmatrix}, \quad P_3(\theta) = \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

证明: 对于任意 $A \in \mathbb{R}^{3 \times 3}$, 存在 $\theta_1, \theta_2, \theta_3$, 使得 $P_1(\theta_1)P_2(\theta_2)P_3(\theta_3)A$ 是上三角方阵.

证明 注意到 $P_i(\theta)$ ($i = 1, 2, 3$) 左乘任意 $A \in \mathbb{R}^{3 \times 3}$ 不改变 A 的第 i 行. 记 $A = (a_{ij})$. 下面分别用 s 和 c 简记 $\sin \theta$ 和 $\cos \theta$.

$$\textcircled{1} \text{ 若 } a_{11} \neq 0, \text{ 取 } P_3(\theta_3) = \begin{pmatrix} c & -s & 0 \\ s & c & 0 \\ 0 & 0 & 1 \end{pmatrix}, \text{ 则}$$

$$A^{(1)} := P_3(\theta)A = \begin{pmatrix} ca_{11} - sa_{21} & ca_{12} - sa_{22} & ca_{13} - sa_{23} \\ sa_{11} + ca_{21} & sa_{12} + ca_{22} & sa_{13} + ca_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

取

$$s = -\frac{a_{21}}{\sqrt{a_{11}^2 + a_{21}^2}}, \quad c = \frac{a_{11}}{\sqrt{a_{11}^2 + a_{21}^2}},$$

则 $a_{11}^{(1)} = \sqrt{a_{11}^2 + a_{21}^2} > 0$, $a_{21}^{(1)} = 0$. 同理, 由于 $a_{11}^{(1)} \neq 0$, 可选取 $P_2(\theta_2)$ 左乘 $P_3(\theta_3)A$ 得到 $A^{(2)}$, 使 $a_{11}^{(2)} > 0$, $a_{21}^{(2)} = a_{31}^{(2)} = 0$. 再左乘 $P_1(\theta_1)$ 得到

$$A^{(3)} := P_1(\theta_1)P_2(\theta_2)P_3(\theta_3)A = \begin{pmatrix} a_{11}^{(2)} & a_{12}^{(2)} & a_{13}^{(2)} \\ 0 & ca_{22}^{(2)} - sa_{32}^{(2)} & ca_{23}^{(2)} - sa_{33}^{(2)} \\ 0 & sa_{22}^{(2)} + ca_{32}^{(2)} & sa_{23}^{(2)} + ca_{33}^{(2)} \end{pmatrix}.$$

- 若 $a_{22}^{(2)} \neq 0$, 取

$$s = -\frac{a_{32}^{(2)}}{\sqrt{\left(a_{22}^{(2)}\right)^2 + \left(a_{32}^{(2)}\right)^2}}, \quad c = \frac{a_{22}^{(2)}}{\sqrt{\left(a_{22}^{(2)}\right)^2 + \left(a_{32}^{(2)}\right)^2}},$$

此时 $A^{(3)}$ 即为所要的上三角方阵.

- 若 $a_{22}^{(2)} = 0$, 直接取 $c = 0, s = 1$, 就有 $a_{32}^{(2)} = 0$, 也得到所要的上三角方阵.

② 若 $a_{11} = 0$, 不妨设 A 第 1 列不全为 0 (否则取 $\theta_3 = \theta_2 = 0$ 就变成 ① 中处理第 2 列的情形).

• 若 $a_{21} \neq 0$, 在左乘 $P_3(\theta_3)$ 时, 取 $s = -1, c = 0$, 就使 $a_{11}^{(1)} = a_{21} \neq 0, a_{21}^{(1)} = 0$. 接下来两次左乘化为 ① 中情形.

• 若 $a_{21} = 0, a_{31} \neq 0$, 取 $\theta_3 = 0$, 接着在左乘 $P_2(\theta_2)$ 时, 取 $s = 1, c = 0$, 便完成了对第 1 列的处理, 剩余 1 次左乘化为 ① 中情形.

综上可知对任意 $A \in \mathbb{R}^{3 \times 3}$, 存在 $\theta_1, \theta_2, \theta_3$, 使得 $P_1(\theta_1)P_2(\theta_2)P_3(\theta_3)A$ 是上三角方阵. \square

注 2.0.4 本题的背景是 QR 分解的 Givens 方法.

习题 2.4.12 设 n 阶复数方阵 $A = (a_{ij})$ 满足 $|a_{ii}| > \sum_{j \neq i} |a_{ij}|, \forall i$. 证明: 线性方程组 $Ax = \mathbf{0}$ 只有零解, 从而 A 是可逆方阵.

证明 只需证 A 的行向量线性无关. 用反证法, 假设 A 的行向量 $\alpha_1, \dots, \alpha_n$ 线性相关, 则在 \mathbb{C} 中存在一组不全为 0 的数 k_1, \dots, k_n , 使得

$$k_1 \alpha_1 + \dots + k_n \alpha_n = \mathbf{0}.$$

不妨设

$$|k_l| = \max \{|k_1|, \dots, |k_n|\}.$$

对于第 l 个分量的等式

$$k_1 a_{l1} + \dots + k_l a_{ll} + \dots + k_n a_{ln} = 0,$$

从 $k_l \neq 0$ 可得

$$a_{ll} = -\frac{k_1}{k_l} a_{l1} - \dots - \frac{k_{l-1}}{k_l} a_{l,l-1} - \frac{k_{l+1}}{k_l} a_{l,l+1} - \dots - \frac{k_n}{k_l} a_{ln} = -\sum_{j \neq l} \frac{k_j}{k_l} a_{lj},$$

于是

$$|a_{ll}| = \left| \sum_{j \neq l} \frac{k_j}{k_l} a_{lj} \right| \leq \sum_{j \neq l} \frac{|k_j|}{|k_l|} |a_{lj}| \leq \sum_{j \neq l} |a_{lj}|,$$

而这与已知条件矛盾. \square

习题 2.4.10 (Sherman–Morrison–Woodbury 公式) 设 A 是 m 阶可逆方阵, B 是 $m \times n$ 矩阵, C 是 $n \times m$ 矩阵. 证明: $A + BC$ 是可逆方阵当且仅当 $I + CA^{-1}B$ 是可逆方阵, 并且

$$(A + BC)^{-1} = A^{-1} - A^{-1}B(I + CA^{-1}B)^{-1}CA^{-1}.$$

特别地, $I_m - BC$ 是可逆方阵当且仅当 $I_n - CB$ 是可逆方阵, 并且

$$(I_m - BC)^{-1} = I_m + B(I_n - CB)^{-1}C.$$

证明 考虑方程 $(A + BC)\mathbf{x} = \mathbf{b}$. 记 $\boldsymbol{\xi} := C\mathbf{x}$, 则方程转化为 $(A + BC)\mathbf{x} = A\mathbf{x} + B\boldsymbol{\xi}$, 也即

$$\begin{pmatrix} A & B \\ C & -I_n \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \boldsymbol{\xi} \end{pmatrix} = \begin{pmatrix} \mathbf{b} \\ \mathbf{0} \end{pmatrix}.$$

对左边的矩阵进行分解:

$$\begin{pmatrix} A & B \\ C & -I_n \end{pmatrix} = \begin{pmatrix} I_m & O \\ CA^{-1} & I_n \end{pmatrix} \begin{pmatrix} A & B \\ O & -I_n - CA^{-1}B \end{pmatrix}.$$

由 $\begin{pmatrix} I_m & O \\ CA^{-1} & I_n \end{pmatrix}^{-1} = \begin{pmatrix} I_m & O \\ -CA^{-1} & I_n \end{pmatrix}$ 可得

$$\begin{pmatrix} A & B \\ O & -I_n - CA^{-1}B \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \boldsymbol{\xi} \end{pmatrix} = \begin{pmatrix} I_m & O \\ -CA^{-1} & I_n \end{pmatrix} \begin{pmatrix} \mathbf{b} \\ \mathbf{0} \end{pmatrix} = \begin{pmatrix} \mathbf{b} \\ -CA^{-1}\mathbf{b} \end{pmatrix}.$$

于是

$$\begin{aligned} I_n + CA^{-1}B \text{ 可逆} &\iff \det(-I_n - CA^{-1}B) \neq 0 \xLeftrightarrow{\text{Laplace 展开}} \\ \det \begin{pmatrix} A & B \\ O & -I_n - CA^{-1}B \end{pmatrix} &= \det(A) \cdot \det(-I_n - CA^{-1}B) \neq 0 \\ &\iff \begin{pmatrix} A & B \\ O & -I_n - CA^{-1}B \end{pmatrix} \text{ 可逆} \\ &\iff \begin{pmatrix} A & B \\ O & -I_n - CA^{-1}B \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \boldsymbol{\xi} \end{pmatrix} = \begin{pmatrix} \mathbf{b} \\ -CA^{-1}\mathbf{b} \end{pmatrix} \text{ 有唯一解} \\ &\stackrel{*}{\iff} (A + BC)\mathbf{x} = \mathbf{b} \text{ 有唯一解} \iff A + BC \text{ 可逆}. \end{aligned}$$

关于 * 处 \iff 的说明: ① \Rightarrow : 由 $\boldsymbol{\xi} = C\mathbf{x}$ 知显然. ② \Leftarrow : 只需注意到 $A\mathbf{x} + B\boldsymbol{\xi} = \mathbf{b}$ 蕴含着 $(-I_n - CA^{-1}B)\boldsymbol{\xi} = -CA^{-1}\mathbf{b}$. \square

注 2.0.5 更多内容可参见 [Sherman-Morrison-Woodbury 公式](#).

习题 2.4.13 设 $n+1$ 阶杨辉三角矩阵 $A = \begin{pmatrix} C_0^0 & & & \\ C_1^0 & C_1^1 & & \\ \vdots & \vdots & \ddots & \\ C_n^0 & C_n^1 & \cdots & C_n^n \end{pmatrix}$, $B = (b_{ij}) = (A^{-1} + A)^{-1}$,

$C = (c_{ij}) = (I + A)^{-1}$, 其中 $0 \leq i, j \leq n$. 证明: 当 $i - j$ 是奇数时, $b_{ij} = 0$; 当 $i - j$ 是正偶数时, $c_{ij} = 0$.

习题 2.4.14 设 S_1, S_2, \dots, S_{2^n} 是 $\{1, 2, \dots, n\}$ 的所有子集, $A = (a_{ij}) \in \mathbb{F}^{2^n \times 2^n}$,

$$a_{ij} = \begin{cases} 1, & S_i \subset S_j \\ 0, & S_i \not\subset S_j \end{cases}.$$

证明: 存在置换方阵 P 使得 PAP^T 是单位上三角方阵, 并且 $A^{-1} = ((-1)^{|S_i|+|S_j|} a_{ij})$.

习题 2.4.15 设 $A = (a_{ij}) \in \mathbb{F}^{n \times n}$, $a_{ij} = \begin{cases} 1, & j \mid i \\ 0, & j \nmid i \end{cases}$. 证明: $A^{-1} = \left(\mu\left(\frac{i}{j}\right) a_{ij}\right)$, 其中

$$\mu(x) = \begin{cases} (-1)^k, & x \text{ 是 } k \text{ 个不同素数的乘积,} \\ 0, & \text{否则.} \end{cases}$$

习题 2.4.16 首先给出几个定义.

- 设 $M \in \mathbb{F}^{n \times n}$. 若存在置换方阵 P , 使得 $P^T M P = \begin{pmatrix} M_{11} & M_{12} \\ O & M_{22} \end{pmatrix}$, 其中 M_{11}, M_{22} 都是阶数 ≥ 1 的方阵, 则 M 称为可约的, 否则 M 称为不可约的.

- 设 $M \in \mathbb{R}^{m \times n}$. 若 M 的元素都是非负实数, 则 M 称为非负矩阵. 若 M 的元素都是正实数, 则 M 称为全正矩阵.

- 设 $M \in \mathbb{R}^{n \times n}$ 是非负矩阵. 若存在正整数 k , 使得 M^k 是全正矩阵, 则 M 称为本原的, 否则 M 称为非本原的.

设 $A \in \mathbb{R}^{n \times n}$ 是非负矩阵. 证明:

(1) A 是不可约的 \iff 对于任意 $i \neq j$, 存在正整数 k , 使得 A^k 的 (i, j) 元素 $> 0 \iff$ 存在正整数 $k \leq n$, 使得 $I + A + \cdots + A^{k-1}$ 的元素都 $> 0 \iff$ 存在正实数 λ , 使得 $(\lambda I - A)^{-1}$ 的元素都 > 0 .

- (2) 存在置换方阵 P , 使得 $P^T A P$ 是准上三角方阵, 其中每个准对角块都是不可约方阵.
- (3) 若 A 是本原的, 则 A 是不可约的.
- (4) 若 A 是非本原且不可约的, 则存在正整数 $m \geq 2$ 和置换方阵 P , 使得

$$P^T A P = \begin{pmatrix} O & A_1 & & \\ & O & \ddots & \\ & & \ddots & A_{m-1} \\ A_m & & & O \end{pmatrix},$$

其中每个准对角块都是零方阵, 空白处元素都是 0.

提示 设 $S_i = \{k \in \mathbb{N} \mid (A^k)_{ii} > 0\}$, d_i 是 S_i 中所有元素的最大公约数.

注意到 $p, q \in S_i \Rightarrow p + q \in S_i$. 故当 k 充分大时, $kd_i \in S_i$.

对于任意 $i \neq j$, 由于 A 是不可约的, 存在正整数 p, q 使得 $(A^p)_{ij} > 0$, $(A^q)_{ji} \neq 0$. 当 k 充分大时, $(kd_i + p + q) \in S_j$. 得 $d_j \mid d_i$. 故 $d_1 = \cdots = d_n = m$.

综上, 对于任意 i, j , 存在 n_{ij} 使得 $(A^{n_{ij}+km})_{ij} > 0$, $(A^{n_{ij}+km+r})_{ij} = 0$, $\forall k \in \mathbb{N}$, $r \in \{1, \cdots, m-1\}$.

若 $m = 1$, 则当 k 充分大时, A^k 的元素都 > 0 , 与 A 非本原矛盾. 故 $m \geq 2$.

注意到 $n_{ii} \equiv 0$, $n_{ij} + n_{jk} \equiv n_{ik} \pmod{m}$, $\forall i, j, k$. 在 $\{1, 2, \cdots, n\}$ 上定义等价关系 $i \sim j \Leftrightarrow n_{ij} \equiv 0 \pmod{m}$, 得到 m 个等价类 I_1, \cdots, I_m .

$$A \begin{bmatrix} I_1 & \cdots & I_m \\ I_1 & \cdots & I_m \end{bmatrix} = \begin{pmatrix} O & A_1 & & \\ & O & \ddots & \\ & & \ddots & A_{m-1} \\ A_m & & & O \end{pmatrix}.$$

□

习题 3.1.7 (4) 计算以下方阵的行列式.

$$A = (a_{ij}) = \begin{pmatrix} 0 & a & b & c \\ -a & 0 & d & e \\ -b & -d & 0 & f \\ -c & -e & -f & 0 \end{pmatrix}.$$

解 1 观察到 A 为 4 阶反对称阵, 写出

$$A_4 = \{(1, 2, 3, 4), (1, 3, 2, 4), (1, 4, 2, 3)\},$$

A 的 Pfaffian 多项式 $\text{Pf}(A) = a_{12}a_{34} - a_{13}a_{24} + a_{14}a_{23} = af + cd - be$, 于是 $|A| = (af + cd - be)^2$.

□

解 2 将 A 按 2×2 分块:

$$\left(\begin{array}{cc|cc} 0 & a & b & c \\ -a & 0 & d & e \\ \hline -b & -d & 0 & f \\ -c & -e & -f & 0 \end{array} \right).$$

由 Schur 公式,

$$\begin{aligned} \det(A) &= \begin{vmatrix} 0 & a \\ -a & 0 \end{vmatrix} \cdot \left| \begin{pmatrix} 0 & f \\ -f & 0 \end{pmatrix} - \begin{pmatrix} -b & -d \\ -c & -e \end{pmatrix} \begin{pmatrix} 0 & a \\ -a & 0 \end{pmatrix}^{-1} \begin{pmatrix} b & c \\ d & e \end{pmatrix} \right| \\ &= a^2 \cdot \left| f \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} - \frac{1}{a} \begin{pmatrix} -b & -d \\ -c & -e \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} b & c \\ d & e \end{pmatrix} \right| \\ &= a^2 \cdot \left| f \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} + \frac{1}{a} \begin{pmatrix} b & d \\ c & e \end{pmatrix} \begin{pmatrix} -d & -e \\ b & c \end{pmatrix} \right| \\ &= a^2 \cdot \left| f \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} + \frac{1}{a} \begin{pmatrix} 0 & cd - be \\ be - cd & 0 \end{pmatrix} \right| \\ &= a^2 \cdot \left| \frac{af + cd - be}{a} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \right| \\ &= (af + cd - be)^2. \end{aligned}$$

□

注 2.0.6 本题的背景是反对称阵的 Pfaffian 多项式, 可参见[谢启鸿高等代数官方博客](https://zhuanlan.zhihu.com/p/269149418)以及 <https://zhuanlan.zhihu.com/p/269149418>.

习题 3.1.7 (7) 计算以下方阵的行列式.

$$A = (a_{ij}) = \begin{pmatrix} a & b & c & d \\ -b & a & -d & c \\ -c & d & a & -b \\ -d & -c & b & a \end{pmatrix}.$$

解 1 ① 首先考虑 $a = 0$ 的情形. 这时 A 为 4 阶反对称阵, 与习题 3.1.7 (4) 类似, 可以写出 A 的 Pfaffian 多项式

$$\text{Pf}(A) = a_{12}a_{34} - a_{13}a_{24} + a_{14}a_{23} = -b^2 - c^2 - d^2,$$

于是 $\det(A) = \text{Pf}(A)^2 = (b^2 + c^2 + d^2)^2$.

② 考虑 a 的一般情形. 注意到

$$\det(A) = \begin{vmatrix} a & b & c & d \\ -b & a & -d & c \\ -c & d & a & -b \\ -d & -c & b & a \end{vmatrix} \xrightarrow[\text{交换第 3、4 列}]{\text{交换第 1、2 列}} \begin{vmatrix} b & a & d & c \\ a & -b & c & -d \\ d & -c & -b & a \\ -c & -d & a & b \end{vmatrix} = \begin{vmatrix} b & a & d & c \\ -a & b & -c & d \\ -d & c & b & -a \\ -c & -d & a & b \end{vmatrix},$$

即在 A 中进行 $a \leftrightarrow b, c \leftrightarrow d$ 的对换后不改变 $\det(A)$. 因此 $\det(A)$ 作为以 a, b, c, d 为变元的多元多项式, 其中 a 与 b, c 与 d 地位等价, 且在 a 取 0 时化为 $(b^2 + c^2 + d^2)^2$. 于是 $\det(A) = (a^2 + b^2 + c^2 + d^2)^2$. \square

解 2 注意到 A 的任意两行 (列) 均正交, 所以 $\tilde{A} := \frac{1}{\sqrt{a^2 + b^2 + c^2 + d^2}}A$ 是正交方阵, 从而 $\det(A) = (a^2 + b^2 + c^2 + d^2)^2$. \square

习题 3.1.7 (10) 计算以下方阵的行列式.

$$A = \begin{pmatrix} 1+a & 1+a^2 & 1+a^3 & 1+a^4 \\ 1+b & 1+b^2 & 1+b^3 & 1+b^4 \\ 1+c & 1+c^2 & 1+c^3 & 1+c^4 \\ 1+d & 1+d^2 & 1+d^3 & 1+d^4 \end{pmatrix}.$$

解 1 通过加边可得

$$\begin{aligned}
 \det(A) &= \begin{vmatrix} 1 & -1 & -1 & -1 & -1 \\ 0 & 1+a & 1+a^2 & 1+a^3 & 1+a^4 \\ 0 & 1+b & 1+b^2 & 1+b^3 & 1+b^4 \\ 0 & 1+c & 1+c^2 & 1+c^3 & 1+c^4 \\ 0 & 1+d & 1+d^2 & 1+d^3 & 1+d^4 \end{vmatrix} = \begin{vmatrix} 1 & -1 & -1 & -1 & -1 \\ 1 & a & a^2 & a^3 & a^4 \\ 1 & b & b^2 & b^3 & b^4 \\ 1 & c & c^2 & c^3 & c^4 \\ 1 & d & d^2 & d^3 & d^4 \end{vmatrix} \\
 &= 2 \begin{vmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & a & a^2 & a^3 & a^4 \\ 1 & b & b^2 & b^3 & b^4 \\ 1 & c & c^2 & c^3 & c^4 \\ 1 & d & d^2 & d^3 & d^4 \end{vmatrix} - \begin{vmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & a & a^2 & a^3 & a^4 \\ 1 & b & b^2 & b^3 & b^4 \\ 1 & c & c^2 & c^3 & c^4 \\ 1 & d & d^2 & d^3 & d^4 \end{vmatrix} \\
 &= [2abcd - (a-1)(b-1)(c-1)(d-1)](b-a)(c-a)(d-a)(c-b)(d-b)(d-c).
 \end{aligned}$$

□

解 2 注意到

$$\begin{aligned}
 A &= \begin{pmatrix} a & & & \\ & b & & \\ & & c & \\ & & & d \end{pmatrix} \begin{pmatrix} 1 & a & a^2 & a^3 \\ 1 & b & b^2 & b^3 \\ 1 & c & c^2 & c^3 \\ 1 & d & d^2 & d^3 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 & 1 \end{pmatrix} \\
 &= \begin{pmatrix} a & & & \\ & b & & \\ & & c & \\ & & & d \end{pmatrix} \begin{pmatrix} 1 & a & a^2 & a^3 \\ 1 & b & b^2 & b^3 \\ 1 & c & c^2 & c^3 \\ 1 & d & d^2 & d^3 \end{pmatrix} \left[I_4 + \begin{pmatrix} 1 & a & a^2 & a^3 \\ 1 & b & b^2 & b^3 \\ 1 & c & c^2 & c^3 \\ 1 & d & d^2 & d^3 \end{pmatrix}^{-1} \begin{pmatrix} \frac{1}{a} \\ \frac{1}{b} \\ \frac{1}{c} \\ \frac{1}{d} \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 & 1 \end{pmatrix} \right].
 \end{aligned}$$

由 Sylvester 行列式恒等式得

$$\left| I_4 + \begin{pmatrix} 1 & a & a^2 & a^3 \\ 1 & b & b^2 & b^3 \\ 1 & c & c^2 & c^3 \\ 1 & d & d^2 & d^3 \end{pmatrix}^{-1} \begin{pmatrix} \frac{1}{a} \\ \frac{1}{b} \\ \frac{1}{c} \\ \frac{1}{d} \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 & 1 \end{pmatrix} \right| = \left| I_1 + \begin{pmatrix} 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & a & a^2 & a^3 \\ 1 & b & b^2 & b^3 \\ 1 & c & c^2 & c^3 \\ 1 & d & d^2 & d^3 \end{pmatrix}^{-1} \begin{pmatrix} \frac{1}{a} \\ \frac{1}{b} \\ \frac{1}{c} \\ \frac{1}{d} \end{pmatrix} \right|.$$

于是

$$\det(A) = abcd(b-a)(c-a)(d-a)(c-b)(d-b)(d-c) \left[1 + \begin{pmatrix} 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & a & a^2 & a^3 \\ 1 & b & b^2 & b^3 \\ 1 & c & c^2 & c^3 \\ 1 & d & d^2 & d^3 \end{pmatrix}^{-1} \begin{pmatrix} \frac{1}{a} \\ \frac{1}{b} \\ \frac{1}{c} \\ \frac{1}{d} \end{pmatrix} \right].$$

$$\text{要求 } \mathbf{x} := \begin{pmatrix} 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & a & a^2 & a^3 \\ 1 & b & b^2 & b^3 \\ 1 & c & c^2 & c^3 \\ 1 & d & d^2 & d^3 \end{pmatrix}^{-1}, \text{ 只需解方程 } \begin{pmatrix} 1 & 1 & 1 & 1 \\ a & b & c & d \\ a^2 & b^2 & c^2 & d^2 \\ a^3 & b^3 & c^3 & d^3 \end{pmatrix} \mathbf{x}^\top = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}. \text{ 运用}$$

Cramer 法则,

$$\Delta = (b-a)(c-a)(d-a)(c-b)(d-b)(d-c),$$

$$\Delta_1 = \Delta|_{a=1} = (b-1)(c-1)(d-1)(c-b)(d-b)(d-c),$$

$$\Delta_2 = \Delta|_{b=1} = (1-a)(c-a)(d-a)(c-1)(d-1)(d-c),$$

$$\Delta_3 = \Delta|_{c=1} = (b-a)(1-a)(d-a)(1-b)(d-b)(d-1),$$

$$\Delta_4 = \Delta|_{d=1} = (b-a)(c-a)(1-a)(c-b)(1-b)(1-c),$$

于是

$$\mathbf{x} = \left(\frac{(b-1)(c-1)(d-1)}{(b-a)(c-a)(d-a)}, \frac{(1-a)(c-1)(d-1)}{(b-a)(c-b)(d-b)}, \frac{(1-a)(1-b)(d-1)}{(c-a)(c-b)(d-c)}, \frac{(1-a)(1-b)(1-c)}{(d-a)(d-b)(d-c)} \right).$$

进而可得

$$\begin{aligned} \det(A) = & abcd(b-a)(c-a)(d-a)(c-b)(d-b)(d-c) \\ & + bcd(b-1)(c-1)(d-1)(c-b)(d-b)(d-c) \\ & + acd(1-a)(c-1)(d-1)(c-a)(d-a)(d-c) \\ & + abd(1-a)(1-b)(d-1)(b-a)(d-a)(d-b) \\ & + abc(1-a)(1-b)(1-c)(b-a)(c-a)(c-b). \end{aligned}$$

□

习题 3.2.1 设 $A \in \mathbb{C}^{m \times n}$. 利用 Binet–Cauchy 公式证明 $\det(AA^H) \geq 0$.

证明 设 M 是一个复方阵, 则 $\det(MM^H) = \det(M) \cdot \det(M^H) = \det(M) \cdot \overline{\det(M)} \geq 0$. 不妨设 $m \leq n$ (否则 $\det(AA^H) = 0$ 自然使结论成立). 于是由 Binet–Cauchy 公式, $\det(AA^H)$

可分解为 C_n^m 组 m 阶方阵乘积的行列式之和, 其中每组的两个方阵互为共轭转置, 它们的乘积的行列式为非负实数. 故 $\det(AA^H) \geq 0$. \square

习题 3.2.3 (1) 计算行列式:

$$\begin{vmatrix} 1 & \cos \theta_1 & \cdots & \cos(n-1)\theta_1 \\ 1 & \cos \theta_2 & \cdots & \cos(n-1)\theta_2 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \cos \theta_n & \cdots & \cos(n-1)\theta_n \end{vmatrix}.$$

解 可以证明第 I 类 Chebyshev 多项式的三角形式 $T_n(\cos \theta) = \cos(n\theta)$: 比较 De Moivre 公式 $(\cos \theta + i \sin \theta)^n = \cos n\theta + i \sin n\theta$ 两边实部可得

$$\cos n\theta = \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} \binom{n}{2k} \cos^{n-2k} \theta \cdot i^{2k} \sin^{2k} \theta = \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} (-1)^k \binom{n}{2k} \cos^{n-2k} \theta (1 - \cos^2 \theta)^k.$$

由此还能得到 $\deg(T_n(x)) = n$, 且

$$T_n(x) \text{ 的首项系数} = \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} (-1)^k \binom{n}{2k} (-1)^k = 2^{n-1}.$$

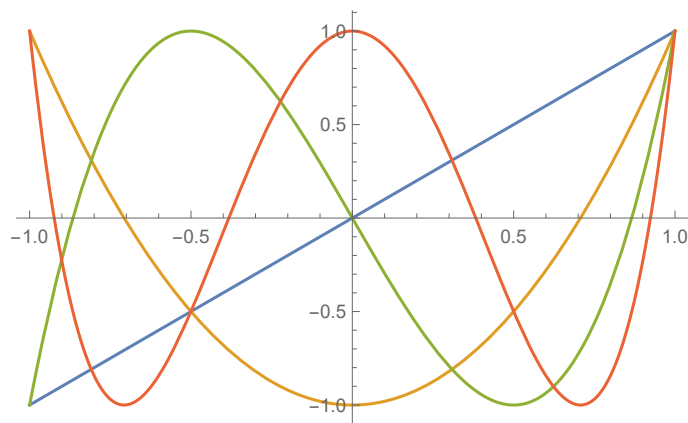


图 2.1: 第 I 类 Chebyshev 多项式

由此可将原来的方阵分解为以下形式:

$$\begin{pmatrix} 1 & \cos \theta_1 & \cdots & \cos(n-1)\theta_1 \\ 1 & \cos \theta_2 & \cdots & \cos(n-1)\theta_2 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \cos \theta_n & \cdots & \cos(n-1)\theta_n \end{pmatrix} = \begin{pmatrix} 1 & \cos \theta_1 & \cdots & \cos^{n-1} \theta_1 \\ 1 & \cos \theta_2 & \cdots & \cos^{n-2} \theta_2 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \cos \theta_n & \cdots & \cos^{n-1} \theta_n \end{pmatrix} \begin{pmatrix} 1 & * & * & \cdots & * \\ & 1 & * & \cdots & * \\ & & 2 & \cdots & * \\ & & & \ddots & \vdots \\ & & & & 2^{n-2} \end{pmatrix}.$$

故结果为 $2^{\frac{(n-1)(n-2)}{2}} \prod_{1 \leq i < j \leq n} (\cos \theta_j - \cos \theta_i)$. □

习题 3.2.3 (2) 计算行列式:

$$\begin{vmatrix} \cos \frac{1}{2}\theta_1 & \cos \frac{3}{2}\theta_1 & \cdots & \cos \frac{2n-1}{2}\theta_1 \\ \cos \frac{1}{2}\theta_2 & \cos \frac{3}{2}\theta_2 & \cdots & \cos \frac{2n-1}{2}\theta_2 \\ \vdots & \vdots & \ddots & \vdots \\ \cos \frac{1}{2}\theta_n & \cos \frac{3}{2}\theta_n & \cdots & \cos \frac{2n-1}{2}\theta_n \end{vmatrix}.$$

解 由于 $\cos \frac{2k+1}{2}\theta + \cos \frac{2k-1}{2}\theta = 2 \cos k\theta \cos \frac{1}{2}\theta$, 从最右一列起, 将前一列加到该列上, 可将原行列式转化为

$$\begin{vmatrix} \cos \frac{1}{2}\theta_1 & 2 \cos \theta_1 \cos \frac{1}{2}\theta_1 & \cdots & 2 \cos(n-1)\theta_1 \cos \frac{1}{2}\theta_1 \\ \cos \frac{1}{2}\theta_2 & 2 \cos \theta_2 \cos \frac{1}{2}\theta_2 & \cdots & 2 \cos(n-1)\theta_2 \cos \frac{1}{2}\theta_2 \\ \vdots & \vdots & \ddots & \vdots \\ \cos \frac{1}{2}\theta_n & 2 \cos \theta_n \cos \frac{1}{2}\theta_n & \cdots & 2 \cos(n-1)\theta_n \cos \frac{1}{2}\theta_n \end{vmatrix} \\ = 2^{n-1} \prod_{k=1}^n \cos \frac{1}{2}\theta_k \cdot \begin{vmatrix} 1 & \cos \theta_1 & \cdots & \cos(n-1)\theta_1 \\ 1 & \cos \theta_2 & \cdots & \cos(n-1)\theta_2 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \cos \theta_n & \cdots & \cos(n-1)\theta_n \end{vmatrix}.$$

由习题 3.2.3 (1) 结果可得本题结果为 $2^{\frac{n(n-1)}{2}} \prod_{k=1}^n \cos \frac{1}{2}\theta_k \prod_{1 \leq i < j \leq n} (\cos \theta_j - \cos \theta_i)$. □

习题 3.2.3 (3) 计算行列式:

$$\begin{vmatrix} \sin \theta_1 & \sin 2\theta_1 & \cdots & \sin n\theta_1 \\ \sin \theta_2 & \sin 2\theta_2 & \cdots & \sin n\theta_2 \\ \vdots & \vdots & \ddots & \vdots \\ \sin \theta_n & \sin 2\theta_n & \cdots & \sin n\theta_n \end{vmatrix}.$$

解 可以证明第 II 类 Chebyshev 多项式的三角形 $U_n(\cos \theta) = \frac{\sin(n+1)\theta}{\sin \theta}$: 比较 De Moivre 公式 $(\cos \theta + i \sin \theta)^n = \cos n\theta + i \sin n\theta$ 两边虚部可得

$$\sin(n+1)\theta = \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} \binom{n+1}{2k+1} \cos^{n-2k} \theta \cdot i^{2k} \sin^{2k+1} \theta = \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} (-1)^k \binom{n+1}{2k+1} \cos^{n-2k} \theta (1 - \cos^2 \theta)^k \sin \theta.$$

由此还能得到 $\deg(U_n(x)) = n$, 且

$$U_n(x) \text{ 的首项系数} = \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} (-1)^k \binom{n+1}{2k+1} (-1)^k = 2^n.$$

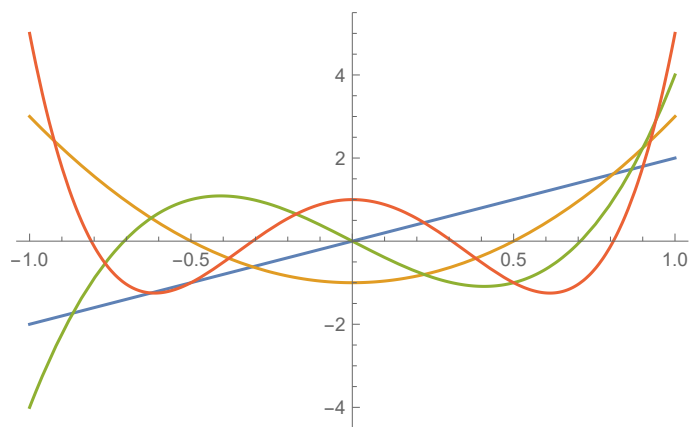


图 2.2: 第 II 类 Chebyshev 多项式

由此可将原来的方阵分解为以下形式:

$$\begin{pmatrix} \sin \theta_1 & \sin 2\theta_1 & \cdots & \sin n\theta_1 \\ \sin \theta_2 & \sin 2\theta_2 & \cdots & \sin n\theta_2 \\ \vdots & \vdots & \ddots & \vdots \\ \sin \theta_n & \sin 2\theta_n & \cdots & \sin n\theta_n \end{pmatrix} = \begin{pmatrix} 1 & \cos \theta_1 & \cdots & \cos^{n-1} \theta_1 \\ 1 & \cos \theta_2 & \cdots & \cos^{n-2} \theta_2 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \cos \theta_n & \cdots & \cos^{n-1} \theta_n \end{pmatrix} \begin{pmatrix} \sin \theta_1 & * & \cdots & * \\ & 2 \sin \theta_2 & \cdots & * \\ & & \ddots & \vdots \\ & & & 2^{n-1} \sin \theta_n \end{pmatrix}.$$

故结果为 $2^{\frac{n(n-1)}{2}} \prod_{k=1}^n \sin \theta_k \prod_{1 \leq i < j \leq n} (\cos \theta_j - \cos \theta_i)$ □

习题 3.2.3 (4) 计算行列式:

$$\begin{vmatrix} \sin \frac{1}{2}\theta_1 & \sin \frac{3}{2}\theta_1 & \cdots & \sin \frac{2n-1}{2}\theta_1 \\ \sin \frac{1}{2}\theta_2 & \sin \frac{3}{2}\theta_2 & \cdots & \sin \frac{2n-1}{2}\theta_2 \\ \vdots & \vdots & \ddots & \vdots \\ \sin \frac{1}{2}\theta_n & \sin \frac{3}{2}\theta_n & \cdots & \sin \frac{2n-1}{2}\theta_n \end{vmatrix}.$$

解 由于 $\sin \frac{2k+1}{2}\theta - \sin \frac{2k-1}{2}\theta = 2 \sin \frac{1}{2}\theta \cos k\theta$, 从最右一列起, 将该列减去前一列, 可将

原行列式转化为

$$\begin{vmatrix} \sin \frac{1}{2}\theta_1 & 2 \sin \frac{1}{2}\theta_1 \cos \theta_1 & \cdots & 2 \sin \frac{1}{2}\theta_1 \cos(n-1)\theta_1 \\ \sin \frac{1}{2}\theta_2 & 2 \sin \frac{1}{2}\theta_2 \cos \theta_2 & \cdots & 2 \sin \frac{1}{2}\theta_2 \cos(n-1)\theta_2 \\ \vdots & \vdots & \ddots & \vdots \\ \sin \frac{1}{2}\theta_n & 2 \sin \frac{1}{2}\theta_n \cos \theta_n & \cdots & 2 \sin \frac{1}{2}\theta_n \cos(n-1)\theta_n \end{vmatrix}$$

$$= 2^{n-1} \prod_{k=1}^n \sin \frac{1}{2}\theta_k \cdot \begin{vmatrix} 1 & \cos \theta_1 & \cdots & \cos(n-1)\theta_1 \\ 1 & \cos \theta_2 & \cdots & \cos(n-1)\theta_2 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \cos \theta_n & \cdots & \cos(n-1)\theta_n \end{vmatrix}.$$

由习题 3.2.3 (1) 结果可得本题结果为 $2^{\frac{n(n-1)}{2}} \prod_{k=1}^n \sin \frac{1}{2}\theta_k \prod_{1 \leq i < j \leq n} (\cos \theta_j - \cos \theta_i)$. \square

习题 3.2.3 (5) 旧 计算行列式:

$$\begin{vmatrix} 1^2 & 2^2 & 3^2 & \cdots & n^2 \\ 2^2 & 3^2 & 4^2 & \cdots & 1^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ n^2 & 1^2 & 2^2 & \cdots & (n-1)^2 \end{vmatrix}.$$

解 先从首尾开始向中间依次对调两行, 将行列式转化为循环方阵行列式, 再将第一列进行 $(n-1)$ 次列交换后换到最后一列:

$$(-1)^{\frac{n(n-1)}{2}} \begin{vmatrix} n^2 & 1^2 & 2^2 & \cdots & (n-1)^2 \\ (n-1)^2 & n^2 & 1^2 & \cdots & (n-2)^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1^2 & 2^2 & 3^2 & \cdots & n^2 \end{vmatrix} = (-1)^{\frac{(n+2)(n-1)}{2}} \begin{vmatrix} 1^2 & 2^2 & \cdots & (n-1)^2 & n^2 \\ n^2 & 1^2 & \cdots & (n-2)^2 & (n-1)^2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 2^2 & 3^2 & \cdots & n^2 & 1^2 \end{vmatrix}.$$

由讲义例 3.10 结论, 上式可化为

$$(-1)^{\frac{n(n-1)}{2}} \prod_{k=0}^{n-1} f(\omega^k),$$

其中 $f(x) = \sum_{k=1}^n k^2 x^{k-1}$, $\omega = \cos \frac{2\pi}{n} + i \sin \frac{2\pi}{n}$ 是 n 次单位根. 我们可以用母函数法得到 $f(x)$

的封闭表达式:

$$\begin{aligned} \sum_{k=1}^n x^k &= \frac{x^{n+1} - x}{x - 1} \xrightarrow{\text{求导}} \sum_{k=1}^n kx^{k-1} = \frac{nx^{n+1} - (n+1)x^n + 1}{(x-1)^2} \\ &\xrightarrow{\text{两边乘}x} \sum_{k=1}^n kx^k = \frac{nx^{n+2} - (n+1)x^{n+1} + x}{(x-1)^2} \\ &\xrightarrow{\text{求导}} \sum_{k=1}^n k^2 x^{k-1} = \frac{(-2n^2 - 2n + 1)x^{n+1} + n^2 x^{n+2} + (n+1)^2 x^n - x - 1}{(x-1)^3} \end{aligned}$$

□

习题 3.2.3 (6) 计算行列式:
$$\begin{vmatrix} s_0 & s_1 & \cdots & s_{n-1} \\ s_1 & s_2 & \cdots & s_n \\ \vdots & \vdots & \ddots & \vdots \\ s_{n-1} & s_n & \cdots & s_{2n-2} \end{vmatrix}, \text{ 其中 } s_k = \sum_{i=1}^n x_i^k.$$

解 观察到此方阵为 $(s_{(i-1)+(j-1)})$. 若记 $V = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{n-1} & x_2^{n-1} & \cdots & x_n^{n-1} \end{pmatrix}$, 则原方阵即 VV^T .

于是行列式为 $\prod_{1 \leq i < j \leq n} (x_j - x_i)^2$.

□

习题 3.2.3 (7) 计算行列式:
$$\begin{vmatrix} c_0 & c_1 & \cdots & c_{n-1} \\ -c_{n-1} & c_0 & \cdots & c_{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ -c_1 & -c_2 & \cdots & c_0 \end{vmatrix}.$$

解 设 $f(x) = \sum_{k=0}^{n-1} c_k x^k$, $Z = \begin{pmatrix} & & & I_{n-1} \\ -1 & & & \end{pmatrix}$. 则原方阵 $C = f(Z)$, $Z^n = -I$. 设 n 阶复方阵 $\Omega = (\omega^{(i-1)(2j-1)})$, 其中 $\omega = \cos \frac{\pi}{n} + i \sin \frac{\pi}{n}$ 为 $2n$ 次单位根. 记 $A = (a_{ij}) = \Omega \bar{\Omega}$, 则

$$a_{ij} = \sum_{k=1}^n \omega^{(i-1)(2k-1)} \bar{\omega}^{(2k-1)(j-1)} = \sum_{k=1}^n (\omega \bar{\omega})^{(2k-1)(j-1)} \omega^{(2k-1)(i-j)} = \sum_{k=1}^n \omega^{(2k-1)(i-j)} = n \delta_{ij}.$$

这说明 $A = \Omega \bar{\Omega} = nI$, 故 Ω 可逆, $\Omega^{-1} = \frac{1}{n} \bar{\Omega}$. 又因为

$$\begin{aligned} Z\Omega &= \begin{pmatrix} \omega & \omega^3 & \cdots & \omega^{2n-1} \\ \omega^2 & \omega^{2 \times 3} & \cdots & \omega^{2(2n-1)} \\ \vdots & \vdots & \ddots & \vdots \\ \omega^{(n-1)} & \omega^{(n-1) \times 3} & \cdots & \omega^{(n-1)(2n-1)} \\ -1 & -1 & \cdots & -1 \end{pmatrix} = \begin{pmatrix} \omega & \omega^3 & \cdots & \omega^{2n-1} \\ \omega^2 & \omega^{2 \times 3} & \cdots & \omega^{2(2n-1)} \\ \vdots & \vdots & \ddots & \vdots \\ \omega^{(n-1)} & \omega^{(n-1) \times 3} & \cdots & \omega^{(n-1)(2n-1)} \\ \omega^n & \omega^{n \times 3} & \cdots & \omega^{n(2n-1)} \end{pmatrix} \\ &= \underbrace{\begin{pmatrix} 1 & 1 & \cdots & 1 \\ \omega & \omega^3 & \cdots & \omega^{2n-1} \\ \omega^2 & \omega^{2 \times 3} & \cdots & \omega^{2(2n-1)} \\ \vdots & \vdots & \ddots & \vdots \\ \omega^{(n-1)} & \omega^{(n-1) \times 3} & \cdots & \omega^{(n-1)(2n-1)} \end{pmatrix}}_{\Omega} \begin{pmatrix} \omega & & & \\ & \omega^3 & & \\ & & \ddots & \\ & & & \omega^{2n-1} \end{pmatrix}, \end{aligned}$$

所以 $Z = \Omega \operatorname{diag}(\omega, \omega^3, \dots, \omega^{2n-1}) \Omega^{-1}$. 记 $D = \operatorname{diag}(\omega, \omega^3, \dots, \omega^{2n-1})$, 则

$$Z^k = \Omega D^k \Omega^{-1} \implies C = f(Z) = \Omega f(D) \Omega^{-1}.$$

于是

$$\det(C) = \det(f(D)) = \det \begin{pmatrix} f(\omega) & & & \\ & f(\omega^3) & & \\ & & \ddots & \\ & & & f(\omega^{2n-1}) \end{pmatrix} = \prod_{k=1}^n f(\omega^{2k-1}).$$

□

习题 3.2.3 (8) 计算行列式: $\begin{vmatrix} C_{p+1}^{q+1} & C_{p+1}^{q+2} & \cdots & C_{p+1}^{q+n} \\ C_{p+2}^{q+1} & C_{p+2}^{q+2} & \cdots & C_{p+2}^{q+n} \\ \vdots & \vdots & \cdots & \vdots \\ C_{p+n}^{q+1} & C_{p+n}^{q+2} & \cdots & C_{p+n}^{q+n} \end{vmatrix}$, 其中 C_{p+i}^{q+j} 是组合数.

解 为了将矩阵 $\left(\frac{(p+i)!}{(q+j)!(p-q+i-j)!} \right)$ 中 $(p+1)!$ 和 $(q+j)!$ 这两个“ i 与 j 不混合”的系数提出来, 我们借助矩阵分块的视角, 将该矩阵分解为

$$\operatorname{diag}((p+i)!) \left(\frac{1}{(p-q+i-j)!} \right) \operatorname{diag} \left(\frac{1}{(q+j)!} \right).$$

为了构造 Vandermonde 行列式, 我们进一步改写为

$$\text{diag} \left(\frac{(p+i)!}{(p-q+i-1)!} \right) \left(\frac{(p-q+i-1)!}{[(p-q+i-1)-j+1]!} \right) \text{diag} \left(\frac{1}{(q+j)!} \right).$$

记 $x_i = p - q + i - 1$, 则已经构造出 $f_j(x_i) = \frac{x_i!}{(x_i - j + 1)!}$ 的形式, $\deg(f_j(x)) = j - 1$. 再利用次数高低顺序与带余除法进行列变换:

$$\begin{vmatrix} f_1(x_1) & f_2(x_1) & \cdots & f_n(x_1) \\ f_1(x_2) & f_2(x_2) & \cdots & f_n(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ f_1(x_n) & f_2(x_n) & \cdots & f_n(x_n) \end{vmatrix} = \begin{vmatrix} 1 & x_1 & \cdots & x_1^{n-1} \\ 1 & x_2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & \cdots & x_n^{n-1} \end{vmatrix} = \prod_{1 \leq i < j \leq n} (j-i) = \prod_{i=1}^{n-1} (n-i)! = \prod_{i=1}^n (i-1)!.$$

故所求行列式结果为 $\prod_{k=1}^n \frac{(k-1)!(p+k)!}{(p-q+k-1)!(q+k)!}$. □

习题 3.2.4 证明:
$$\begin{vmatrix} 1+a^2-b^2-c^2 & 2(ab+c) & 2(ca-b) \\ 2(ab-c) & 1+b^2-c^2-a^2 & 2(bc+a) \\ 2(ca+b) & 2(bc-a) & 1+c^2-a^2-b^2 \end{vmatrix} = (1+a^2+b^2+c^2)^3.$$

证明

$$\begin{aligned} \text{LHS} & \stackrel{\textcircled{1}+=(b \times \textcircled{3}-c \times \textcircled{2})}{=} \begin{vmatrix} 1+a^2+b^2+c^2 & c(1+a^2+b^2+c^2) & -b(1+a^2+b^2+c^2) \\ 2ab-2c & 1+b^2-c^2-a^2 & 2bc+2a \\ 2ac+2b & 2bc-2a & 1+c^2-a^2-b^2 \end{vmatrix} \\ & = (1+a^2+b^2+c^2) \cdot \begin{vmatrix} 1 & c & -b \\ 2ab-2c & 1+b^2-c^2-a^2 & 2bc+2a \\ 2ac+2b & 2bc-2a & 1+c^2-a^2-b^2 \end{vmatrix} \\ & = (1+a^2+b^2+c^2) \cdot \begin{vmatrix} 1 & 0 & 0 \\ 2ab-2c & 1+b^2+c^2-a^2-2abc & 2ab^2+2a \\ 2ac+2b & -2a-2ac^2 & 1+c^2-a^2+b^2+2abc \end{vmatrix} \\ & = (1+a^2+b^2+c^2) [(1+b^2+c^2-a^2)^2 - 4a^2b^2c^2 + 4a^2(b^2+1)(c^2+1)] \\ & = (1+a^2+b^2+c^2)^3. \end{aligned}$$

□

习题 3.2.5 设 n 阶复数方阵 $\Omega = \left(\omega^{(i-1)(j-1)} \right)$, 其中 $\omega = \cos \frac{2\pi}{n} + i \sin \frac{2\pi}{n}$ 是 n 次单位根.

证明: $\det(\Omega) = n^{\frac{n}{2}} e^{i\theta}$, 其中 $\theta = \frac{(n-1)(3n-2)}{4}\pi$.

证明 ① 先证明 $\det(\Omega)$ 模长为 $n^{\frac{n}{2}}$. 只需证 $|\det(\Omega)|^2 = \det(\Omega)\overline{\det(\Omega)} = \det(\Omega\overline{\Omega}) = n^n$. 记 $A = (a_{ij}) \Omega\overline{\Omega}$, 则

$$a_{ij} = \sum_{k=1}^n \omega^{(i-1)(k-1)} \overline{\omega}^{(k-1)(j-1)} = \sum_{k=1}^n (\omega\overline{\omega})^{(k-1)(j-1)} \omega^{(k-1)(i-j)} = \sum_{k=1}^n \omega^{(k-1)(i-j)} = n\delta_{ij}.$$

这说明 $A = \Omega\overline{\Omega} = nI$. 故 $\det(\Omega\overline{\Omega}) = n^n$.

② 再证明辐角 $\theta = \frac{(n-1)(3n-2)}{4}\pi$. 如图 2.3, 由几何关系, 当 $0 \leq i < j \leq n-1$ 时,

$$\begin{aligned} \arg(\omega^j - \omega^i) &= \arg(\omega^j) + \frac{1}{2} [\pi - (\arg(\omega^j) - \arg(\omega^i))] \\ &= \frac{2\pi j}{n} + \frac{1}{2} \left[\pi - \frac{2\pi(j-i)}{n} \right] = \frac{\pi(2j+2i+n)}{2n}. \end{aligned}$$

于是

$$\begin{aligned} \arg(\det(\Omega)) &= \arg\left(\prod_{0 \leq i < j \leq n-1} (\omega^j - \omega^i)\right) = \sum_{0 \leq i < j \leq n-1} \arg(\omega^j - \omega^i) \\ &= \sum_{0 \leq i < j \leq n-1} \frac{\pi(2j+2i+n)}{2n} = \frac{\pi}{2n} \sum_{i=0}^{n-2} \sum_{j=i+1}^{n-1} (2j+2i+n) \\ &= \frac{(n-1)(3n-2)}{4}\pi. \end{aligned}$$

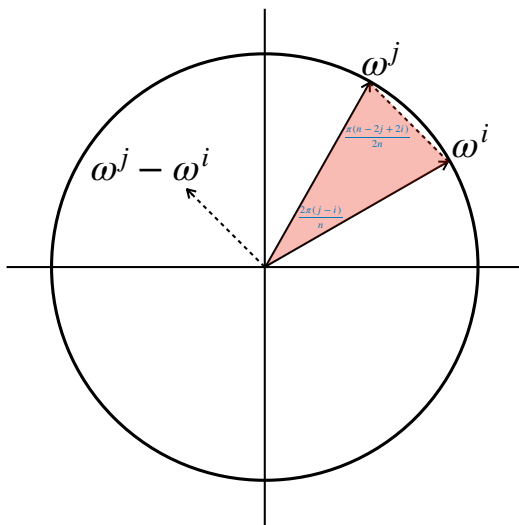


图 2.3

□

习题 3.2.7 (1) 计算行列式和逆矩阵: $A = \begin{pmatrix} 1+a_1b_1 & a_1b_2 & \cdots & a_1b_n \\ a_2b_1 & 1+a_2b_2 & \cdots & a_2b_n \\ \vdots & \vdots & \ddots & \vdots \\ a_nb_1 & a_nb_2 & \cdots & 1+a_nb_n \end{pmatrix}.$

解 由 Sylvester 行列式恒等式,

$$\begin{aligned}\det(A) &= \det \left(I_n + \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \begin{pmatrix} b_1 & b_2 & \cdots & b_n \end{pmatrix} \right) \\ &= \det \left(I_1 + \begin{pmatrix} b_1 & b_2 & \cdots & b_n \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \right) = 1 + \sum_{i=1}^n a_i b_i.\end{aligned}$$

代数余子式 $A_{ii} = 1 - a_i b_i + \sum_{k=1}^n a_k b_k$, 而当 $i \neq j$ 时, 若 $i < j$, 提取因子后 A_{ij} 化为

$$(-1)^{i+j} a_j b_i \begin{vmatrix} 1+a_1 b_1 & a_1 b_2 & \cdots & a_1 & \cdots & a_1 b_n \\ a_2 b_1 & 1+a_2 b_2 & \cdots & a_2 & \cdots & a_2 b_n \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 1+a_{i-1} b_{i-1} & a_{i-1} & \cdots & a_{i-1} & \cdots & a_{i-1} b_n \\ a_{i+1} & 1+a_{i+1} b_{i+1} & \cdots & 1+a_{j-1} b_{j-1} & \cdots & b_{j-1} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ b_1 & b_2 & \cdots & \textcircled{1} & \cdots & b_n \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_n b_1 & a_n b_2 & \cdots & a_n & \cdots & 1+a_n b_n \end{vmatrix}$$

利用第 i 列 (元素 1 所在列) 消元

$$(-1)^{i+j} a_j b_i \begin{vmatrix} 1 & 0 & \cdots & a_1 & \cdots & 0 \\ 0 & 1 & \cdots & a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 1 & a_{i-1} & \cdots & a_{i-1} & \cdots & 0 \\ a_{i+1} & 1 & \cdots & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n & \cdots & 1 \end{vmatrix}$$

将红色分块内的第 1 列经 $j-i-1$ 次换列调至最右列

$$(-1)^{i+j} \cdot (-1)^{j-i-1} a_j b_i = (-1)^{2j-1} a_j b_i = -a_j b_i.$$

同理, 当 $i > j$ 时, 也有 $A_{ij} = -a_j b_i$. 故可写出 $A^{-1} = \frac{1}{\det(A)} A^*$. □

习题 3.2.16 设 $f(x) = \sum_{i=0}^m f_i x^i$ 和 $g(x) = \sum_{i=0}^n g_i x^i$ 都是复系数多项式, $f_m g_n \neq 0$. 如下 $m+n$

阶方阵

$$S = \left(\begin{array}{cccc} f_0 & f_1 & \cdots & f_m \\ & \ddots & \ddots & \ddots & \ddots \\ & & f_0 & f_1 & \cdots & f_m \\ g_0 & g_1 & \cdots & g_n \\ & \ddots & \ddots & \ddots & \ddots \\ & & g_0 & g_1 & \cdots & g_n \end{array} \right) \left. \begin{array}{l} \\ \\ \\ \end{array} \right\} \begin{array}{l} n \text{ 行} \\ \\ m \text{ 行} \end{array}$$

称为 Sylvester 方阵, $\det(S)$ 称为 Sylvester 结式. 证明:

(1) $\det(S) = f_m^n g_n^m \prod_{i=1}^m \prod_{j=1}^n (v_j - u_i)$, 其中 u_1, \dots, u_m 和 v_1, \dots, v_n 分别是 f 和 g 的所有根.

(2) $\det(S) \neq 0$ 当且仅当 $\gcd(f, g) = 1$.

习题 4.1.1 (4)(5)(6) 计算下列矩阵的秩.

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 4 & 5 & 6 \\ 3 & 4 & 5 & 6 & 7 \\ 4 & 5 & 6 & 7 & 8 \end{pmatrix} \quad \begin{pmatrix} 1^2 & 2^2 & 3^2 & 4^2 & 5^2 \\ 2^2 & 3^2 & 4^2 & 5^2 & 6^2 \\ 3^2 & 4^2 & 5^2 & 6^2 & 7^2 \\ 4^2 & 5^2 & 6^2 & 7^2 & 8^2 \end{pmatrix} \quad \begin{pmatrix} 1^3 & 2^3 & 3^3 & 4^3 & 5^3 \\ 2^3 & 3^3 & 4^3 & 5^3 & 6^3 \\ 3^3 & 4^3 & 5^3 & 6^3 & 7^3 \\ 4^3 & 5^3 & 6^3 & 7^3 & 8^3 \end{pmatrix}$$

解 考虑 $m \times n$ 矩阵 $A = (a_{ij})$, 其中 $a_{ij} = (i + j - 1)^k, k \in \mathbb{N}$. 由

$$(i + j - 1)^k = \sum_{s=0}^k \binom{k}{s} (i - 1)^{k-s} j^s$$

将 i 与 j 分离可得

$$A = \underbrace{\left((i - 1)^{k-j+1} \right)_{m \times (k+1)}}_B \underbrace{\text{diag} (C_k^0, C_k^1, \dots, C_k^k)}_D \underbrace{\left(j^{i-1} \right)_{(k+1) \times n}}_C.$$

因为 B 中可找到由 $0, 1, \dots, \min\{m, k+1\} - 1$ 生成的 Vandermonde 方阵作为子矩阵, 所以 $\text{rank}(B) = \min\{m, k+1\}$. 因为 C 中可找到由 $1, 2, \dots, \min\{n, k+1\}$ 生成的 Vandermonde 方阵作为子矩阵, 所以 $\text{rank}(C) = \min\{n, k+1\}$. 又因为 $\text{rank}(D) = k+1$, 所以

$$\begin{aligned} \text{rank}(A) &= \text{rank}(BDC) \leq \min\{\text{rank}(B), \text{rank}(D), \text{rank}(C)\} \\ &= \min\{\min\{m, k+1\}, k+1, \min\{n, k+1\}\} = \min\{m, n, k+1\}. \end{aligned}$$

另一方面, 记 $r = \min\{m, n, k+1\}$, 往证 $\text{rank}(A) \geq r$. 注意到: ① 行满秩矩阵只需经过列变换即可化成标准形; ② 列满秩矩阵只需经过行变换即可化成标准形; ③ 任一矩阵左乘列满秩或者右乘行满秩矩阵不改变原矩阵的秩. 未完待续 □

习题 4.1.4 证明: 任意非零矩阵 A 都可以表示为 BC 的形式, 其中 B 是列满秩的, C 是行满秩的.

证明 设 $A = P \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} Q$, 其中 $P = \begin{pmatrix} P_1 & P_2 \\ P_3 & P_4 \end{pmatrix}$, $Q = \begin{pmatrix} Q_1 & Q_2 \\ Q_3 & Q_4 \end{pmatrix}$ 是可逆方阵. 由 $A \neq O$ 知 $r \geq 1$, 于是

$$A = \begin{pmatrix} P_1 & P_2 \\ P_3 & P_4 \end{pmatrix} \begin{pmatrix} I_r & \\ & O \end{pmatrix} \begin{pmatrix} Q_1 & Q_2 \\ Q_3 & Q_4 \end{pmatrix} = \begin{pmatrix} P_1 \\ P_3 \end{pmatrix} (Q_1 \quad Q_2).$$

由 P, Q 是满秩方阵, 以及线性无关向量组的任一部分组线性无关, 可知 $\begin{pmatrix} P_1 \\ P_3 \end{pmatrix}$ 列满秩, $(Q_1 \quad Q_2)$ 行满秩. 这就是所要的分解. \square

习题 4.1.5 (1) 设 $A \in \mathbb{F}^{m \times n}$. 证明以下命题彼此等价:

- ① A 有右逆.
- ② A 是行满秩的.
- ③ $A^T \mathbf{x} = \mathbf{0}$ 只有零解.
- ④ 存在 $B \in \mathbb{F}^{(n-m) \times n}$, 使得 $\begin{pmatrix} A \\ B \end{pmatrix}$ 可逆.

证明 ① \Rightarrow ②: 设 B 为 A 的一个右逆: $AB = I_m$. 则 $m = \text{rank}(AB) \leq \text{rank}(A) \leq m$, 因此 $\text{rank}(A) = m$, A 行满秩.

② \Rightarrow ③: 因为行满秩矩阵只需经过列变换即可化成标准形, 所以可设 $A = (I_m \quad O)Q$, 其中 $Q \in \mathbb{F}^{n \times n}$ 是可逆方阵. 于是 $A^T \mathbf{x} = Q^T \begin{pmatrix} I_m \\ O \end{pmatrix} \mathbf{x} = \mathbf{0} \iff \begin{pmatrix} I_m \\ O \end{pmatrix} \mathbf{x} = \mathbf{0} \iff \mathbf{x} = \mathbf{0}$.

③ \Rightarrow ④: 因为 $A^T \mathbf{x} = \mathbf{0}$ 只有零解, 所以可经过行变换得到 $PA^T = \begin{pmatrix} I_m \\ O \end{pmatrix}$, 其中 $A \in \mathbb{F}^{n \times n}$ 是可逆方阵. 于是 $A^T = P^{-1} \begin{pmatrix} I_m \\ O \end{pmatrix}$. 取 $B = (O \quad I_{n-m})(P^{-1})^T$, 则 $B^T = P^{-1} \begin{pmatrix} O \\ I_{n-m} \end{pmatrix}$, $\begin{pmatrix} A^T & B^T \end{pmatrix} = P^{-1} \begin{pmatrix} I_m & \\ & I_{n-m} \end{pmatrix} = P^{-1} I_n = P^{-1}$. 故 $\begin{pmatrix} A \\ B \end{pmatrix} = (P^{-1})^T$ 可逆.

④ \Rightarrow ①: 设 $\begin{pmatrix} A \\ B \end{pmatrix} \begin{pmatrix} C_{n \times m} & D_{n \times (n-m)} \end{pmatrix} = I_n$, 则由分块乘法得 $AC = I_m$. \square

习题 4.1.6 (1)(2) 设 $A, C \in \mathbb{F}^{m \times n}$, $B, D \in \mathbb{F}^{n \times p}$. 证明:

- (1) 若 A, B 都是行满秩的, 则 AB 也是行满秩的.
 (2) 若 C, D 都是列满秩的, 则 CD 也是列满秩的.

证明 只需证明:

- (1) 右乘行满秩矩阵不改变原矩阵的秩.
 (2) 左乘列满秩矩阵不改变原矩阵的秩.

下面给出 (1) 的证明, (2) 同理. 设 $A_{m \times n} B_{n \times p} = E_{m \times p}$. 一方面, $\text{rank}(E) = \text{rank}(AB) \leq \text{rank}(A)$. 另一方面, 由 Sylvester 秩不等式, $\text{rank}(E) = \text{rank}(AB) \geq \text{rank}(A) + \text{rank}(B) - n = \text{rank}(A)$. 故 $\text{rank}(E) = \text{rank}(A)$. \square

习题 4.1.7 (1) 设 $A \in \mathbb{F}^{n \times n}$, A^* 是 A 的伴随方阵. 证明: $\text{rank}(A^*) = \begin{cases} n, & \text{rank}(A) = n \\ 1, & \text{rank}(A) = n - 1 \\ 0, & \text{rank}(A) \leq n - 2 \end{cases}$.

证明 设 $A = P \begin{pmatrix} I_r & \\ & O \end{pmatrix} Q$, 其中 $P, Q \in \mathbb{F}^{n \times n}$ 是可逆方阵. 则

$$\text{rank}(A^*) = \text{rank} \left(P \begin{pmatrix} I_r & \\ & O \end{pmatrix} Q \right)^* \stackrel{\textcircled{1}}{=} \text{rank} \left(Q^* \begin{pmatrix} I_r & \\ & O \end{pmatrix}^* P^* \right) \stackrel{\textcircled{2}}{=} \text{rank} \begin{pmatrix} I_r & \\ & O \end{pmatrix}^*.$$

上式中等号 ① 来自习题 3.3.7, 等号 ② 是因为 $P^* = \det(P)P^{-1}$, $Q^* = \det(Q)Q^{-1}$ 可逆, 而左 (右) 乘可逆方阵不改变原矩阵的秩. 注意到 $\begin{pmatrix} I_r & \\ & O \end{pmatrix}$ 中非零元素 (若有) 只能位于主对

角线上, 若 $r = n$, 则 $\begin{pmatrix} I_r & \\ & O \end{pmatrix}^*$ 主对角线全为非零元素, $\text{rank} \begin{pmatrix} I_r & \\ & O \end{pmatrix}^* = n$; 若 $r = n - 1$, $\begin{pmatrix} I_r & \\ & O \end{pmatrix}^* = \text{diag}(0, \dots, 0, 1)$, $\text{rank} \begin{pmatrix} I_r & \\ & O \end{pmatrix}^* = 1$; 若 $r \leq n - 2$, 则 $\begin{pmatrix} I_r & \\ & O \end{pmatrix}^* = O_{n \times n}$, $\text{rank} \begin{pmatrix} I_r & \\ & O \end{pmatrix}^* = 0$. \square

习题 4.1.7 (2) 求所有 $A, B \in \mathbb{C}^{n \times n}$ 满足 $A^* = B^* \neq O$.

证明 由习题 4.1.7 (1) 结论, $A^* \neq O \iff \text{rank}(A^*) \geq 1 \iff \text{rank}(A) \geq n - 1$. 由习题 3.3.7, $(A^*)^* = (\det(A))^{n-2} A$, 因此 $A^* = B^* \iff (\det(A))^{n-2} A = (\det(B))^{n-2} B$.

① 若 $\text{rank}(A) = n$, 则 $\det(A) \neq 0$, 从而 $\det(B) \neq 0$. 记 $\lambda = \frac{(\det(A))^{n-2}}{(\det(B))^{n-2}}$, 则 $B = \lambda A$, 由习题 3.3.7 可知

$$B^* = (\lambda A)^* = \lambda^{n-1} A^* = A^* \iff \lambda^{n-1} = 1.$$

故 $B = \lambda A$, 其中 $\lambda = \cos \frac{2k\pi}{n-1} + i \sin \frac{2k\pi}{n-1}$ ($k = 0, \dots, n-2$).

② 若 $\text{rank}(A) = n-1$, 则 $\text{rank}(A^*) = 1$. 设 $A^* = P \text{diag}(1, O_{n-1}) Q$, 其中 P, Q 是可逆方阵, 则

$$A^* = P \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & \cdots & 0 \end{pmatrix} Q = \alpha \beta^T,$$

其中 α 是 P 的最后一列, β^T 是 Q 的最后一行. 未完待续

□

习题 4.1.8 (3) 证明秩等式:

$$\text{rank}(A \otimes B) = \text{rank}(A) \text{rank}(B).$$

证明 设 $A = P_1 \begin{pmatrix} I_{r_1} & \\ & O \end{pmatrix} Q_1$, $B = P_2 \begin{pmatrix} I_{r_2} & \\ & O \end{pmatrix} Q_2$, 其中 P_1, Q_1, P_2, Q_2 都是可逆方阵. 由 Kronecker 积的混合积性质 (习题 2.2.7) 可得

$$A \otimes B = (P_1 \otimes P_2) \left(\begin{pmatrix} I_{r_1} & \\ & O \end{pmatrix} \otimes \begin{pmatrix} I_{r_2} & \\ & O \end{pmatrix} \right) (Q_1 \otimes Q_2).$$

而 $P_1 \otimes P_2$ 与 $Q_1 \otimes Q_2$ 均可逆, 因此

$$\text{rank}(A \otimes B) = \text{rank} \left(\begin{pmatrix} I_{r_1} & \\ & O \end{pmatrix} \otimes \begin{pmatrix} I_{r_2} & \\ & O \end{pmatrix} \right) = r_1 r_2 = \text{rank}(A) \text{rank}(B).$$

□

习题 4.1.8 (5) 证明秩不等式:

$$\text{rank}(AC - BD) \leq \text{rank}(A - B) + \text{rank}(C - D),$$

其中 $A, B \in \mathbb{F}^{m \times n}$, $C, D \in \mathbb{F}^{n \times p}$.

证明 ① 若 $B = O$, 则欲证不等式变为 $\text{rank}(AC) \leq \text{rank}(A) + \text{rank}(C - D)$, 这可由 $\text{rank}(AC) \leq \text{rank}(A)$ 直接得到. 类似地, 若 $C = O$, 则欲证不等式也平凡地成立.

② 下面考虑 $B \neq O$ 且 $C \neq O$ 的情形. 由初等变换可得

$$\begin{aligned} \text{rank}(A - B) + \text{rank}(C - D) &= \text{rank} \begin{pmatrix} A - B & \\ & C - D \end{pmatrix} = \text{rank} \begin{pmatrix} AC - BC & \\ & BC - BD \end{pmatrix} \\ &= \text{rank} \begin{pmatrix} AC - BC & BC - BD \\ & BC - BD \end{pmatrix} = \text{rank} \begin{pmatrix} AC - BD & BC - BD \\ BC - BD & BC - BD \end{pmatrix} \geq \text{rank}(AC - BD). \end{aligned}$$

□

习题 4.1.8 (6) 证明秩不等式:

$$-\frac{n}{2} \leq \text{rank}(AB) - \text{rank}(BA) \leq \frac{m}{2},$$

其中 $A \in \mathbb{F}^{m \times n}$, $B \in \mathbb{F}^{n \times m}$.

证明 不妨设 $A = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$, $B = \begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix}$, 其中 $B_1 \in \mathbb{F}^{r \times r}$. 若 $\text{rank}(A) \leq \frac{m}{2}$, 则

$$\text{rank}(AB) - \text{rank}(BA) \leq \text{rank}(AB) \leq \text{rank}(A) \leq \frac{m}{2},$$

右边不等式已经成立, 故可不妨设 $r = \text{rank}(A) > \frac{m}{2}$, 从而 $\text{rank}(B_2) \leq m - r < \frac{m}{2}$. 由

$$AB = \begin{pmatrix} B_1 & B_2 \\ O & O \end{pmatrix}, \quad BA = \begin{pmatrix} B_1 & O \\ B_3 & O \end{pmatrix}$$

可知

$$\text{rank}(AB) = \text{rank} \begin{pmatrix} B_1 & B_2 \end{pmatrix} \leq \text{rank}(B_1) + \text{rank}(B_2),$$

$$\text{rank}(BA) = \text{rank} \begin{pmatrix} B_1 \\ B_3 \end{pmatrix} \geq \text{rank}(B_1).$$

于是

$$\text{rank}(AB) - \text{rank}(BA) \leq \text{rank}(B_1) + \text{rank}(B_2) - \text{rank}(B_1) = \text{rank}(B_2) \leq \frac{m}{2}.$$

将 A, B 互换即可得到左边不等式.

□

习题 4.1.10 设复系数多项式 $f(x) = \sum_{i=0}^m f_i x^i$, $g(x) = \sum_{i=0}^n g_i x^i$, 构造 $m+n$ 阶复数方阵

$$S = \left(\begin{array}{cccc} f_0 & f_1 & \cdots & f_m \\ & \ddots & \ddots & \ddots & \ddots \\ & & f_0 & f_1 & \cdots & f_m \\ g_0 & g_1 & \cdots & g_n \\ & \ddots & \ddots & \ddots & \ddots \\ & & g_0 & g_1 & \cdots & g_n \end{array} \right) \left. \begin{array}{l} \\ \\ \\ \\ \end{array} \right\} \begin{array}{l} n \text{ 行} \\ \\ \\ m \text{ 行} \end{array}$$

求证: $\text{rank}(S) = \max\{\deg(f) + n, \deg(g) + m\} - \deg(\gcd(f, g))$.

证明 设 $u(x) = \sum_{k=0}^{n-1} u_k x^k$, $v(x) = \sum_{k=0}^{m-1} v_k x^k$. 注意到

$$\begin{pmatrix} u_0 & \cdots & u_{n-1} & v_0 & \cdots & v_{m-1} \end{pmatrix} S \begin{pmatrix} 1 \\ x \\ x^2 \\ \vdots \\ x^{m+n-1} \end{pmatrix} = u(x)f(x) + v(x)g(x),$$

考虑解空间, 视 \mathbf{x} 为 $\begin{pmatrix} u_0 & \cdots & u_{n-1} & v_0 & \cdots & v_{m-1} \end{pmatrix}^\top$, 就有

$$\begin{aligned} \text{rank}(S) &= \text{rank}(S^\top) = m+n - \dim\{\mathbf{x} \mid S^\top \mathbf{x} = \mathbf{0}\} \\ &= m+n - \dim\{(u_0, \cdots, u_{n-1}, v_0, \cdots, v_{m-1}) \mid u(x)f(x) + v(x)g(x) = 0\}. \end{aligned}$$

而

$$u(x)f(x) + v(x)g(x) = 0 \iff u(x) \cdot \frac{f(x)}{\gcd(f, g)} = -v(x) \cdot \frac{g(x)}{\gcd(f, g)},$$

因为 $\gcd\left(\frac{f(x)}{\gcd(f, g)}, \frac{g(x)}{\gcd(f, g)}\right) = 1$, 所以上式又等价于

$$\frac{f(x)}{\gcd(f, g)} \Big| v(x), \quad \frac{g(x)}{\gcd(f, g)} \Big| u(x).$$

记

$$u(x) = w(x) \cdot \frac{g(x)}{\gcd(f, g)}, \quad v(x) = -w(x) \cdot \frac{f(x)}{\gcd(f, g)},$$

则 $w(x)$ 成为自由变量, 且满足约束条件

$$\begin{cases} \deg(w) = \deg(u) + \deg(\gcd(f, g)) - \deg(g) \leq n - 1 + \deg(\gcd(f, g)) - \deg(g), \\ \deg(w) = \deg(v) + \deg(\gcd(f, g)) - \deg(f) \leq m - 1 + \deg(\gcd(f, g)) - \deg(f). \end{cases}$$

故

$$\begin{aligned} \text{rank}(S) &= m + n - (\max \{\deg(w)\} + 1) \\ &= m + n - (\min \{n - \deg(g), m - \deg(f)\} + \deg(\gcd(f, g)) - 1 + 1) \\ &= m + n + \max \{\deg(g) - n, \deg(f) - m\} - \deg(\gcd(f, g)) \\ &= \max \{\deg(f) + n, \deg(g) + m\} - \deg(\gcd(f, g)). \end{aligned}$$

□

习题 4.1.11 设复系数多项式 $f(x) = \sum_{i=0}^n f_i x^i$, $g(x) = \sum_{i=0}^n g_i x^i$, 构造 n 阶复数方阵 $B = (b_{ij})$,

$$\frac{f(x)g(y) - f(y)g(x)}{x - y} = \sum_{i,j=1}^n b_{ij} x^{i-1} y^{j-1}.$$

求证: $\text{rank}(B) = \max\{\deg(f), \deg(g)\} - \deg(\gcd(f, g))$.

习题 4.2.2 设 $A_1, A_2 \in \mathbb{F}^{m \times n}$, $b_1, b_2 \in \mathbb{F}^{m \times 1}$. 证明: 若线性方程组 $A_1 \mathbf{x} = \mathbf{b}_1$ 与 $A_2 \mathbf{x} = \mathbf{b}_2$ 在 $\mathbb{F}^{n \times 1}$ 中有相同的非空解集, 则存在可逆方阵 $P \in \mathbb{F}^{m \times m}$, 使得 $PA_1 = A_2$ 且 $P\mathbf{b}_1 = \mathbf{b}_2$.

证明 设 $A_1 \mathbf{x} = \mathbf{b}_1$ 与 $A_2 \mathbf{x} = \mathbf{b}_2$ 的解空间为 $\{\alpha_0 + \lambda_1 \alpha_1 + \cdots + \lambda_{n-r} \alpha_{n-r} \mid \lambda_1, \cdots, \lambda_{n-r} \in \mathbb{F}\}$, 其中 $r = \text{rank}(A_1) = \text{rank}(A_2)$. 由 $\alpha_1, \cdots, \alpha_{n-r}$ 线性无关可设 $\alpha_1, \cdots, \alpha_{n-r} = Q \begin{pmatrix} I_{n-r} \\ O \end{pmatrix}$, 其

中 $Q \in \mathbb{F}^{(n-r) \times (n-r)}$ 为可逆方阵. 再由 $A_i Q \begin{pmatrix} I_{n-r} \\ O \end{pmatrix} = O$ 可设 $A_i Q = \begin{pmatrix} O & C_i \end{pmatrix}$ ($i = 1, 2$). 因

为 $\text{rank}(C_i) = \text{rank}(A_i) = r$, 所以可设 $C_i = P_i \begin{pmatrix} I_r \\ O \end{pmatrix}$, 其中 $P_i \in \mathbb{F}^{m \times m}$ 为可逆方阵, $i = 1, 2$.

取 $P = P_2 P_1^{-1}$ 就有 $PA_1 = P_2 P_1^{-1} \begin{pmatrix} O & C_1 \end{pmatrix} Q^{-1} = \begin{pmatrix} O & C_2 \end{pmatrix} Q^{-1} = A_2$. □

习题 4.2.3 设矩阵 A 是行 (列) 满秩. 证明: B 是 A 的广义逆矩阵 $\iff B$ 是 A 的右 (左) 逆.

证明 以 A 行满秩为例:

\Rightarrow : 设 $A = P \begin{pmatrix} I & O \end{pmatrix} Q$, 其中 P, Q 是可逆方阵. 则 $B = Q^{-1} \begin{pmatrix} I \\ O \end{pmatrix} P^{-1}$,

$$AB = P \begin{pmatrix} I & O \end{pmatrix} Q Q^{-1} \begin{pmatrix} I \\ O \end{pmatrix} P^{-1} = P I P^{-1} = I.$$

\Leftarrow : 若 B 是 A 的一个右逆, 则

$$ABA = (AB)A = IA = A, \quad BAB = B(AB) = BI = B.$$

这即是广义逆矩阵的等价定义. □

习题 4.2.4 设矩阵 $A = PQ$, 其中 P 列满秩, Q 行满秩. 证明:

(1) 若 $A = \tilde{P}\tilde{Q}$, \tilde{P} 列满秩, \tilde{Q} 行满秩, 则存在可逆方阵 R , 使得 $\tilde{P} = PR$, $\tilde{Q} = R^{-1}Q$.

(2) 若 P^\dagger, Q^\dagger 分别是 P, Q 的广义逆矩阵, 则 $A^\dagger = Q^\dagger P^\dagger$ 是 A 的广义逆矩阵.

(3) A 的每个广义逆矩阵都可以表示为 $Q^\dagger P^\dagger$ 的形式, 其中 P^\dagger, Q^\dagger 分别是 P, Q 的广义逆矩阵.

证明 (1) 设 $\tilde{P}^\dagger, \tilde{Q}^\dagger$ 为 \tilde{P}, \tilde{Q} 的广义逆矩阵. 由习题 4.2.3 知 $\tilde{P}^\dagger \tilde{P} = \tilde{Q} \tilde{Q}^\dagger = I_r$, 其中 $r = \text{rank}(A)$. 由 $\tilde{P}\tilde{Q} = PQ$, 两边同时左乘 \tilde{P}^\dagger 得到 $\tilde{Q} = \tilde{P}^\dagger PQ$. 同理可得 $\tilde{P} = PQ\tilde{Q}^\dagger$. 将它们代入 $PQ = \tilde{P}\tilde{Q}$ 即得 $PQ = PQ\tilde{Q}^\dagger \tilde{P}^\dagger PQ$. 设 P^\dagger, Q^\dagger 为 P, Q 的广义逆矩阵, 将它们分别左乘、右乘于等式两端得到

$$P^\dagger PQQ^\dagger = P^\dagger PQ\tilde{Q}^\dagger \tilde{P}^\dagger PQQ^\dagger \implies I_r = Q\tilde{Q}^\dagger \tilde{P}^\dagger P.$$

故取 $R = Q\tilde{Q}^\dagger$, 就有 $R^{-1} = \tilde{P}^\dagger P$. 由上可知 R 满足所需性质.

(2) 只需验证其满足广义逆矩阵的等价定义:

$$AA^\dagger A = PQQ^\dagger P^\dagger PQ = PQ = A, \quad A^\dagger AA^\dagger = Q^\dagger P^\dagger PQQ^\dagger P^\dagger = Q^\dagger P^\dagger = A^\dagger.$$

(3) 设 B 是 A 的任一广义逆矩阵. 由广义逆矩阵的等价定义 (的第一条),

$$ABA = A \implies PQBPQ = PQ.$$

设 P^\dagger, Q^\dagger 为 P, Q 的广义逆矩阵. 由习题 4.2.3 知 $P^\dagger P = QQ^\dagger = I_r$. 将 P^\dagger, Q^\dagger 分别左乘、右乘于等式两端得到 $QBP = I_r$. 再在两边同时分别左乘、右乘 Q^\dagger, P^\dagger 就得到 $B = Q^\dagger P^\dagger$. □

习题 4.2.5 (1) 设 A 是对称方阵. 证明: 对于任意正整数 $k \leq \text{rank}(A)$, A 有 k 或 $k+1$ 阶可逆主子矩阵.

证明 ① 先证明 A 有 $\text{rank}(A)$ 阶可逆主子矩阵. 设 $A = P \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} Q$, 其中 P, Q 是可逆方阵. 则

$$A = P \left(\underbrace{\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}}_{\text{选前 } r \text{ 行, 前 } r \text{ 列}} QP^{-\top} \right) P^{\top} = P \underbrace{\begin{pmatrix} R & O \\ O & O \end{pmatrix}}_{\text{记为 } S} P^{\top}.$$

由 $A^{\top} = P \begin{pmatrix} R^{\top} & O \\ O & O \end{pmatrix} P^{\top} = A$ 可知 $R^{\top} = R$, 即 R 是对称可逆方阵. 将 P 分块为 $(P_1 \ P_2)$,

其中 P_1 列数为 r . 那么 $A = (P_1 \ O) \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} \begin{pmatrix} P_1^{\top} \\ O \end{pmatrix}$. 因为 P_1 列满秩, 所以可选出 P_1 的 r 阶可逆子矩阵 $P_1 [i_1, \dots, i_r]$. 由 $A = PSP^{\top}$ 及 Binet-Cauchy 公式,

$$\det(A [i_1, \dots, i_r]) = \det(P [i_1, \dots, i_r]) \det(R) \det(P^{\top} [i_1, \dots, i_r]) = [\det(P [i_1, \dots, i_r])]^2 \det(R) \neq 0.$$

故 $A [i_1, \dots, i_r]$ 就是 A 的 $\text{rank}(A)$ 阶可逆主子矩阵.

② 设 $A \in \mathbb{F}^{n \times n}$. 当 $n = 1$ 时显然成立. 下面对 n 归纳, 设当 $n - 1$ 时命题成立. 由

$$\begin{aligned} A &= \left(\begin{array}{c|c} A_1 & \begin{smallmatrix} * \\ \vdots \\ * \end{smallmatrix} \\ \hline * \dots * & a_{nn} \end{array} \right) \\ &= \left(\begin{array}{c|c} A_1 & \begin{smallmatrix} 0 \\ \vdots \\ 0 \end{smallmatrix} \\ \hline 0 \dots 0 & 0 \end{array} \right) + \left(\begin{array}{c|c} O & \begin{smallmatrix} * \\ \vdots \\ * \end{smallmatrix} \\ \hline 0 \dots 0 & a_{nn} \end{array} \right) + \left(\begin{array}{c|c} O & \begin{smallmatrix} 0 \\ \vdots \\ 0 \end{smallmatrix} \\ \hline * \dots * & 0 \end{array} \right) \end{aligned}$$

知 $\text{rank}(A) \leq \text{rank}(A_1) + 2$. 下面分 3 类讨论:

- 若 $\text{rank}(A) = \text{rank}(A_1)$. 由 A_1 的可逆主子矩阵也是 A 的可逆主子矩阵即得证.
- 若 $\text{rank}(A) = \text{rank}(A_1) + 1$. 只需证当 $k = \text{rank}(A)$ 时命题成立, 只需注意到 A 的 $\text{rank}(A)$ 阶可逆主子矩阵即 k 阶可逆主子矩阵.
- 若 $\text{rank}(A) = \text{rank}(A_1) + 2$. 只需证当 $k = \text{rank}(A_1) + 1$ 时命题成立. 只需注意到 A 的 $\text{rank}(A)$ 阶可逆主子矩阵即 $k + 1$ 阶可逆主子矩阵. \square

推论 2.0.7 若 A 是对称方阵, 则 A 一定有 $\text{rank}(A)$ 阶可逆主子矩阵.

注 2.0.8 设

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

则 A 的任意奇数阶主子矩阵均不可逆 (因此 A 的任意偶数阶主子矩阵均可逆), 由此说明习题 4.2.5 (1) 的结论不能再改进.

习题 4.2.5 (3) 设 A 是反对称方阵且对角元素都是 0¹. 证明: $r = \text{rank}(A)$ 是偶数, 并且对于任意正偶数 $k \leq r$, A 有 k 阶可逆主子矩阵.

证明

□

习题 4.2.11 设方阵 A 满足 $A^m = O$, $A^{m-1} \neq O$, $m \geq 1$. A 称为幂零方阵, m 称为 A 的幂零指数. 证明:

(1) 存在可逆方阵 P , 使得 $P^{-1}AP$ 是上三角方阵.

(2) 存在可逆方阵 P , 使得 $B = P^{-1}AP = \begin{pmatrix} O & B_1 & & \\ & O & \ddots & \\ & & \ddots & B_{m-1} \\ & & & O \end{pmatrix}$, 其中每个对角块是零

方阵, B_i 形如 $\begin{pmatrix} I & O \end{pmatrix}$, 其余空白处元素都是 0.

(3) 存在可逆方阵 P , 使得 $C = P^{-1}AP = \begin{pmatrix} O & C_1 & & \\ & O & \ddots & \\ & & \ddots & C_{m-1} \\ & & & O \end{pmatrix}$, 其中每个对角块是零

方阵, C_i 形如 $\begin{pmatrix} I \\ O \end{pmatrix}$, 其余空白处元素都是 0. C 称为 Weyr 块.

¹ 对角元素都是 0 并不能由 A 是反对称方阵推出, 因为当 $\text{char } \mathbb{F} = 2$ 时, $1 = -1$.

(4) 存在置换方阵 P , 使得 $P^T B P = \text{diag}(J_1, J_2, \dots, J_s)$, 其中每个 $J_i = \begin{pmatrix} I_{n_i-1} \\ 0 \end{pmatrix}$ 是 n_i 阶方阵, $m = n_1 \geq n_2 \geq \dots \geq n_s$. $\text{diag}(J_1, J_2, \dots, J_s)$ 称为 Jordan 方阵, J_i 称为 Jordan 块.

证明 (1) 对幂零指数 (或最小多项式) 归纳. 当幂零指数为 1 时, 该方阵本身即为零方阵, 结论成立.

假设当幂零指数为 $m-1$ ($m \geq 2$) 时结论成立. 设 $A = P_0 \begin{pmatrix} I & \\ & O \end{pmatrix} Q_0$, 其中 P_0, Q_0 是可逆方阵, 并将 $Q_0 P_0$ 对应地分块为 $\begin{pmatrix} R_1 & R_2 \\ R_3 & R_4 \end{pmatrix}$, 则

$$A = P_0 \begin{pmatrix} I & \\ & O \end{pmatrix} Q_0 P_0^{-1} = P_0 \begin{pmatrix} I & \\ & O \end{pmatrix} \begin{pmatrix} R_1 & R_2 \\ R_3 & R_4 \end{pmatrix} P_0^{-1} = P_0 \begin{pmatrix} R_1 & R_2 \\ O & O \end{pmatrix} P_0^{-1}.$$

由此可见对任意 $k \geq 1$,

$$A^k = P_0 \begin{pmatrix} R_1 & R_2 \\ O & O \end{pmatrix}^k P_0^{-1} = P_0 \begin{pmatrix} R_1^k & R_1^{k-1} R_2 \\ O & O \end{pmatrix} P_0^{-1}.$$

由 A 的幂零指数为 m 可知

$$R_1^{m-1} \begin{pmatrix} R_1 & R_2 \end{pmatrix} = O \quad \text{且} \quad R_1^{m-2} \begin{pmatrix} R_1 & R_2 \end{pmatrix} \neq O.$$

由 R_1, R_2 的设法可知 $\begin{pmatrix} R_1 & R_2 \end{pmatrix}$ 行满秩, 因此将以上两式的等号两边同时右乘 $\begin{pmatrix} R_1 & R_2 \end{pmatrix}$ 的一个右逆就得到 $R_1^{m-2} \neq O$ 且 $R_1^{m-1} = O$, 即 R_1 的幂零指数为 $m-1$. 由归纳假设, 存在可逆方阵 Q 使 $Q^{-1} R_1 Q = U$ 是上三角方阵. 故

$$A = P_0 \begin{pmatrix} Q U Q^{-1} & * \\ O & O \end{pmatrix} P_0^{-1} = P_0 \begin{pmatrix} Q & \\ & 1 \end{pmatrix} \begin{pmatrix} U & * \\ O & O \end{pmatrix} \begin{pmatrix} Q^{-1} & \\ & 1 \end{pmatrix} P_0^{-1}.$$

将 $P_0 \begin{pmatrix} Q & \\ & 1 \end{pmatrix}$ 取为 P 即为所求.

(2) 从 (1) 的证明过程可见, 上述结论可加强为: 存在可逆方阵 P , 使得 $P^{-1} A P = (B_{ij})_{m \times m}$ 是准上三角方阵, 其中对角块 B_{ii} 都是零方阵. 对幂零指数归纳. ① 当幂零指数为 1 时, 该方阵本身即为零方阵, 结论成立. ② 假设当幂零指数为 $m-1$ ($m \geq 2$) 时结论成立. 按照 (1) 的记号可得

$$A = P_0 \begin{pmatrix} R_1 & R_2 \\ O & O \end{pmatrix} P_0^{-1}.$$

由归纳假设, 存在可逆方阵 Q 使 $Q^{-1}R_1Q = \begin{pmatrix} O & B_1 & & \\ & O & \ddots & \\ & & \ddots & B_{m-2} \\ & & & O \end{pmatrix}$. 于是

$$A = P_0 \begin{pmatrix} Q & \\ & I \end{pmatrix} \begin{pmatrix} O & B_1 & & * \\ & O & \ddots & * \\ & & \ddots & B_{m-2} & * \\ & & & O & * \\ & & & & O \end{pmatrix} \begin{pmatrix} Q^{-1} & \\ & I \end{pmatrix} P_0^{-1}.$$

由此可设 $A \sim \begin{pmatrix} O & B_1 & & C_1 \\ & O & \ddots & C_2 \\ & & \ddots & B_{m-2} & \vdots \\ & & & O & C_{m-1} \\ & & & & O \end{pmatrix}$ 将 $P_0 \begin{pmatrix} Q & \\ & I \end{pmatrix}$. 注意到利用 B_i 形如 $\begin{pmatrix} I & O \end{pmatrix}$

的特点, 可以这样将 C_1 变为 O :

$$\underbrace{\begin{pmatrix} I & & & \\ & I & & C_1 \\ & & \ddots & \\ & & & I \\ & & & & I \end{pmatrix}}_{\text{只是为了凑相似, 实际无效果}} \begin{pmatrix} O & B_1 & & C_1 \\ & O & \ddots & C_2 \\ & & \ddots & B_{m-2} & \vdots \\ & & & O & C_{m-1} \\ & & & & O \end{pmatrix} \begin{pmatrix} I & & & \\ & I & & -C_1 \\ & & \ddots & \\ & & & I \\ & & & & I \end{pmatrix}$$

$$= \begin{pmatrix} O & B_1 & & O \\ & O & \ddots & C_2 \\ & & \ddots & B_{m-2} & \vdots \\ & & & O & C_{m-1} \\ & & & & O \end{pmatrix},$$

重复类似操作可得 $A \sim \begin{pmatrix} O & B_1 & & O \\ & O & \ddots & O \\ & & \ddots & B_{m-2} & \vdots \\ & & & O & C_{m-1} \\ & & & & O \end{pmatrix}$. 注意到该矩阵从第一行到 C_{m-1} 的最
后一行所在行构成的子矩阵行满秩, 从而 C_{m-1} 行满秩. 因此存在可逆方阵 R 使 $C_{m-1}R^{-1} = \begin{pmatrix} I & O \end{pmatrix}$. 故

$$\underbrace{\begin{pmatrix} I & & & \\ & I & & \\ & & \ddots & \\ & & & I \\ & & & & R \end{pmatrix}}_{\text{只是为了凑相似, 实际无效果}} \begin{pmatrix} O & B_1 & & O \\ & O & \ddots & O \\ & & \ddots & B_{m-2} & \vdots \\ & & & O & C_{m-1} \\ & & & & O \end{pmatrix} \begin{pmatrix} I & & & \\ & I & & \\ & & \ddots & \\ & & & I \\ & & & & R^{-1} \end{pmatrix} \\ = \begin{pmatrix} O & B_1 & & \\ & O & \ddots & \\ & & \ddots & B_{m-2} \\ & & & O & B_{m-1} \\ & & & & O \end{pmatrix}.$$

由相似是等价关系即得证.

(3) 由 (2) 可知, $A^T \sim \begin{pmatrix} O & B_1 & & \\ & O & \ddots & \\ & & \ddots & B_{m-1} \\ & & & O \end{pmatrix}$, 于是

$$A \sim \begin{pmatrix} O & & & \\ B_1^T & O & & \\ & \ddots & \ddots & \\ & & B_{m-1}^T & O \end{pmatrix} \sim \begin{pmatrix} O & B_{m-1}^T & & \\ & O & \ddots & \\ & & \ddots & B_1^T \\ & & & O \end{pmatrix},$$

其中第二个相似用到了 $\begin{pmatrix} & I \\ \ddots & \\ I & \end{pmatrix}$ (起“中心对称”效果). 取 $C_i = B_{m-i}^T$ 即得证.

(4) 注意到 (2) 中 B 的特点是每行最多有 1 个 1. 我们从第一行开始, 找到最靠左的 1, 沿其所在列向下找到主对角线上的 0, 再向右找第一个 1, 接着向下找主对角线上的 0……如此重复直至找不到 1; 然后在前面未选中的行中找到最靠上的含 1 的行, 从这个 1 开始重复上述步骤. 最终就能根据以上的选法选出所有的 1, 将 B 重排成 Jordan 方阵, 这个选主子矩阵的过程自然对应了一个置换相似. \square

习题 5.1.5 设 $A \in \mathbb{R}^{n \times n}$. 证明:

- (1) $A + A^T$ 的特征值都是实数.
- (2) $A - A^T$ 的特征值都是纯虚数或 0.
- (3) 若 n 是奇数, 则 A 必有实特征值.
- (4) 若 A 无实特征值, 则 $\det(A) > 0$.

证明 (1) 等价于证明“若 A 是对称方阵, 则 A 的特征值都是实数”. 设 (λ, α) 是 A 的一个特征对 (在 \mathbb{C} 和 $\mathbb{C}^{n \times 1}$ 上考虑), 则

$$\begin{aligned} A\alpha = \lambda\alpha &\implies \alpha^H A\alpha = \lambda\alpha^H\alpha \implies (\alpha^H A\alpha)^H = \bar{\lambda}\alpha^H\alpha \\ &\implies \alpha^H A\alpha = \bar{\lambda}\alpha^H\alpha \implies \lambda\alpha^H\alpha = \bar{\lambda}\alpha^H\alpha \xrightarrow{\alpha^H\alpha > 0} \lambda = \bar{\lambda}. \end{aligned}$$

(2) 等价于证明“若 A 是反对称方阵”, 则 A 的特征值都是纯虚数或 0. 设 (λ, α) 是 A 的一个特征对 (在 \mathbb{C} 和 $\mathbb{C}^{n \times 1}$ 上考虑), 则

$$\begin{aligned} A\alpha = \lambda\alpha &\implies \alpha^H A\alpha = \lambda\alpha^H\alpha \implies (\alpha^H A\alpha)^H = \bar{\lambda}\alpha^H\alpha \\ &\implies -\alpha^H A\alpha = \bar{\lambda}\alpha^H\alpha \implies -\lambda\alpha^H\alpha = \bar{\lambda}\alpha^H\alpha \xrightarrow{\alpha^H\alpha > 0} \lambda + \bar{\lambda} = 0 \iff \operatorname{Re} \lambda = 0. \end{aligned}$$

(3) 因为 $\varphi_A(x)$ 在 \mathbb{C} 上有奇数个根, 而共轭虚根成对出现, 所以必有实根.

(4) 由 A 无实特征值可知 n 是偶数. 设 $\lambda_1, \dots, \lambda_{\frac{n}{2}}, \bar{\lambda}_1, \dots, \bar{\lambda}_{\frac{n}{2}} \in \mathbb{C} \setminus \mathbb{R}$ 为 A 的所有特征值. 由相似三角化可得 $\det(A) = \prod_{i=1}^{\frac{n}{2}} \lambda_i \bar{\lambda}_i > 0$. \square

习题 5.1.8 设 $A \in \mathbb{F}^{n \times n}$ 不是纯量方阵. 证明:

(1) 对于任意 $b_1, b_2, \dots, b_{n-1} \in \mathbb{F}$, 存在可逆方阵 $P \in \mathbb{F}^{n \times n}$ 使得 $P^{-1}AP$ 的前 $n-1$ 个对角元素依次是 b_1, b_2, \dots, b_{n-1} .

(2) 若 $\operatorname{tr}(A) = 0$ 且 $|\mathbb{F}| > n$, 则存在可逆方阵 $X, Y \in \mathbb{F}^{n \times n}$, 使得 $A = XY - YX$.

证明 (1) 对 n 归纳. 设 $A = (a_{ij})$. 当 $n = 2$ 时,

- 若 $a_{12} \neq 0$, 则 $\begin{pmatrix} 1 & 0 \\ -x & 1 \end{pmatrix} A \begin{pmatrix} 1 & 0 \\ x & 1 \end{pmatrix} = \begin{pmatrix} a_{11} + xa_{12} & * \\ * & * \end{pmatrix}$, 存在 x 使 $a_{11} + xa_{12} = b_1$.
- 若 $a_{21} \neq 0$, 则 $\begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} A \begin{pmatrix} 1 & -x \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} a_{11} + xa_{21} & * \\ * & * \end{pmatrix}$, 存在 x 使 $a_{11} + xa_{21} = b_1$.
- 若 $a_{12} = a_{21} = 0$, 则 $a_{11} \neq a_{22}$, $\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} A \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} a_{11} & 0 \\ a_{11} - a_{22} & a_{22} \end{pmatrix}$, 化为前述情形.

当 $n \geq 3$ 时, A 有 $n-1$ 阶主子矩阵 B 不是纯量方阵. 对 B 应用归纳假设, A 可相似于 $\begin{pmatrix} b_1 & \alpha \\ \beta & C \end{pmatrix}$ 形式. 若 C 不是纯量方阵, 对 C 应用归纳假设, 结论成立. 下设 C 是纯量方阵.

- 若 $\alpha \neq 0$, 则 $\begin{pmatrix} 1 & \\ \gamma & I \end{pmatrix} \begin{pmatrix} b_1 & \alpha \\ \beta & C \end{pmatrix} \begin{pmatrix} 1 & \\ -\gamma & I \end{pmatrix} = \begin{pmatrix} b_1 - \alpha\gamma & * \\ * & C + \gamma\alpha \end{pmatrix}$, 存在 $\gamma \neq 0$ 使 $\alpha\gamma = 0$, $C + \gamma\alpha$ 不是纯量方阵. 根据归纳假设, 结论成立.
- 若 $\beta \neq 0$, 则 $\begin{pmatrix} 1 & -\gamma \\ & I \end{pmatrix} \begin{pmatrix} b_1 & \alpha \\ \beta & C \end{pmatrix} \begin{pmatrix} 1 & \gamma \\ & I \end{pmatrix} = \begin{pmatrix} b_1 - \gamma\beta & * \\ * & C + \beta\gamma \end{pmatrix}$, 存在 $\gamma \neq 0$ 使 $\gamma\beta = 0$, $C + \beta\gamma$ 不是纯量方阵. 根据归纳假设, 结论成立.

(2) 由 (1) 及 $|\mathbb{F}| > n$ 知, 存在可逆方阵 $P \in \mathbb{F}^{n \times n}$ 使得 $B = (b_{ij}) := P^{-1}AP$ 的前 $n-1$ 个对角元素均为 0, 再由 $\text{tr}(A) = 0$ 知 B 的对角线元素全为 0. 设 $X = \text{diag}(1, \dots, n)$, $Y = (y_{ij})$, 其中

$$y_{ij} = \begin{cases} \frac{b_{ij}}{i-j}, & \text{若 } i \neq j, \\ 1, & \text{若 } i = j. \end{cases}$$

则 $XY - YX = ((i-j)y_{ij}) = B = P^{-1}AP$, 从而 A 是 PXP^{-1} 和 PYP^{-1} 的换位子. \square

习题 5.1.10 设 $A, B \in \mathbb{F}^{n \times n}$ 满足 $AB = O$, $M_1 = \text{diag}(A+B, O)$, $M_2 = \text{diag}(A, B)$. 证明: $\varphi_{M_1} = \varphi_{M_2}$. 并举例说明 M_1 与 M_2 有可能不相似.

证明

$$\begin{aligned} \varphi_{M_1} &= \det \left(xI_{2n} - \begin{pmatrix} A+B & \\ & O \end{pmatrix} \right) = \det \begin{pmatrix} xI_n - (A+B) & \\ & xI_n \end{pmatrix} = x^n \det(xI_n - (A+B)), \\ \varphi_{M_2} &= \det \begin{pmatrix} xI_n - A & \\ & xI_n - B \end{pmatrix} = \det(xI_n - A) \det(xI_n - B) = \det(x^2 I_n - x(A+B)) \\ &= x^n \det(xI_n - (A+B)) = \varphi_{M_1}. \end{aligned}$$

M_1 与 M_2 不相似的例子: 取 $A = E_{1n} \in \mathbb{F}^{n \times n}$, 则 $A^2 = O$. 再取 $B = -A$. 则 $\text{rank}(M_1) = 0$, 而 $\text{rank}(M_2) = 2$. M_1 与 M_2 不相似. \square

注 2.0.9 本题结论的等价表述: $\varphi_A(x)\varphi_B(x) = x^n\varphi_{A+B}(x)$.

习题 5.1.11 设 $A, B \in \mathbb{F}^{n \times n}$ 满足 $\text{rank}(A) = \text{rank}(B) = 1$. 证明:

- (1) $\varphi_A(x) = x^n - \text{tr}(A)x^{n-1}$.
- (2) $\text{tr}(A) \neq 0$ 是 A 可以相似于对角方阵的充分必要条件.
- (3) $\text{tr}(A) = \text{tr}(B)$ 是 A 与 B 相似的充分必要条件.

证明 (1) 设 $A = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \begin{pmatrix} b_1 & \cdots & b_n \end{pmatrix}$ 是 A 的满秩分解. 由例 (3.12) 结论,

$$\varphi_A(x) = x^{n-1} \left(x - \sum_{i=1}^n a_i b_i \right) = x^n - \text{tr}(A)x^{n-1}.$$

(2) 由 (1) 知 A 的特征值为 $n-1$ 个 0 加上 $\text{tr}(A)$. ① 必要性: 若 A 可相似成对角方阵且 $\text{tr}(A) = 0$, 则 $\text{rank}(A) = \text{rank}(O) = 0$, 矛盾. ② 充分性: 当 $\text{tr}(A) \neq 0$ 时, A 的特征值 $\text{tr}(A)$ 的几何重数 $m_1 \geq 1$, 特征值 0 的几何重数 $m_2 = n - \text{rank}(0I - A) = n - \text{rank}(A) = n - 1$, 故 $m_1 + m_2 \geq n$, 说明 A 可相似对角化.

(3) 只需证充分性: ① 若 $\text{tr}(A) = \text{tr}(B) \neq 0$, 由 (2) 知 A 和 B 分别可相似对角化为 $\text{diag}(0, \dots, 0, \text{tr}(A))$ 和 $\text{diag}(0, \dots, 0, \text{tr}(B))$. ② 若 $\text{tr}(A) = \text{tr}(B) = 0$, 由 $\text{rank}(A) = 1$ 知存在线性无关的 $\alpha_1, \dots, \alpha_{n-1} \in \{x \mid Ax = 0\}$. 将该向量组扩充为可逆方阵 $P = (\alpha_1 \cdots \alpha_{n-1} \quad \alpha_n)$,

则结合 $\text{tr}(A) = 0$ 可知 $AP = \begin{pmatrix} 0 & \cdots & 0 & * \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & * \end{pmatrix}$, 故可设

$$AP = P \left(\begin{array}{c|c} O & \beta \\ \hline 0 \cdots \cdots 0 & 0 \end{array} \right), \quad \text{即} \quad P^{-1}AP = \left(\begin{array}{c|c} O & \beta \\ \hline 0 \cdots \cdots 0 & 0 \end{array} \right).$$

由 $\text{rank}(A) = 1$ 可知 $\beta \neq \mathbf{0}$, 因此存在 $n-1$ 阶可逆方阵 Q 使 $Q\beta = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$. 于是

$$\underbrace{\left(\begin{array}{c|c} Q & \\ \hline & 1 \end{array} \right)}_{\tilde{Q}} \underbrace{\left(\begin{array}{c|c} O & \beta \\ \hline 0 \cdots \cdots 0 & 0 \end{array} \right)}_{\tilde{Q}^{-1}} \underbrace{\left(\begin{array}{c|c} Q^{-1} & \\ \hline & 1 \end{array} \right)}_{\tilde{Q}^{-1}} = \left(\begin{array}{c|c} O & Q\beta \\ \hline 0 \cdots \cdots 0 & 0 \end{array} \right),$$

也即 $\tilde{Q}P^{-1}AP\tilde{Q}^{-1} = E_{1n}$. 同理 B 也可相似为 E_{1n} . 由相似是等价关系知 A 与 B 相似. \square

习题 5.1.12 (2) 设 $A, B \in \mathbb{F}^{n \times n}$. 证明: 若 $\text{rank}(ABA) = \text{rank}(A)$, 则 AB 与 BA 相似.

证明 设 $A = P \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} Q$, 对应于分块地设 $B = Q^{-1} \begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix} P^{-1}$. 则

$$\text{rank}(ABA) = \text{rank}(A) \implies \text{rank} \left(\begin{pmatrix} I_r & O \\ O & O \end{pmatrix} QBP \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} \right) = \text{rank} \begin{pmatrix} B_1 & O \\ O & O \end{pmatrix} = r.$$

因此 B_1 是可逆方阵. 另外有

$$AB = P \begin{pmatrix} B_1 & B_2 \\ O & O \end{pmatrix} P^{-1} \underset{\sim}{\sim} \begin{pmatrix} B_1 & B_2 \\ O & O \end{pmatrix},$$

$$BA = Q^{-1} \begin{pmatrix} B_1 & O \\ B_3 & O \end{pmatrix} Q \underset{\sim}{\sim} \begin{pmatrix} B_1 & O \\ B_3 & O \end{pmatrix}.$$

下面通过待定系数法证明以上两式右端都相似于 $\begin{pmatrix} B_1 & O \\ O & O \end{pmatrix}$. 设

$$\begin{pmatrix} B_1 & B_2 \\ O & O \end{pmatrix} = \begin{pmatrix} I & X \\ & I \end{pmatrix} \begin{pmatrix} B_1 & O \\ O & O \end{pmatrix} \begin{pmatrix} I & -X \\ & I \end{pmatrix} = \begin{pmatrix} B_1 & -B_1X \\ O & O \end{pmatrix},$$

$$\begin{pmatrix} B_1 & O \\ B_3 & O \end{pmatrix} = \begin{pmatrix} I & \\ Y & I \end{pmatrix} \begin{pmatrix} B_1 & O \\ O & O \end{pmatrix} \begin{pmatrix} I & \\ -Y & I \end{pmatrix} = \begin{pmatrix} B_1 & O \\ YB_1 & O \end{pmatrix}.$$

观察可知应取 $X = -B_1^{-1}B_2$, $Y = B_3B_1^{-1}$. 这就证明了 AB 与 BA 相似.

【也可利用定理 5.8, 由 $\text{Spec}(B_1) \cap \{0\} = \emptyset$ 直接得到 AB 与 BA 都与 $\begin{pmatrix} B_1 & O \\ O & O \end{pmatrix}$ 相似.】 \square

习题 5.1.12 (3) 设 $A, B \in \mathbb{F}^{n \times n}$. 证明: 若 $\text{rank}(ABA) = \text{rank}(B)$, 则 AB 与 BA 相似.

证明 设 $A = P \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} Q$, 对应于分块地设 $B = Q^{-1} \begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix} P^{-1}$. 则

$$\text{rank}(ABA) = \text{rank}(B) \implies \text{rank} \left(\begin{pmatrix} I_r & O \\ O & O \end{pmatrix} QBP \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} \right) = \text{rank}(B_1) = \text{rank}(B).$$

由例 (4.6) 结论知 $\text{rank}(B) = \text{rank}(B_1) \iff$ 存在矩阵 X, Y , 使得

$$B_2 = B_1X, \quad B_3 = YB_1, \quad B_4 = YB_1X.$$

于是

$$AB = P \begin{pmatrix} B_1 & B_1X \\ O & O \end{pmatrix} P^{-1}, \quad BA = Q^{-1} \begin{pmatrix} B_1 & O \\ YB_1 & O \end{pmatrix} Q.$$

由

$$\begin{pmatrix} I & X \\ & I \end{pmatrix} \begin{pmatrix} B_1 & B_1X \\ O & O \end{pmatrix} \begin{pmatrix} I & -X \\ & I \end{pmatrix} = \begin{pmatrix} B_1 & O \\ O & O \end{pmatrix}$$

与

$$\begin{pmatrix} I & \\ -Y & I \end{pmatrix} \begin{pmatrix} B_1 & O \\ YB_1 & O \end{pmatrix} \begin{pmatrix} I & \\ Y & I \end{pmatrix} = \begin{pmatrix} B_1 & O \\ O & O \end{pmatrix}$$

可知 AB 与 BA 均相似于 $\begin{pmatrix} B_1 & O \\ O & O \end{pmatrix}$. \square

习题 5.1.12 (4) 设 $A, B \in \mathbb{F}^{n \times n}$. 举例: $\text{rank}(AB) = \text{rank}(BA)$, AB 与 BA 不相似.

$$\text{解 } A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \quad AB = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad BA = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

假设矩阵 P 满足 $PAB = BAP$, 计算易知 P 必有零行, 因此不存在可逆方阵 P , 使得 $PAB = BAP$, 即 AB 与 BA 不相似. \square

习题 5.1.13 设 $A = (a_{ij}) \in \mathbb{C}^{n \times n}$, $\varphi_A(x) = \sum_{k=0}^n c_k x^k$, λ 是 A 的任意特征值. 证明:

$$(1) |\lambda| \leq 1 + \max_{0 \leq k \leq n-1} |c_k|.$$

$$(2) |\lambda| \leq 2 \max_{0 \leq k \leq n-1} |c_k|^{\frac{1}{n-k}}.$$

$$(3) \text{ (Gershgorin 圆盘定理) 存在 } k \text{ 使得 } |\lambda - a_{kk}| \leq \sum_{j \neq k} |a_{kj}|.$$

证明 (1) 不妨设 $|\lambda| > 1$. 记 $c = \max_{0 \leq k \leq n-1} |c_k|$. 由定理 5.2 知 $\varphi_A(x)$ 为 n 次首一多项式, 即 $c_n = 1$. 于是

$$\varphi_A(\lambda) = 0 \implies \lambda^n = - \sum_{k=0}^{n-1} c_k \lambda^k.$$

由三角不等式可得

$$|\lambda|^n = \left| \sum_{k=0}^{n-1} c_k \lambda^k \right| \leq \sum_{k=0}^{n-1} |c_k| |\lambda|^k \leq c \sum_{k=0}^{n-1} |\lambda|^k = c \frac{|\lambda|^n - 1}{|\lambda| - 1} \leq \frac{c |\lambda|^n}{|\lambda| - 1}.$$

整理即得 $|\lambda| \leq 1 + \max_{0 \leq k \leq n-1} |c_k|$.

(2) 不妨设 $\lambda \neq 0$. 记 $c = \max_{0 \leq k \leq n-1} |c_k|^{\frac{1}{n-k}}$, 则 $|c_k| \leq c^{n-k}$. 用反证法, 假设 $\frac{c}{|\lambda|} < \frac{1}{2}$, 则由

$$|\lambda|^n = \left| \sum_{k=0}^{n-1} c_k \lambda^k \right| \text{ 可得}$$

$$1 = \left| \sum_{k=0}^{n-1} \frac{c_k}{\lambda^{n-k}} \right| \leq \sum_{k=0}^{n-1} \frac{|c_k|}{|\lambda|^{n-k}} \leq \sum_{k=0}^{n-1} \left(\frac{c}{|\lambda|} \right)^{n-k} = \sum_{k=1}^n \left(\frac{c}{|\lambda|} \right)^k < \sum_{k=1}^n \frac{1}{2^k} < 1,$$

矛盾. 故原不等式得证.

(3) 设 $\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ 是 A 的与 λ 对应的特征向量. 不妨设 $n \geq 2$. 则

$$A\mathbf{x} = \lambda\mathbf{x} \iff \lambda x_i = \sum_{j=1}^n a_{ij} x_j \iff (\lambda - a_{ii})x_i = \sum_{j \neq i} a_{ij} x_j, \quad i = 1, \dots, n.$$

设 $k \in \{1, \dots, n\}$ 是任意一个满足 $|x_k| = \|\mathbf{x}\|_\infty$ 的指标, 由 $\mathbf{x} \neq \mathbf{0}$ 可知 $|x_k| \neq 0$. 于是

$$(\lambda - a_{kk})x_k = \sum_{j \neq k} a_{kj} x_j \iff \lambda - a_{kk} = \sum_{j \neq k} a_{kj} \frac{x_j}{x_k}.$$

由三角不等式可得

$$|\lambda - a_{kk}| = \left| \sum_{j \neq k} a_{kj} \frac{x_j}{x_k} \right| \leq \sum_{j \neq k} |a_{kj}| \frac{|x_j|}{\|\mathbf{x}\|_\infty} \leq \sum_{j \neq k} |a_{kj}|.$$

□

习题 5.2.2 设 $\lambda_1, \lambda_2, \dots, \lambda_n$ 是 n 阶方阵 $A = (a_{ij})$ 的所有特征值. 证明: $\sum_{i=1}^n \lambda_i^2 = \sum_{i,j=1}^n a_{ij}a_{ji}$.

证明 因为 $\lambda_1^2, \lambda_2^2, \dots, \lambda_n^2$ 是 A^2 的所有特征值, 所以

$$\sum_{i=1}^n \lambda_i^2 = \operatorname{tr}(A^2) = \sum_{i=1}^n \sum_{j=1}^n a_{ij}a_{ji}.$$

□

习题 5.2.3 设正整数 $m \geq 2$, $\omega = \cos \frac{2\pi}{m} + i \sin \frac{2\pi}{m}$. 证明:

(1) 对于任意 $f \in \mathbb{C}[x]$, 存在 $g \in \mathbb{C}[x]$ 使得 $\prod_{k=1}^m f(\omega^k x) = g(x^m)$.

(2) 对于任意 $A \in \mathbb{C}^{n \times n}$, $B = A^m$ 的特征多项式 $\varphi_B(x) = (-1)^{(m-1)n} \prod_{k=1}^m \varphi_A(\omega^k x^{\frac{1}{m}})$.

证明 (1) 设 $\prod_{k=1}^m f(\omega^k x) = \sum_{k=0}^n a_k x^k$. 注意到 $\prod_{k=1}^m f(\omega^k \omega x) = \prod_{k=1}^m f(\omega^k x)$, 因此

$$\sum_{k=0}^n a_k x^k = \sum_{k=0}^n a_k (\omega x)^k.$$

由于 $\omega = \cos \frac{2\pi}{m} + i \sin \frac{2\pi}{m}$, 对任意满足 $m \nmid k$ 的 k , $\omega^k \neq 1$, 此时从 $a_k(1 - \omega^k) = 0$ 推出 $a_k = 0$. 故 $\prod_{k=1}^m f(\omega^k x)$ 一定可写成 $g(x^m)$ 的形式.

(2) 由 $x^m - a^m = -x^m \left[\left(\frac{a}{x} \right)^m - 1 \right] = -x^m \prod_{k=1}^m \left(\frac{a}{x} - \omega^k \right) = -\prod_{k=1}^m (a - \omega^k x)$ 以及 A 与 I_n 乘积可交换得

$$x^m I_n - A^m = -\prod_{k=1}^m (A - \omega^k x I_n).$$

因此

$$\begin{aligned} \varphi_B(x^m) &= \det(x^m I_n - A^m) = \det \left(-\prod_{k=1}^m (A - \omega^k x I_n) \right) = (-1)^n \prod_{k=1}^m \det(A - \omega^k x I_n) \\ &= (-1)^n \prod_{k=1}^m (-1)^n \varphi_A(\omega^k x) = (-1)^{(m+1)n} \prod_{k=1}^m \varphi_A(\omega^k x) = (-1)^{(m-1)n} \prod_{k=1}^m \varphi_A(\omega^k x). \end{aligned}$$

把上式中的 x 用 $x^{\frac{1}{m}}$ 代入即得 $\varphi_B(x) = (-1)^{(m-1)n} \prod_{k=1}^m \varphi_A(\omega^k x^{\frac{1}{m}})$. □

习题 5.2.9 设 A 是可逆复数方阵, 并且存在正整数 $i < j$, 使得 A^i 与 A^j 相似. 证明: 存在正整数 k , 使得 A^k 的特征值都是 1.

证明 设 $\lambda_1, \dots, \lambda_n$ 是 $A \in \mathbb{C}^{n \times n}$ 的所有特征值. 注意到从 $A^i \stackrel{\text{相似}}{\sim} A^j$ 可推出

$$A^{is} \stackrel{\text{相似}}{\sim} A^{i^{s-1}j} \stackrel{\text{相似}}{\sim} \dots \stackrel{\text{相似}}{\sim} A^{ij^{s-1}} \stackrel{\text{相似}}{\sim} A^{js}, \quad \forall s \geq 1.$$

由 $i \neq j$ 可知 $i^s, i^{s-1}j, \dots, ij^{s-1}, j^s$ 这 $s+1$ 个数两两不同, 于是当 $s+1 > n!$ 时, 一定存在 $\alpha < \beta$, 使得 A^α 的特征值 $\lambda_1^\alpha, \dots, \lambda_n^\alpha$ 与 A^β 的特征值 $\lambda_1^\beta, \dots, \lambda_n^\beta$ 满足 $\lambda_i^\alpha = \lambda_i^\beta$ ($i = 1, \dots, n$). 取 $k = \beta - \alpha$. 由 A 可逆知 $\lambda_i \neq 0$, 因此 $\lambda_i^k = 1$ ($i = 1, \dots, n$), 也即 $\text{Spec}(A^k) = \{1\}$. \square

习题 5.2.10 设 I 是指标集合, $\{A_i \mid i \in I\}$ 是一组两两乘积可交换的复数方阵.

- (1) 证明: 存在复数向量 α 是所有 A_i 的公共特征向量.
- (2) 证明: 存在可逆复数方阵 P 使得 $P^{-1}A_iP$ 都是上三角方阵.
- (3) 对实数域 \mathbb{R} 和任意数域 \mathbb{F} 上的方阵, 推广上述结论.

解 (1) ① 先证明乘积可交换的两个复数方阵的情形. 不妨设 $I = \{1, 2\}$, $A_1A_2 = A_2A_1$. 设 λ 作为 A_1 的特征值的几何重数是 r . 令 $\mathbf{x}_1, \dots, \mathbf{x}_r$ 是 A 的与 λ 相伴的特征空间的一组基, 设 $X = (\mathbf{x}_1 \ \dots \ \mathbf{x}_r)$. 对每一个 $j = 1, \dots, r$, 有

$$A_1(A_2\mathbf{x}_j) = A_1A_2\mathbf{x}_j = A_2A_1\mathbf{x}_j = A_2\lambda\mathbf{x}_j = \lambda(A_2\mathbf{x}_j).$$

由此可见 $A_2\mathbf{x}_j$ 也在 A 的与 λ 相伴的特征空间中, 因此 $A_2\mathbf{x}_j$ 是 X 的列向量的线性组合, 从而存在 $C_2 \in \mathbb{C}^{r \times r}$, 使得 $A_2X = XC_2$. 令 (μ, \mathbf{u}) 是 C_2 的一个特征对. 由 X 的列向量组线性无关及 $\mathbf{u} \neq \mathbf{0}$ 可知 $X\mathbf{u} \neq \mathbf{0}$. 因为

$$A_2(X\mathbf{u}) = A_2X\mathbf{u} = XC_2\mathbf{u} = X\mu\mathbf{u} = \mu(X\mathbf{u}),$$

同时 $X\mathbf{u}$ 也在 A 的与 λ 相伴的特征空间中, 即

$$A_1(X\mathbf{u}) = \lambda(X\mathbf{u}),$$

所以 $\alpha = X\mathbf{u}$ 就是 A_1 与 A_2 的公共特征向量.

② 假设对某个 $m \geq 2$, 命题已经对由不超过 m 个矩阵组成的交换族成立. 不妨设 $I = \{1, \dots, m, m+1\}$. 设 λ 是 A_1 的一个特征值. 令 $\mathbf{x}_1, \dots, \mathbf{x}_r$ 是 A 的与 λ 相伴的特征空间的一组基, 设 $X = (\mathbf{x}_1 \ \dots \ \mathbf{x}_r)$. 同 ① 可得, 对任意 $j \in \{2, \dots, m+1\}$, 存在 $C_j \in \mathbb{C}^{r \times r}$, 使得 $A_jX = XC_j$. 于是由

$$A_iA_jX = A_iXC_j = XC_iC_j \quad \text{与} \quad A_jA_iX = A_jXC_i = XC_jC_i$$

可知 $X(C_i C_j - C_j C_i) = O$, 而 X 列满秩, 故 $C_i C_j = C_j C_i$, 即 C_2, \dots, C_{m+1} 是交换族. 由归纳假设, 它们有公共特征向量 \mathbf{u} . 同 ① 可知 $\alpha = X\mathbf{u}$ 就是 A_1, \dots, A_{m+1} 的一个公共特征向量.

(2) 对 A_i 的大小归纳.

① 当 $A_i \in \mathbb{C}^{1 \times 1}$ 时命题自动成立.

② 设对某个 $n \geq 1$, 命题已对大小不超过 $n \times n$ 的矩阵组成的交换族成立. 设 α 是所有 A_i 的公共特征向量, 且 $A\alpha = \lambda_i \alpha$. 因为 $\alpha \neq \mathbf{0}$, 所以存在可逆方阵 P 以 α 为第一列. 于是

$$A_i P = A_i \begin{pmatrix} \alpha & * \end{pmatrix} = \begin{pmatrix} \alpha & * \end{pmatrix} \begin{pmatrix} \lambda_i & * \\ & B_i \end{pmatrix}, \quad \forall i \in I.$$

由 $A_i = P \begin{pmatrix} \lambda_i & * \\ & B_i \end{pmatrix} P^{-1}$ 及乘积可交换的性质得

$$A_i A_j = P \begin{pmatrix} \lambda_i & * \\ & B_i \end{pmatrix} \begin{pmatrix} \lambda_j & * \\ & B_j \end{pmatrix} P^{-1} = P \begin{pmatrix} \lambda_j & * \\ & B_j \end{pmatrix} \begin{pmatrix} \lambda_i & * \\ & B_i \end{pmatrix} P^{-1} = A_j A_i,$$

由此可得 $\begin{pmatrix} \lambda_i & * \\ & B_i \end{pmatrix} \begin{pmatrix} \lambda_j & * \\ & B_j \end{pmatrix} = \begin{pmatrix} \lambda_j & * \\ & B_j \end{pmatrix} \begin{pmatrix} \lambda_i & * \\ & B_i \end{pmatrix}$, 进而得到 $B_i B_j = B_j B_i$. 对 B_i 运用归纳假设即可.

(3) ① 实数域 \mathbb{R} 上的交换族里的矩阵在 \mathbb{C} 上有公共特征向量.

② 任意数域 \mathbb{F} 上的交换族里的矩阵在 \mathbb{F} 的代数闭包上有公共特征向量. □

习题 5.3.9 证明: 任意复数方阵 A 可以唯一地表示为 $A = B + C$ 的形式, 其中 B 可以相似于对角方阵, C 是幂零方阵, 并且 B, C 都可以表示为 A 的多项式. 这种表示形式称为 A 的 Jordan–Chevalley 分解.

证明 ① 先对 A 的 Jordan 标准形证明结论. 设 A 的全体不同特征值为 $\lambda_1, \dots, \lambda_s$ 且 $J = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_s \end{pmatrix}$, 其中 J_i 是 A 以 λ_i 为特征值的所有 Jordan 块的直和, 其阶数为 λ_i 的代数重数 n_i . 则对每个 i 都有 $J_i = M_i + N_i$, 其中 $M_i = \lambda_i I$ 是对角方阵, N_i 是幂零方阵, 且 $M_i N_i = N_i M_i$. 设

$$M = \begin{pmatrix} M_1 & & \\ & \ddots & \\ & & M_s \end{pmatrix}, \quad N = \begin{pmatrix} N_1 & & \\ & \ddots & \\ & & N_s \end{pmatrix},$$

则 $J = M + N$, 其中 M 是对角方阵, N 是幂零方阵, 且 $MN = NM$. 因为 $\lambda_1, \dots, \lambda_s$ 两两不同, 所以多项式 $(x - \lambda_1)^{n_1}, \dots, (x - \lambda_s)^{n_s}$ 两两互素. 由中国剩余定理, 存在多项式 g 满足

$$g(x) = h_i(x)(x - \lambda_i)^{n_i} + \lambda_i \quad (i = 1, \dots, s, h_i \in \mathbb{C}[x]).$$

因为 $g(J_i) = h_i(J_i)(J_i - \lambda_i I)^{n_i} + \lambda_i I = \lambda_i I = M_i$, 所以

$$g(J) = \begin{pmatrix} g(J_1) & & \\ & \ddots & \\ & & g(J_s) \end{pmatrix} = \begin{pmatrix} M_1 & & \\ & \ddots & \\ & & M_s \end{pmatrix} = M.$$

从而 $N = J - M = J - g(J)$ 也是 J 的多项式.

② 再考虑一般情形. 设 $P^{-1}AP = J$, 则 $A = PJP^{-1} = P(M + N)P^{-1}$. 令 $B = PMP^{-1}$, $C = PNP^{-1}$, 则 B 可以相似于对角方阵, C 是幂零方阵. 又

$$g(A) = g(PJP^{-1}) = Pg(J)P^{-1} = PMP^{-1} = B,$$

从而 $C = A - g(A)$ 也是 A 的多项式.

③ 下证上述分解的唯一性. 假设 $A = B + C = B_1 + C_1$ 是两种满足要求的分解. 由 B_1 与 C_1 乘积可交换及 C_1 可表成 A 的多项式可得 $AB_1 = B_1A$ (对非零多项式的次数归纳可得), 同理 $AC_1 = C_1A$. 由 A 与 B_1 乘积可交换及 $B = g(A)$ 可得 $BB_1 = B_1B$, 同理 $CC_1 = C_1C$. 设 C 与 C_1 的幂零指数分别为 r, t , 则 $(C_1 - C)^{r+t} = O$. 于是

$$B - B_1 = C_1 - C \implies (B - B_1)^{r+t} = (C_1 - C)^{r+t} = O.$$

引理 设 $A, B \in \mathbb{F}^{n \times n}$, 且 $AB = BA$. 若 A, B 都可对角化, 则存在可逆方阵 $Q \in \mathbb{F}^{n \times n}$, 使得 $Q^{-1}AQ$ 与 $Q^{-1}BQ$ 同为对角方阵.

【引理的证明: 因为 A 可对角化, 所以存在可逆方阵 P , 使 $P^{-1}AP = \begin{pmatrix} \lambda_1 I & & \\ & \ddots & \\ & & \lambda_s I \end{pmatrix}$, 其中

$\lambda_1, \dots, \lambda_s$ 是 A 的全体两两不同的特征值. 由 $AB = BA$ 可得

$$(P^{-1}AP)(P^{-1}BP) = (P^{-1}BP)(P^{-1}AP).$$

对 $P^{-1}BP$ 作相应分块 $P^{-1}BP = (B_{ij})_{s \times s}$, 代入上式可得 $\lambda_i IB_{ij} = B_{ij} \lambda_j I$ 即 $(\lambda_i - \lambda_j)B_{ij} =$

O . 因此当 $i \neq j$ 时 $B_{ij} = O$, 即 $P^{-1}BP = \begin{pmatrix} B_{11} & & \\ & \ddots & \\ & & B_{ss} \end{pmatrix}$. 因为 B 可对角化, 所以 $P^{-1}BP$

也可对角化, 从而每个分块 B_{ii} 都可对角化, 即存在可逆方阵 P_i , 使 $P_i^{-1}B_{ii}P_i$ 为对角方阵 ($i = 1, \dots, s$). 因此

$$\begin{pmatrix} P_1^{-1} & & \\ & \ddots & \\ & & P_s^{-1} \end{pmatrix} P^{-1}BP \begin{pmatrix} P_1 & & \\ & \ddots & \\ & & P_s \end{pmatrix}$$

为对角方阵. 同时,

$$\begin{pmatrix} P_1^{-1} & & \\ & \ddots & \\ & & P_s^{-1} \end{pmatrix} P^{-1}AP \begin{pmatrix} P_1 & & \\ & \ddots & \\ & & P_s \end{pmatrix} = \begin{pmatrix} \lambda_1 I & & \\ & \ddots & \\ & & \lambda_s I \end{pmatrix}.$$

故 $Q = P \begin{pmatrix} P_1 & & \\ & \ddots & \\ & & P_s \end{pmatrix}$ 为所需的可逆方阵.】

由以上引理, 设 Q 是使得 $Q^{-1}BQ$ 与 $Q^{-1}B_1Q$ 都是对角方阵的可逆方阵, 则

$$(Q^{-1}BQ - Q^{-1}B_1Q)^{r+t} = Q^{-1}(B - B_1)^{r+t}Q = O.$$

由 $Q^{-1}BQ - Q^{-1}B_1Q$ 是对角方阵可知它只能是零方阵, 进而 $B = B_1$, $C = C_1$. □

注 2.0.10 B, C 都可以表示为 A 的多项式 $\implies BC = CB$.

习题 5.3.11 设 $\lambda_1, \lambda_2, \dots, \lambda_n$ 是 n 阶复数方阵 A 的所有特征值. $\rho(A) = \max_{1 \leq i \leq n} |\lambda_i|$ 称为 A 的谱半径.

(1) 证明: $\lim_{k \rightarrow +\infty} A^k = O$ 的充分必要条件是 $\rho(A) < 1$.

(2) 设 r 是幂级数 $\sum_{k=0}^{\infty} c_k z^k$ 的收敛半径. 证明: 当 $\rho(A) < r$ 时, 矩阵幂级数 $\sum_{k=0}^{\infty} c_k A^k$ 收敛.

证明 (1) 不妨设 A 即为 Jordan 方阵.

$$\implies \lim_{k \rightarrow +\infty} A^k = O \implies \lim_{k \rightarrow +\infty} \lambda_i^k = 0 \implies |\lambda_i| < 1 \ (i = 1, \dots, n) \implies \rho(A) < 1.$$

\Leftarrow : 不妨只考虑 A 是一个 Jordan 方块 $J_n(\lambda) =: \lambda I + J$ 的情形. 则

$$\begin{aligned} \lim_{k \rightarrow +\infty} A^k &= \lim_{k \rightarrow +\infty} \left(\lambda I + \underbrace{\begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix}}_J \right)^k = \lim_{k \rightarrow +\infty} \sum_{i=0}^k \lambda^{k-i} \binom{k}{i} J^i \\ &= \lim_{k \rightarrow +\infty} \sum_{i=0}^{n-1} \lambda^{k-i} \binom{k}{i} J^i = O. \end{aligned}$$

(2) 不妨只考虑 A 是一个 Jordan 方块 $J_n(\lambda) =: \lambda I + J$ 的情形. 设 $f(z) = \sum_{k=0}^{\infty} c_k z^k$, 由

Cauchy-Hadamard 定理可见 $f'(z) = \sum_{k=1}^{\infty} k c_k z^{k-1}$ 的收敛半径也是 r . 则当 $\rho(A) < r$ 时,

$$f(A) = f(\lambda I + J) = \sum_{k=0}^{n-1} \frac{f^{(k)}(\lambda)}{k!} J^k = \begin{pmatrix} 0 & f(\lambda) & f'(\lambda) & \cdots & \frac{f^{(n-1)}(\lambda)}{(n-1)!} \\ & \ddots & \ddots & \ddots & \vdots \\ & & \ddots & f'(\lambda) & \\ & & & f(\lambda) & \\ & & & & 0 \end{pmatrix}.$$

□

习题 5.3.12 (1) 求所有满足 $e^X = I$ 的 $X \in \mathbb{R}^{n \times n}$.

(2) 对哪些 $A \in \mathbb{R}^{n \times n}$, 存在 $X \in \mathbb{R}^{n \times n}$, 使得 $e^X = A$?

(3) 对哪些 $A \in \mathbb{R}^{n \times n}$, 存在 $X \in \mathbb{C}^{n \times n}$, 使得 $e^X = A$?

证明 (1) 设 $X = P \begin{pmatrix} X_1 & & \\ & \ddots & \\ & & X_k \end{pmatrix} P^{-1}$, 其中 P 是可逆方阵, X_1, \dots, X_k 是 Jordan 块.

则 $e^X = P \begin{pmatrix} e^{X_1} & & \\ & \ddots & \\ & & e^{X_k} \end{pmatrix} P^{-1}$. 若 $e^X = I$, 则 $e^{X_i} = I$, $i = 1, \dots, k$. 下面考虑 $e^{J_n(\lambda)} =$

$$e^\lambda \begin{pmatrix} \frac{1}{0!} & \frac{1}{1!} & \frac{1}{2!} & \cdots & \frac{1}{(n-1)!} \\ & \ddots & \ddots & \ddots & \vdots \\ & & \ddots & \ddots & \frac{1}{2!} \\ & & & \ddots & \frac{1}{1!} \\ & & & & \frac{1}{0!} \end{pmatrix}, \text{ 当 } n \geq 2 \text{ 时, 由于副对角线的存在, } e^{J_n(\lambda)} \neq I, \text{ 故 } n = 1, \text{ 且}$$

$e^\lambda = 1$, 即 $\lambda = 2k\pi i$ ($k \in \mathbb{Z}$). 容易发现, 为了使 $X \in \mathbb{R}^{n \times n}$, $2k\pi i$ 必与 $-2k\pi i$ 成对出现, 且这时 (经过相似将它们调整到对角线上相邻位置后) 有

$$\begin{pmatrix} 1 & i \\ -\frac{i}{2} & \frac{3}{2} \end{pmatrix} \begin{pmatrix} 2k\pi i & \\ & -2k\pi i \end{pmatrix} \begin{pmatrix} \frac{3}{2} & -i \\ \frac{i}{2} & 1 \end{pmatrix} = \begin{pmatrix} & 2k\pi \\ -2k\pi & \end{pmatrix}.$$

故所求的 $X = P \operatorname{diag}(O, Y_1, \dots, Y_t) P^{-1}$, 其中 $Y_i = \begin{pmatrix} & 2k_i\pi \\ -2k_i\pi & \end{pmatrix}$ ($k_i \in \mathbb{Z}$), O 为 $n - 2t$ 阶零方阵, P 为任意 n 阶可逆方阵.

(2) 设 $e^X = A$ 有解 $X = P \begin{pmatrix} X_1 & & \\ & \ddots & \\ & & X_k \end{pmatrix} P^{-1} \in \mathbb{R}^{n \times n}$, 其中 P 是可逆方阵, X_1, \dots, X_k 是 Jordan 块. 则 $A = P \begin{pmatrix} e^{X_1} & & \\ & \ddots & \\ & & e^{X_k} \end{pmatrix} P^{-1}$. 由 (1) 中 $e^{J_n(\lambda)}$ 的表达式以及 $\det(e^{J_n(\lambda)}) = e^{\operatorname{tr}(J_n(\lambda))} > 0$ 可知

(3) 由 $\det(e^X) = e^{\operatorname{tr}(X)}$ 可知 $\det(A) \neq 0$, 即 A 须为可逆方阵. 反过来, 对于任一可逆方阵 A , **未完待续** □

习题 5.4.5 设 $f_1, f_2, \dots, f_k \in \mathbb{F}[x]$ 是两两互素的首一多项式, A_i 是 f_i 的友方阵. 证明: $f_1 f_2 \cdots f_k$ 的友方阵与 $\operatorname{diag}(A_1, A_2, \dots, A_k)$ 相似.

证明 不妨设 $k = 2$. 记 $B = \begin{pmatrix} A_1 & \\ & A_2 \end{pmatrix}$. 则 $d_B(x) = \operatorname{lcm}(d_{A_1}(x), d_{A_2}(x))$. 又 $d_{A_i}(x) = \varphi_{A_i}(x) = f_i(x)$, 所以 $d_B(x) = \operatorname{lcm}(f_1(x), f_2(x)) = f_1(x)f_2(x)$. 而由 $\deg(d_B) = \deg(f_1) + \deg(f_2)$ 也可知 $d_B = \varphi_B$, 因此 B 与 φ_B 即 $f_1 f_2$ 的友方阵相似. $k \geq 3$ 的情形归纳即得. □

习题 5.4.7 (1) 设 $A \in \mathbb{F}^{n \times n}$, $\alpha, \beta \in \mathbb{F}^{n \times 1}$ 满足 $\gcd(d_{A, \alpha}, d_{A, \beta}) = 1$. 证明: $d_{A, \alpha + \beta} = d_{A, \alpha} d_{A, \beta}$.

证明 设 f 是 A 在 \mathbb{F} 上关于 $\alpha + \beta$ 的任一化零多项式, 即 $f(A)(\alpha + \beta) = \mathbf{0}$, 那么 $f(A)\alpha = -f(A)\beta$. 两边同乘 $d_{A,\alpha}(A)$ 得到 $d_{A,\alpha}(A) \cdot f(A)\alpha = -d_{A,\alpha}(A) \cdot f(A)\beta = \mathbf{0}$. 这说明 $d_{A,\alpha} \cdot f$ 是 A 在 \mathbb{F} 上关于 β 的一个化零多项式 (注意以上 \cdot 均表示多项式乘积). 于是 $d_{A,\beta} \mid (d_{A,\alpha} \cdot f)$. 结合 $d_{A,\alpha}$ 与 $d_{A,\beta}$ 互素可知 $d_{A,\beta} \mid f$. 同理, $d_{A,\alpha} \mid f$. 再次运用 $d_{A,\alpha}$ 与 $d_{A,\beta}$ 互素可得 $d_{A,\alpha}d_{A,\beta} \mid f$. 令 $f = d_{A,\alpha+\beta}$, 就得到 $d_{A,\alpha}d_{A,\beta} \mid d_{A,\alpha+\beta}$. 另一方面, 由 $d_{A,\alpha}(A) \cdot d_{A,\beta}(A)(\alpha + \beta) = d_{A,\alpha}(A) \cdot d_{A,\beta}(A)\alpha + d_{A,\alpha}(A) \cdot d_{A,\beta}(A)\beta = \mathbf{0}$ 可得 $d_{A,\alpha+\beta} \mid d_{A,\alpha}d_{A,\beta}$. 综上, $d_{A,\alpha+\beta} = d_{A,\alpha}d_{A,\beta}$. \square

习题 5.4.8 设 Hessenberg 方阵 $A = (a_{ij}) \in \mathbb{F}^{n \times n}$ 满足 $\prod_{i=2}^n a_{i,i-1} \neq 0$. 证明: $d_A = \varphi_A$.

证明 注意到 $(\lambda I - A) \begin{bmatrix} 2, \dots, n \\ 1, \dots, n-1 \end{bmatrix}$ 是上三角方阵, 因此

$$\det((\lambda I - A) \begin{bmatrix} 2, \dots, n \\ 1, \dots, n-1 \end{bmatrix}) = (-1)^{n-1} \prod_{i=2}^n a_{i,i-1} \neq 0,$$

从而对任意 $\lambda \in \mathbb{F}$, 都有 $\text{rank}(\lambda I - A) \geq n-1$. 故当 $\lambda \in \text{Spec}(A)$ 时, 就有 $\text{rank}(\lambda I - A) = n-1$, 即以 λ 为特征值的 Jordan 块只有 1 个. 又因为 Jordan 块的最小多项式与特征多项式相等, 且 A 的最小多项式是其 Jordan 标准形中所有 Jordan 块最小多项式的最小公倍式, 所以 $d_A = \varphi_A$. \square

习题 5.4.9 设 $A = \begin{pmatrix} B & I & & \\ & B & \ddots & \\ & & \ddots & I \\ & & & B \end{pmatrix}$, 其中 $B \in \mathbb{F}^{n \times n}$ 满足 $d_B = \varphi_B$ 在 $\mathbb{F}[x]$ 中不可约. 证

明: $d_A = \varphi_A$.

证明 设如上的 A 是 $m \times m$ 分块. 设 f 是 A 的任一化零多项式, 即

$$f(A) = \begin{pmatrix} f(B) & f'(B) & \cdots & \frac{f^{(m-1)}(B)}{(m-1)!} \\ & \ddots & \ddots & \vdots \\ & & \ddots & f'(B) \\ & & & f(B) \end{pmatrix} = O \implies f^{(k)}(B) = O, \quad k = 0, 1, \dots, m-1.$$

因此 $d_B \mid \gcd(f, f', \dots, f^{(m-1)})$. 再令 $f = d_A$, 由 d_B 在 $\mathbb{F}[x]$ 中不可约知 $d_B^m \mid d_A$. 注意到 $\deg(d_B^m) = m \deg(d_B) = m \deg(\varphi_B) = mn$, 而后者即为 A 的阶数, 故 $d_A = d_B^m = \varphi_A$. \square

习题 5.4.10 证明: 对于任意 $A \in \mathbb{F}^{n \times n}$, 存在 $f \in \mathbb{F}[x]$, 使得 A 的伴随方阵 $A^* = f(A)$.

证明 ① 若 A 可逆, 设 A 的特征多项式 $\varphi_A(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$, 其中 $a_0 = (-1)^n \det(A) \neq 0$. 由 Cayley-Hamilton 定理,

$$\varphi_A(x) = 0 \implies A \left(-\frac{1}{a_0}A^{n-1} - \frac{a_{n-1}}{a_0}A^{n-2} - \cdots - \frac{a_1}{a_0}I \right) = I,$$

因此

$$A^{-1} = -\frac{1}{a_0}A^{n-1} - \frac{a_{n-1}}{a_0}A^{n-2} - \cdots - \frac{a_1}{a_0}I.$$

故

$$A^* = \det(A)A^{-1} = -\det(A) \left(\frac{1}{a_0}A^{n-1} + \frac{a_{n-1}}{a_0}A^{n-2} + \cdots + \frac{a_1}{a_0}I \right).$$

② 若 A 不可逆, 设 y 是变元, 则由①, 存在 $f \in \mathbb{F}[x]$, 使得 $(yI + A)^* = f(yI + A)$. 令 $y = 0$ 就得到 $A^* = f(A)$. \square

习题 5.4.11 设 $\text{char } \mathbb{F} = 0$, $A, B \in \mathbb{F}^{n \times n}$ 满足 $AB - BA = A$. 证明: A 是幂零方阵.

证明 ① 先证明对 $1 \leq k \leq n$, $\text{tr}(A^k) = 0$. 当 $k = 1$ 时显然成立. 当 $k \geq 2$ 时,

$$\text{tr}(A^k) = \text{tr}(A^{k-1}(AB - BA)) = \text{tr}(A^k B) - \text{tr}(A^{k-1}BA) = \text{tr}(A^k B) - \text{tr}(AA^{k-1}B) = 0.$$

② 再证明若对 $1 \leq k \leq n$, $\text{tr}(A^k) = 0$, 则 A 是幂零方阵. 不妨设 A 是 Jordan 标准形, A 的对角元为 $\lambda_1, \cdots, \lambda_n$. 由 $\text{tr}(A^k) = 0$ 得到方程组

$$\lambda_1^k + \cdots + \lambda_n^k = 0, \quad k = 1, \cdots, n.$$

若 A 不是幂零方阵, 则存在 $\lambda_i \neq 0$, 结合上述方程组以及 $\text{char } \mathbb{F} = 0$ 知至少有另一个 $\lambda_j \neq \lambda_i$ 满足 $\lambda_j \neq 0$. 故 (必要时进行重新编号后) 不妨设

$$\lambda_1^k + \cdots + \lambda_r^k = 0, \quad k = 1, \cdots, r,$$

其中 $\lambda_1, \cdots, \lambda_r$ 两两不同. 也即

$$\begin{pmatrix} \lambda_1 & \cdots & \lambda_r \\ \vdots & \ddots & \vdots \\ \lambda_1^r & \cdots & \lambda_r^r \end{pmatrix} \mathbf{x} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \text{ 有解 } \mathbf{x} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}.$$

而由 Vandermonde 行列式可知

$$\det \begin{pmatrix} \lambda_1 & \cdots & \lambda_r \\ \vdots & \ddots & \vdots \\ \lambda_1^r & \cdots & \lambda_r^r \end{pmatrix} = \prod_{k=1}^r \lambda_k \prod_{1 \leq i < j \leq r} (\lambda_j - \lambda_i) \neq 0,$$

于是上述方程组不可能有非零解, 矛盾. 故 A 是幂零方阵. \square

习题 5.5.1 举例: “ $\forall x \in \mathbb{F}, xI - A$ 与 $xI - B$ 相抵” 不是 “ A 与 B 相似” 的充分条件.

解 $A = \begin{pmatrix} 1 & 1 & & \\ & 1 & & \\ & & 1 & 1 \\ & & & 1 \end{pmatrix}, B = \begin{pmatrix} 1 & 1 & & \\ & 1 & 1 & \\ & & 1 & \\ & & & 1 \end{pmatrix}$. 由 Jordan 标准形可知 A 与 B 不相似. 而对于

$$xI - A = \begin{pmatrix} x-1 & -1 & & \\ & x-1 & & \\ & & x-1 & 1 \\ & & & x-1 \end{pmatrix}, B = \begin{pmatrix} x-1 & -1 & & \\ & x-1 & -1 & \\ & & x-1 & \\ & & & x-1 \end{pmatrix} : \textcircled{1} \text{ 当 } x = 1$$

时, 直接验证可知 $xI - A$ 与 $xI - B$ 相抵. $\textcircled{2}$ 当 $x \neq 1$ 时, 利用 $x-1 \neq 0$ 可由初等变换将这两个矩阵都相抵成 $\text{diag}(x-1, x-1, x-1, x-1)$. \square

习题 5.5.2 设 $A \in \mathbb{F}^{n \times n}$, $xI - A$ 与 $\text{diag}(f_1, f_2, \dots, f_n)$ 在 $\mathbb{F}[x]$ 上模相抵, f_i 是首一多项式. 证明:

- (1) 存在 $P \in \mathbb{F}[x]^{n \times n}$ 使得 $(xI - A)P = \lambda I$, 其中 $\lambda = \text{lcm}(f_1, f_2, \dots, f_n) \in \mathbb{F}[x]$.
- (2) $\lambda(A) = O$.
- (3) A 与 $\text{diag}(A_1, A_2, \dots, A_n)$ 相似, 其中 A_i 是 f_i 的友方阵.

证明 (1) 设 $Q(xI - A)R = \begin{pmatrix} f_1 & & \\ & \ddots & \\ & & f_n \end{pmatrix}$, 其中 $Q, R \in \mathbb{F}[x]^{n \times n}$ 是模方阵. 则

$$Q(xI - A)R \begin{pmatrix} \frac{\lambda}{f_1} & & \\ & \ddots & \\ & & \frac{\lambda}{f_n} \end{pmatrix} = \begin{pmatrix} f_1 & & \\ & \ddots & \\ & & f_n \end{pmatrix} \begin{pmatrix} \frac{\lambda}{f_1} & & \\ & \ddots & \\ & & \frac{\lambda}{f_n} \end{pmatrix} = \lambda I,$$

于是利用 λI 是纯量方阵就有

$$\lambda I = Q^{-1} \lambda I Q = (xI - A) \underbrace{R \begin{pmatrix} \frac{\lambda}{f_1} & & \\ & \ddots & \\ & & \frac{\lambda}{f_n} \end{pmatrix}}_P Q.$$

(2) 设 $(xI - A)P = \lambda I$, 且 $P = \sum_{i=0}^m P_i x^i$, $\lambda I = \sum_{i=0}^m \lambda_i x^i I$, 其中 $P_i \in \mathbb{F}^{n \times n}$, $\lambda_i \in \mathbb{F}$, m 是某个充分大的正整数. 则 $P_{i-1} - P_i A = \lambda_i I$. 由此得到

$$\lambda(A) = \sum_{i=0}^m \lambda_i A^i = \sum_{i=0}^m (P_{i-1} - P_i A) A^i = O \quad (\text{下标越界的 } P_i \text{ 视为零矩阵}).$$

(3) 只需证 $\begin{pmatrix} xI - A_1 & & \\ & \ddots & \\ & & xI - A_n \end{pmatrix}$ 与 $\text{diag}(f_1, \dots, f_n)$ 模相抵. 注意到 $\mathbf{1}$ 多项式的

友方阵是空矩阵, 又由例 (4.15), 对任意非 $\mathbf{1}$ 首一多项式 f_i , 其友方阵 A_i 的特征方阵 $xI - A_i$ 的 Smith 标准形为 $\text{diag}(I_{k-1}, f_i)$, 其中 $k = \deg(f_i)$. 然后经过主对角线元素重排就得证. \square

注 2.0.11 (2) 提供了证明定理 5.18 (2) 的另一途径.

习题 5.5.7 设 $A, B \in \mathbb{F}^{n \times n}$. 证明: 若 $\text{diag}(A, \dots, A)$ 与 $\text{diag}(B, \dots, B)$ 相似, 则 A 与 B 相似.

证明 设 $xI - A$ 的初等因子组为 $p_i^{m_{ij}}$ ($1 \leq i \leq k, 1 \leq m_{ij} \leq n_i$). 则 $xI - \underbrace{\text{diag}(A, \dots, A)}_{k \uparrow A}$ 的初等因子组为 “ k 份” $p_i^{m_{ij}}$ ($1 \leq i \leq k, 1 \leq m_{ij} \leq n_i$). 从而 $xI - \underbrace{\text{diag}(B, \dots, B)}_{k \uparrow B}$ 的初等因子组也为 “ k 份” $p_i^{m_{ij}}$ ($1 \leq i \leq k, 1 \leq m_{ij} \leq n_i$). 故 $xI - B$ 的初等因子组为 $p_i^{m_{ij}}$ ($1 \leq i \leq k, 1 \leq m_{ij} \leq n_i$), A 与 B 相似. \square

习题 6.1.2 设 $P \in \mathbb{R}^{n \times n}$ 满足 $\|P\alpha\| = \|\alpha\|, \forall \alpha \in \mathbb{R}^{n \times 1}$. 证明: P 是正交方阵.

证明 1 先证明保范数可推出保内积. 由题, 对任意 $\alpha, \beta \in \mathbb{R}^{n \times 1}$,

$$(\alpha + \beta)^\top P^\top P (\alpha + \beta) = (\alpha + \beta)^\top (\alpha + \beta),$$

展开化简并注意 $\alpha^\top P^\top P \beta = \beta^\top P^\top P \alpha$ 即得到

$$\alpha^\top P^\top P \beta = \alpha^\top \beta.$$

依次取 $\alpha = e_i, \beta = e_j$ ($i, j = 1, \dots, n$) 就得到 P 的列向量两两正交、长度为 1. \square

证明 2 已知条件可等价转化为

$$\alpha^\top (P^\top P - I) \alpha = 0, \quad \forall \alpha \in \mathbb{R}^{n \times 1}.$$

因为 $P^T P - I$ 是对称实方阵, 所以存在 n 阶正交方阵 Q , 使得 $P^T P - I = Q \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} Q^T$.

于是

$$\alpha^T Q \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} Q^T \alpha = 0, \quad \forall \alpha \in \mathbb{R}^{n \times 1}.$$

取 $\alpha = e_k Q$, 即 $\alpha^T Q = e_k^T$, 得 $\lambda_k = 0$, $(k = 1, \dots, n)$. 故 $P^T P = I$. \square

习题 6.1.4 设 $\alpha, \beta \in \mathbb{R}^{3 \times 1}$, $\alpha \times \beta$ 是向量积运算, P 是 3 阶可逆实数方阵.

(1) 证明: 若 $P \in \text{SO}(3, \mathbb{R})$, 则 $(P\alpha) \times (P\beta) = P(\alpha \times \beta)$, $\forall \alpha, \beta$.

(2) 设 $(P\alpha) \times (P\beta) = P(\alpha \times \beta)$, $\forall \alpha, \beta$. P 是否必为正交方阵?

解 (1) 不妨设 α 与 β 线性无关. 于是存在 $\gamma \in \mathbb{R}^{n \times 1}$, 使得 $(\alpha \ \beta \ \gamma)$ 可逆. 由例 (2.13),

$$(\alpha \ \beta \ \gamma)^{-1} = \frac{1}{\det(\alpha \ \beta \ \gamma)} \begin{pmatrix} * & & \\ & * & \\ (\alpha \times \beta)^T & & \end{pmatrix},$$

$$(\alpha \ \beta \ \gamma)^{-1} P^{-1} = (P(\alpha \ \beta \ \gamma))^{-1} = \frac{1}{\det(P) \det(\alpha \ \beta \ \gamma)} \begin{pmatrix} * & & \\ & * & \\ ((P\alpha) \times (P\beta))^T & & \end{pmatrix}.$$

将第一式两边右乘 P^{-1} 并与第二式对比就得到

$$(\alpha \times \beta)^T P^{-1} = ((P\alpha) \times (P\beta))^T \iff (P\alpha) \times (P\beta) = P(\alpha \times \beta).$$

(2) 同 (1) 中推导可得 $PP^T = \det(P)I$. 两边求行列式得 $(\det(P))^2 = (\det(P))^3$. 故 $\det(P) = 1$, $PP^T = I$, P 是正交方阵. \square

习题 6.1.6 求 n 阶 Givens 方阵 $G_{ij}(\theta)$ 和 Householder 方阵 H_v 在 \mathbb{C} 上的 Jordan 标准形.

证明 ① Givens 方阵: 因为 $\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$ 的特征值为 $\cos \theta \pm i \sin \theta$, 无论 $\sin \theta$ 是否为 0 其均能在 \mathbb{C} 上相似对角化, 因此 $G_{ij}(\theta)$ 一定能在 \mathbb{C} 上相似对角化为

$$\text{diag}(\cos \theta + i \sin \theta, \cos \theta - i \sin \theta, 1, \dots, 1).$$

② Householder 方阵: 因为 $\mathbf{v} \neq \mathbf{0}$, 所以 $\operatorname{tr}(\mathbf{v}\mathbf{v}^\top) = \operatorname{tr}(\mathbf{v}^\top\mathbf{v}) = \|\mathbf{v}\|^2 \neq 0$, 由习题 5.1.11 (2) 知 $\mathbf{v}\mathbf{v}^\top$ 在 \mathbb{C} 上可相似对角化, 进而 H_v 在 \mathbb{C} 上可相似对角化. 而

$$\varphi_{H_v}(x) = \det\left((x-1)I + \frac{2}{\mathbf{v}^\top\mathbf{v}}\mathbf{v}\mathbf{v}^\top\right) = (x-1)^{n-1} \det((x-1)+2) = (x-1)^{n-1}(x+1),$$

因此 H_v 可在 \mathbb{C} 上相似对角化为

$$\operatorname{diag}(-1, 1, \dots, 1).$$

□

习题 6.1.10 设 $\alpha_k, \beta_k, \gamma_k$ 如定义 (6.2) 所述, $1 \leq k \leq n$. 证明:

- (1) $\beta_k = P\alpha_k$, 其中 $P = I - \sum_{i=1}^{k-1} \gamma_i \gamma_i^\top$ 与 $\begin{pmatrix} I_{n-k+1} & \\ & O \end{pmatrix}$ 正交相似.
 (2) $\det(\alpha_1, \dots, \alpha_k, \beta_{k+1}, \dots, \beta_n) = \det(\beta_1, \dots, \beta_k, \alpha_{k+1}, \dots, \alpha_n) = \det(A)$.

证明 (1) 按定义 (6.2), 有

$$\beta_k = \alpha_k - \sum_{i=1}^{k-1} (\gamma_i^\top \alpha_k) \gamma_i = \left(I - \sum_{i=1}^{k-1} \gamma_i \gamma_i^\top\right) \alpha_k.$$

欲证 $P = I - \sum_{i=1}^{k-1} \gamma_i \gamma_i^\top$ 与 $\begin{pmatrix} I_{n-k+1} & \\ & O \end{pmatrix}$ 正交相似, 即证 $-\sum_{i=1}^{k-1} \gamma_i \gamma_i^\top$ 与 $\begin{pmatrix} O & \\ & -I_{k-1} \end{pmatrix}$ 正交相似, 也即证 $\sum_{i=1}^{k-1} \gamma_i \gamma_i^\top$ 与 $\begin{pmatrix} I_{k-1} & \\ & O \end{pmatrix}$ 正交相似. 记 $A = \sum_{i=1}^{k-1} \gamma_i \gamma_i^\top$, 由 $\gamma_i^\top \gamma_j = \delta_{ij}$ 得

$$A^2 = \left(\sum_{i=1}^{k-1} \gamma_i \gamma_i^\top\right) \left(\sum_{i=1}^{k-1} \gamma_i \gamma_i^\top\right) = \sum_{i=1}^{k-1} \gamma_i \gamma_i^\top \gamma_i \gamma_i^\top = \sum_{i=1}^{k-1} \gamma_i \gamma_i^\top = A.$$

由例 (4.10) 中幂等方阵的性质可得 A 相似于 $\begin{pmatrix} I_r & \\ & O \end{pmatrix}$, 其中 $r = \operatorname{rank}(A)$. 下面证明 $\operatorname{rank}(A) = k-1$. 先证 $(\gamma_1 \ \dots \ \gamma_{k-1})$ 列满秩. 假设它不是列满秩的, 不妨设 γ_{k-1} 可由 $\gamma_1, \dots, \gamma_{k-2}$ 线性表出为

$$\gamma_{k-1} = \sum_{i=1}^{k-2} \mu_i \gamma_i,$$

则由 $\gamma_i^\top \gamma_{k-1} = 0$ ($i = 1, \dots, k-2$)、 $\gamma_i^\top \gamma_j = 0$ ($1 \leq i < j \leq k-2$) 及 $\|\gamma_i\| = 1$ ($i = 1, \dots, k-2$) 可得 $\mu_i = 0$ ($i = 1, \dots, k-2$), 与假设矛盾. 于是

$$A = \sum_{i=1}^{k-1} \gamma_i \gamma_i^\top = \underbrace{\begin{pmatrix} \gamma_1 & \cdots & \gamma_{k-1} \end{pmatrix}}_{\text{列满秩}} \underbrace{\begin{pmatrix} \gamma_1^\top \\ \vdots \\ \gamma_{k-1}^\top \end{pmatrix}}_{\text{行满秩}}$$

秩为 $k-1$ (用到了左乘列满秩矩阵或右乘行满秩矩阵不改变原矩阵的秩). 因此 A 相似于 $\begin{pmatrix} I_{k-1} & \\ & O \end{pmatrix}$, 即 A 的全体特征值为

$$\underbrace{1, \dots, 1}_{k-1 \text{ 个}}, \underbrace{0, \dots, 0}_{n-k+1 \text{ 个}}.$$

又因为 $A = \sum_{i=1}^{k-1} \gamma_i \gamma_i^\top$ 是实对称方阵, 由定理 6.7, 它可由正交方阵相似对角化. 由特征值即可

确定 A 正交相似于 $\begin{pmatrix} I_{k-1} & \\ & O \end{pmatrix}$. 从而原命题得证.

(2) 按定义 (6.2), 有

$$\textcircled{1} \beta_k = \alpha_k - \sum_{i=1}^{k-1} \frac{\beta_i^\top \alpha_k}{\beta_i^\top \beta_i} \beta_i \iff \textcircled{2} \alpha_k = \beta_k + \sum_{i=1}^{k-1} \frac{\beta_i^\top \alpha_k}{\beta_i^\top \beta_i} \beta_i.$$

于是利用互作线性表出进行初等变换有

$$\begin{aligned} \det(A) &= \det(\alpha_1, \dots, \alpha_k, \alpha_{k+1}, \dots, \alpha_n) \\ &\stackrel{\textcircled{2}}{=} \det(\beta_1, \dots, \alpha_k, \alpha_{k+1}, \dots, \alpha_n) \\ &\stackrel{\textcircled{2}}{=} \dots \\ &\stackrel{\textcircled{2}}{=} \det(\beta_1, \dots, \beta_k, \alpha_{k+1}, \dots, \alpha_n) \\ &\stackrel{\textcircled{2}}{=} \dots \\ &\stackrel{\textcircled{2}}{=} \det(\beta_1, \dots, \beta_k, \beta_{k+1}, \dots, \beta_n) \\ &\stackrel{\textcircled{1}}{=} \det(\alpha_1, \dots, \beta_k, \beta_{k+1}, \dots, \beta_n) \\ &\stackrel{\textcircled{1}}{=} \dots \\ &\stackrel{\textcircled{1}}{=} \det(\alpha_1, \dots, \alpha_k, \beta_{k+1}, \dots, \beta_n). \end{aligned}$$

□

习题 6.1.12 (1) 设 A 是反对称实数方阵. 证明: $I + A$ 可逆, 并且 $B = (I + A)^{-1}(I - A)$ 是正交方阵.

(2) 设 B 是正交方阵, 并且 $I + B$ 可逆. 证明: 存在反对称实数方阵 A 使得

$$B = (I + A)^{-1}(I - A).$$

证明 (1) 因为 A 是反对称方阵, 所以 A 的特征值不可能为非 0 实数, 即 $-1 \notin \text{Spec}(A)$, 因而 $I + A$ 可逆. 对于 $B = (I + A)^{-1}(I - A)$, 注意到关于 A 的有理函数 $(I + A)^{-1}$ 与 $I - A$ 乘积可交换 (这是因为若 $MN = NM$ 且 N 可逆, 则 $MN^{-1} = N^{-1}M$. 这里取 $M = I - A$, $N = I + A$), 直接验证得

$$B^T B = (I - A^T)(I + A^T)^{-1}(I + A)^{-1}(I - A) = (I + A)(I - A)^{-1}(I + A)^{-1}(I - A) = I.$$

(2) 直接解得 $A = (I + B)^{-1}(I - B)$. 验证: $A^T = (I - B^T)(I + B^T)^{-1} = (I - B^{-1})(I + B^{-1})^{-1}$, 为证 $A^T = -A$, 即证 $(I + B)(I - B^{-1}) + (I - B)(I + B^{-1}) = O$, 直接计算知成立. \square

习题 6.1.13 设正交方阵 $A = \begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix}$. 证明: $\det(A_1 A_1^T) \leq 1$, 等号成立当且仅当 $A_2 = O$.

证明 由 $A^T A = I$ 分块乘法可得 $A_1^T A_1 + A_3^T A_3 = I$. 因为 $A_1^T A_1$ 和 $A_3^T A_3$ 都是半正定的, 即对任意大小匹配的复向量 α , $\alpha^H A_1^T A_1 \alpha \geq 0$, $\alpha^H A_3^T A_3 \alpha \geq 0$. 特别地, 当 (α, λ) 为 $A_1^T A_1$ 的任一特征对时,

$$\alpha^H \alpha = \alpha^H (A_1^T A_1 + A_3^T A_3) \alpha = \lambda \alpha^H \alpha + \alpha^H A_3^T A_3 \alpha \geq \lambda \alpha^H \alpha,$$

于是 $0 \leq \lambda \leq 1$. 又因为 $A_1^T A_1$ 是对称实方阵, 由正交对角化知 $|\det(A_1^T A_1)| \leq 1$. 等号成立当且仅当 $A_3^T A_3 = O$ 即 $A_3 = O$, 由 A 是正交方阵知这又等价于 $A_2 = O$. \square

习题 6.1.14 设正交方阵 $A = \begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix}$, 其中 A_1 是方阵. 证明:

(1) A_1 的任意特征值 λ 满足 $|\lambda| \leq 1$, 从而 $|\det(A_1)| \leq 1$.

(2) $|\det(A_1)| = 1$ 当且仅当 $A = \text{diag}(A_1, A_4)$.

证明 (1) 由 $A^T A = I$ 分块乘法可得 $A_1^T A_1 + A_3^T A_3 = I$. 对 A_1 的任一特征值 λ , 设其对应的特征向量为 α , 则

$$\alpha^H (A_1^T A_1 + A_3^T A_3) \alpha = \alpha^H \alpha.$$

利用 $A_1\alpha = \lambda\alpha$ 展开即

$$|\lambda|^2 \alpha^H \alpha = \alpha^H \alpha - \|A_3 \alpha\|^2 \leq \alpha^H \alpha,$$

故 $|\lambda|^2 \leq 1$, 即 $|\lambda| \leq 1$. 从而 $|\det(A_1)| \leq 1$.

(2) \Leftarrow : 由 $A_3 = O$ 得 $A_1^T A_1 = I$, 故 $|\det(A_1)| = 1$.

\Rightarrow : $|\det(A_1)| = 1$ 即 $\det(A_1 A_1^T) = 1$. 由习题 6.1.12 知 $A_2 = O$. 再由 A 是正交方阵、 A_1 是方阵得 $A_3 = O$. \square

习题 6.1.15 设 $A \in \{-1, 1\}^{n \times n}$ 满足 $AA^T = nI$. 证明: 若 A 有 $p \times q$ 全一子矩阵, 则 $pq \leq n$.

证明 设由该子矩阵所在行除去这 q 列后得到的行向量为 $\alpha_1, \dots, \alpha_p$. 由 $AA^T = nI$ 得

$$\alpha_i \alpha_j^T + q = n\delta_{ij}.$$

故

$$\begin{aligned} 0 &\leq (\alpha_1 + \dots + \alpha_p)(\alpha_1 + \dots + \alpha_p)^T \\ &= p(n - q) + p(p - 1)(-q) \\ &= p(n - pq), \end{aligned}$$

即 $pq \leq n$. \square

习题 6.1.16 设 P 是 n 阶可逆实数方阵, 并且对于任意 n 阶对称实数方阵 A , $P^{-1}AP$ 都是对称方阵. 证明: 存在实数 λ 使得 λP 是正交方阵.

证明 由题, 对任意 n 阶对称实数方阵 A ,

$$P^{-1}AP = P^T A^T P^{-T} \iff APP^T = PP^T A^T = PP^T A,$$

即 PP^T 与任意 n 阶对称实数方阵乘积可交换. 依次取 $A = E_{kk}$ ($k = 1, \dots, n$) 可知 PP^T 必

为对角阵 $\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$. 再依次取 $A = E_{ij} + E_{ji}$ ($i = 1, \dots, n, j = 1, \dots, n, i \neq j$) 可知

$\lambda_i = \lambda_j$ ($i \neq j$). 故存在 $\mu > 0$, 使得 $PP^T = \mu I$. 若取 $\lambda = \frac{1}{\sqrt{\mu}}$, 则 $(\lambda P)(\lambda P)^T = \frac{1}{\mu} PP^T = I$, 此 λ 即为所求. \square

习题 6.2.1 (1) 证明: 任意 3 阶正交方阵 A 可以正交相似于 $\begin{pmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & \det(A) \end{pmatrix}$ 的形式.

(2) 说明上述结论的几何含义.

证明 (1) 由于奇数阶正交方阵的行列式必为其特征值 (定理 6.2 的推论), 所以 $\det(A) \in \text{Spec}(A)$. 由定理 6.6: 若 A 的特征值均为实数, 则 A 相似于 $\begin{pmatrix} \pm I & \\ & \det(A) \end{pmatrix}$; 若 A 有虚特征

值, 则 A 相似于 $\begin{pmatrix} \cos \theta & \sin \theta & \\ -\sin \theta & \cos \theta & \\ & & \det(A) \end{pmatrix}$ (其中 $\sin \theta \neq 0$). 这两种形式都属于所要的形式.

(2) ① 若 $\det(A) = 1$, 即 $1 \in \text{Spec}(A)$, 则存在 $\alpha \neq 0$, 满足 $A\alpha = \alpha$, 于是 A 对应的线性变换是以 α 为轴的旋转.

② 若 $\det(A) = -1$, 即 $-1 \in \text{Spec}(A)$, 则存在 $\alpha \neq 0$, 满足 $A\alpha = -\alpha$, 于是 A 对应的线性变换是以 α 为轴的旋转与以 α 为法向的镜面反射的复合. \square

习题 6.2.2 设对称方阵 $A, B \in \mathbb{C}^{n \times n}$ 满足 $\varphi_A = \varphi_B$. A 与 B 是否一定相似?

解 不一定. 反例: 令 $A = \begin{pmatrix} 1 & i \\ i & -1 \end{pmatrix}, B = O$, 则 $\varphi_A(x) = \varphi_B(x) = x^2$, 但 $\text{rank}(A) \neq \text{rank}(B)$. \square

习题 6.2.3 设 A 是实数方阵, $k \in \mathbb{N}^*$. 证明: A 可以相似于正交方阵 $\iff A^k$ 可以相似于正交方阵.

证明 先证明如下引理:

引理 对于 $k \in \mathbb{N}^*$ 与 $\lambda \in \mathbb{C} \setminus \{0\}$, $J_n(\lambda)^k$ 的 Jordan 标准形为 $J_n(\lambda^k)$.

【引理的证明: 对 $J_n(\lambda) = \lambda I + J_n(0)$ 作二项式展开得

$$J_n(\lambda)^k = \sum_{i=0}^k \binom{k}{i} \lambda^{k-i} J_n(0)^i = \begin{pmatrix} \lambda^k & C_k^1 \lambda^{k-1} & C_k^2 \lambda^{k-2} & \cdots & C_k^{n-1} \lambda^{k-n+1} & \\ & \ddots & \ddots & \ddots & \vdots & \\ & & \ddots & C_k^2 \lambda^{k-2} & & \\ & & & C_k^1 \lambda^{k-1} & & \\ & & & & \lambda^k & \end{pmatrix}.$$

其中当 $m > k$ 时 C_k^m 取为 0. 记 $B = J_n(\lambda)^k - \lambda^k I$. 由于 $n - \text{rank}(B)$ 即 $J_n(\lambda)$ 的 Jordan 标准形中特征值为 λ^k 的 Jordan 块的个数, 所以只需证 $\text{rank}(B) = n - 1$. 注意到当 $\lambda \neq 0$ 时, $\det(B) = 0$ 且 B 的右上角 $n - 1$ 阶子式行列式非零即得证.】

由定理 6.6 知任意正交方阵可正交相似于

$$\text{diag} \left(\begin{pmatrix} \cos \theta_1 & \sin \theta_1 \\ -\sin \theta_1 & \cos \theta_1 \end{pmatrix}, \dots, \begin{pmatrix} \cos \theta_s & \sin \theta_s \\ -\sin \theta_s & \cos \theta_s \end{pmatrix}, I, -I \right),$$

其中 $\theta_k \in (0, \pi)$ ($k = 1, \dots, s$). 于是 $\begin{pmatrix} \cos \theta_k & \sin \theta_k \\ -\sin \theta_k & \cos \theta_k \end{pmatrix}$ 的特征值为 $\cos \theta_k \pm i \sin \theta_k$ (互异), 故任意正交方阵可在 \mathbb{C} 上相似对角化. 由于 A 是实方阵, A 可在 \mathbb{R} 上相似于正交方阵 $\iff A$ 可在 \mathbb{C} 上相似于正交方阵 $\iff A$ 可在 \mathbb{C} 上相似于 $\text{diag}(e^{i\theta_1}, \dots, e^{i\theta_n})$, 其中 $\theta_i \in \mathbb{R}$ ($i = 1, \dots, n$). 而这又等价于 A^k 可在 \mathbb{C} 上相似于 $\text{diag}(e^{i(k\theta_1)}, \dots, e^{i(k\theta_n)})$ ($k \in \mathbb{N}^*$) (\Rightarrow 是显然的; \Leftarrow 是因为 $\det(A^k) \neq 0$ 推出 A 可逆, 对 A 的一个 Jordan 块 $J_s(\lambda)$, 当 $|\lambda| \neq 0$ 时, $J_s(\lambda)^k$ 的 Jordan 标准形为 $J_s(\lambda^k)$, 故由 A^k 可相似对角化能推出 A 的 Jordan 标准形中所有 Jordan 块都是一阶的, 也即 A 可相似对角化, 且可知 A 的特征值模长均为 1). 进而等价于 A^k 可在 \mathbb{C} 上相似于正交方阵, 也即可在 \mathbb{R} 上相似于正交方阵. \square

习题 6.2.4 设 A 是实数方阵. 证明: A 可以在 \mathbb{R} 上相似于规范方阵 $\iff A$ 可以在 \mathbb{C} 上相似于对角方阵.

证明 \Rightarrow : 由定理 6.9 知任意规范实数方阵可正交相似于

$$\text{diag} \left(\begin{pmatrix} a_1 & b_1 \\ -b_1 & a_1 \end{pmatrix}, \dots, \begin{pmatrix} a_s & b_s \\ -b_s & a_s \end{pmatrix}, \lambda_{2s+1}, \dots, \lambda_n \right).$$

对于二阶实方阵 $\begin{pmatrix} a_k & b_k \\ -b_k & a_k \end{pmatrix}$ ($k = 1, \dots, s$), 其特征值 $a_k \pm b_k i$ 互异, 故可在 \mathbb{C} 上相似对角化. 故 A 可在 \mathbb{C} 上相似于对角方阵.

\Leftarrow : 若 A 可在 \mathbb{C} 上相似对角化, 可设 A 相似于 $\text{diag}(\lambda_1, \bar{\lambda}_1, \dots, \lambda_s, \bar{\lambda}_s, \lambda_{2s+1}, \dots, \lambda_n)$, 其中 $\lambda_1, \dots, \lambda_s \notin \mathbb{R}$, $\lambda_{2s+1}, \dots, \lambda_n \in \mathbb{R}$. 对于 $\lambda_k = a_k + b_k i$ ($k = 1, \dots, s$), $\begin{pmatrix} \lambda_k & \\ & \bar{\lambda}_k \end{pmatrix}$ 与 $\begin{pmatrix} a_k & b_k \\ -b_k & a_k \end{pmatrix}$ 在 $\mathbb{C}[x]$ 上的 Smith 标准形均为 $\begin{pmatrix} 1 & \\ & (x-a)^2 + b^2 \end{pmatrix}$, 故它们在 \mathbb{C} 上相似. 最终的形式就说明 A 可以在 \mathbb{R} 上相似于规范方阵. \square

习题 6.2.4 旧 证明: 上三角的规范实数方阵一定是对角方阵.

证明 只需证明如下结论: 若 $\begin{pmatrix} A_1 & A_2 \\ O & A_3 \end{pmatrix}$ 是规范实方阵, 则 A_1, A_3 均为规范实方阵.

由 $AA^T = A^T A$ 得 $A_1 A_1^T + A_2 A_2^T = A_1^T A_1$. 两边求迹并利用 $\text{tr}(A_1 A_1^T) = \text{tr}(A_1^T A_1)$ 得 $\text{tr}(A_2 A_2^T) = A_2$ 所有元素平方和 $= 0$, 故 $A_2 = O$. 从而推出 $A_1 A_1^T = A_1^T A_1$, 即 A_1 规范. 同理 A_3 规范. 这样在对角线上归纳就可证明原命题. \square

习题 6.2.5 (2)(3) 设 A 是 n 阶规范实数方阵, $\lambda \in \mathbb{C}$ 和 $\alpha \in \mathbb{C}^{n \times 1}$ 是 A 的一个特征值和相应的特征向量. 证明:

(2) α 也是 A^T 的特征向量.

(3) 若 $\lambda \notin \mathbb{R}$, $\alpha = u + iv$, 其中 $u, v \in \mathbb{R}^{n \times 1}$, 则 $\|u\| = \|v\|$ 且 $u \perp v$.

证明 (2) **法一** 由定理 6.15, 设酉方阵 P 使得 $P^H A P = D$ 为复对角方阵. 则

- α 是 A 的特征向量 $\iff P^H \alpha$ 是 D 的特征向量;
- α 是 A^T 的特征向量 $\iff P^H \alpha$ 是 D^H 的特征向量.

因此问题转化为证明: 若 β 是复对角方阵 D 的特征向量, 则 β 也是 \bar{D} 的特征向量.

设 $D = \text{diag}(\lambda_1, \dots, \lambda_n)$, $\beta = (b_1 \ \dots \ b_n)^T$. 不妨设 $b_1, \dots, b_s \neq 0, b_{s+1} = \dots = b_n = 0$.

$$\text{则由 } \begin{pmatrix} \lambda_1 b_1 \\ \vdots \\ \lambda_s b_s \end{pmatrix} \text{ 与 } \begin{pmatrix} b_1 \\ \vdots \\ b_s \end{pmatrix} \text{ 共线可得 } \lambda_1 = \dots = \lambda_s =: \lambda. \text{ 于是 } \bar{D}\beta = \begin{pmatrix} \bar{\lambda} b_1 \\ \vdots \\ \bar{\lambda} b_s \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \bar{\lambda} \beta, \text{ 即 } \beta \text{ 是 } \bar{D}$$

的特征向量.(我们还得到了 A 与 A^T 对应于 α 的特征值共轭.)

法二 设 \mathcal{A} 为复内积空间 \mathbb{C}^n 上与 A 对应的规范变换, 则 $\mathcal{A}\alpha = \lambda\alpha$. 下证 $\mathcal{A}^*\alpha = \bar{\lambda}\alpha$. 利用定理 10.30 中 ① \iff ③ 及定理 10.27 中 $(\mathcal{A}^*)^* = \mathcal{A}$ 可得

$$\begin{aligned} \|\mathcal{A}^*\alpha - \bar{\lambda}\alpha\|^2 &= (\mathcal{A}^*\alpha - \bar{\lambda}\alpha, \mathcal{A}^*\alpha - \bar{\lambda}\alpha) \\ &= (\mathcal{A}^*\alpha, \mathcal{A}^*\alpha) - \bar{\lambda}(\mathcal{A}^*\alpha, \alpha) - \lambda(\alpha, \mathcal{A}^*\alpha) + |\lambda|^2(\alpha, \alpha) \\ &= (\mathcal{A}\alpha, \mathcal{A}\alpha) - \bar{\lambda}(\alpha, \mathcal{A}\alpha) - \lambda(\mathcal{A}\alpha, \alpha) + |\lambda|^2(\alpha, \alpha) \\ &= (\mathcal{A}\alpha - \lambda\alpha, \mathcal{A}\alpha - \lambda\alpha) \\ &= \|\mathcal{A}\alpha - \lambda\alpha\|^2 \\ &= 0. \end{aligned}$$

(3) 由 (2) 可设 $A\alpha = \lambda\alpha$ 与 $A^T\alpha = \bar{\lambda}\alpha$. 再设 $\lambda = a + ib$ ($b \neq 0$), $\alpha = u + iv$. 于是

$$A\alpha = \lambda\alpha \implies \begin{cases} Au = au - bv \\ Av = av + bu \end{cases}, \quad A^T\alpha = \bar{\lambda}\alpha \implies \begin{cases} A^T u = au + bv \\ A^T v = av - bu \end{cases}.$$

从而

$$\begin{cases} u^T Au = au^T u - bu^T v \\ u^T A^T u = au^T u + bu^T v \end{cases} \xrightarrow[u^T v = v^T u = 0]{u^T Au = (u^T A^T u)^T} bu^T v = 0 \implies u^T v = 0 \implies u \perp v.$$

类似有

$$\begin{cases} v^T Au = av^T u - bv^T v \\ u^T A^T v = au^T v - bu^T u \end{cases} \xrightarrow[u^T v = v^T u = 0]{v^T Au = (u^T A^T v)^T} u^T u = v^T v \implies \|u\| = \|v\|.$$

□

习题 6.2.6 设 n 阶实数方阵 $A = (a_{ij})$ 满足 $a_{ij} = 0, \forall |i - j| \geq 2$ 并且 $a_{i,i+1}a_{i+1,i} > 0, \forall i < n$.

(1) 证明: A 可以在 \mathbb{R} 上相似于对称方阵, 从而 A 可以相似于对角方阵.

(2) 证明: $d_A = \varphi_A$, 从而 A 有 n 个两两不同的实特征值.

证明 (1) 我们断言 A 可以在 \mathbb{R} 上对角相似于三对角对称方阵. 通过待定系数法确定此相似: 设

$$\begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{pmatrix} \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} \frac{1}{d_1} & & \\ & \ddots & \\ & & \frac{1}{d_n} \end{pmatrix} = \left(\frac{d_i}{d_j} a_{ij} \right)_{n \times n},$$

则由其 (三对角) 对称方阵可得

$$\frac{d_i}{d_j} a_{ij} = \frac{d_j}{d_i} a_{ji} \implies d_i^2 a_{ij} = d_j^2 a_{ji}.$$

只考虑 $i \neq j$ 也即只需考虑 $|i - j| = 1$ 的情形. 由条件得 $a_{ij}a_{ji} > 0$, 若要求每个 $d_i > 0$, 则在取定 d_1 后, 由方程组

$$\begin{cases} \frac{d_1^2 a_{12}}{a_{21}} = d_2^2 \\ \frac{d_2^2 a_{23}}{a_{32}} = d_3^2 \\ \vdots \\ \frac{d_{n-1}^2 a_{n-1,n}}{a_{n,n-1}} = d_n^2 \end{cases}$$

可依次解得 d_2, \dots, d_n . 这就构造出 A 在 \mathbb{R} 上相似于 (三对角) 对称方阵的情形, 从而 A 可相似对角化.

(2) 由条件得 $\prod_{i=2}^n a_{i,i-1} \neq 0$. 故由习题 5.4.8 知 $d_A = \varphi_A$. 从而 A 的每个特征值只对应于 1 个 Jordan 块, 结合 (1) 知这些 Jordan 块都是 1 阶的, 即 A 有 n 个两两不同的实特征值 (还用到了对称方阵的特征值都是实数). \square

习题 6.2.7 设实数方阵 A 与 B 乘积可交换. 证明:

(1) 若 A 是规范方阵, 则 A 与 B^T 乘积可交换.

(2) 若 A 和 B 都是规范方阵, 则 $A+B$ 和 AB 也都是规范方阵.

证明 (1) 设 $C = AB^T - B^T A$. 为证 $C = O$, 只需证 $\text{tr}(C^T C) = 0$. 而

$$\begin{aligned} \text{tr}(C^T C) &= \text{tr}((BA^T - A^T B)(AB^T - B^T A)) \\ &= \text{tr}(BA^T AB^T - BA^T B^T A - A^T BAB^T + A^T BB^T A) \\ &= \text{tr}(BA^T AB^T) - \text{tr}(BA^T B^T A) - \text{tr}(A^T BAB^T) + \text{tr}(A^T BB^T A) \\ &\stackrel{\substack{AB=BA \\ A^T A=AA^T}}{=} \text{tr}(B^T BA^T A) - \text{tr}(BB^T A^T A) - \text{tr}(BB^T A^T A) + \text{tr}(BB^T A^T A) \\ &= \text{tr}(B^T BA^T A) - \text{tr}(BB^T A^T A), \end{aligned}$$

再由 $\text{tr}(B^T BA^T A) = \text{tr}(A^T AB^T B) = \text{tr}(AA^T B^T B) = \text{tr}(AB^T A^T B) = \text{tr}(A^T BAB^T) = \text{tr}(A^T ABB^T) = \text{tr}(BB^T A^T A)$ 即得 $\text{tr}(C^T C) = 0$.

(2) 只需利用 $A^T A = AA^T, B^T B = BB^T, A^T B = BA^T, AB^T = B^T A$. \square

注 2.0.12 (1) 的证明也可由习题 6.2.10 (1) 得到 $A^T = f(A)$ 与 B 乘积可交换, 进而 A 与 B^T 乘积可交换.

习题 6.2.8 分别给出满足下列条件的 n 阶实数方阵 A, B 的例子.

(1) A 与 B 乘积可交换, A 与 B^T 乘积不交换.

(2) A 和 B 都是规范方阵, $A+B$ 和 AB 都不是规范方阵.

解 (1) $A = B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$.

(2) $A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}, B = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. 此时 $A+B = \begin{pmatrix} 0 & 2 \\ 0 & 1 \end{pmatrix}$ 与 $AB = \begin{pmatrix} -1 & 0 \\ -1 & 1 \end{pmatrix}$ 显然都不符合 2 阶规范实方阵的形状. \square

习题 6.2.9 设 A 是 n 阶实数方阵. 证明:

- (1) 若 $A + A^T$ 的特征值都是正数, 则 A 的特征值的实部都是正数.
 (2) 若 A 的特征值的实部都是正数, 则 $A + A^T$ 必有特征值是正数.

证明 (1) 因为对称方阵 $A + A^T$ 的特征值都是正实数, 所以由定理 7.4(2), $A + A^T$ 是正定的. 设 (λ, \mathbf{x}) 是 A 的任一特征对, 即 $A\mathbf{x} = \lambda\mathbf{x}$, 则

$$\begin{cases} A\mathbf{x} = \lambda\mathbf{x} \implies \mathbf{x}^H A\mathbf{x} = \lambda\mathbf{x}^H \mathbf{x}, \\ \mathbf{x}^H A^T = \bar{\lambda}\mathbf{x}^H \implies \mathbf{x}^H A^T \mathbf{x} = \bar{\lambda}\mathbf{x}^H \mathbf{x}. \end{cases}$$

于是 $\mathbf{x}^H (A + A^T) \mathbf{x} = 2\operatorname{Re}(\lambda)\mathbf{x}^H \mathbf{x}$. 由 $A + A^T$ 的正定性与 $\mathbf{x}^H \mathbf{x} > 0$ 知 $\operatorname{Re}(\lambda) > 0$.

(2) 用反证法. 假设 $A + A^T$ 的特征值的实部都非正数, 则 $A + A^T$ 是负定的, 同 (1) 可得对 A 的任一特征对 (λ, \mathbf{x}) , 有 $\mathbf{x}^H (A + A^T) \mathbf{x} = 2\operatorname{Re}(\lambda)\mathbf{x}^H \mathbf{x} < 0$, 于是 $\operatorname{Re}(\lambda) < 0$, 与条件矛盾. \square

习题 6.2.10 (1) 设 A 是 n 阶规范实数方阵, 则 A 与 A^T 正交相似, 并且存在 $f \in \mathbb{R}[x]$ 使得 $A^T = f(A)$.

证明 ① 不妨设 A 已经是定理 6.9 中的准对角方阵. 则只需证规范实方阵 $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$ 与 $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ 正交相似. 而这可由 $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ 立即得到. 【也可用 $\begin{pmatrix} & 1 \\ 1 & \end{pmatrix}$ 做相似.】

② 对于正交方阵 P , $A^T = f(A)$ 能推出 $(P^T A P)^T = f(P^T A P)$, 故仍可假设 A 已经是定理 6.9 中的准对角方阵, 进而只需考虑对所有二阶块 $\begin{pmatrix} a_i & b_i \\ -b_i & a_i \end{pmatrix}$ ($i = 1, \dots, s$) 存在同一

个 $f \in \mathbb{R}[x]$, 使得 $f \begin{pmatrix} a_i & b_i \\ -b_i & a_i \end{pmatrix} = \begin{pmatrix} a_i & -b_i \\ b_i & a_i \end{pmatrix}$, 且对 λ_i ($i = 2s + 1, \dots, n$) 有 $f(\lambda_i) = \lambda_i$. 利用 $(x - a_i)^2 + b_i^2$ 是 $\begin{pmatrix} a_i & b_i \\ -b_i & a_i \end{pmatrix}$ 的化零多项式, 构造 f 满足

$$\begin{cases} f \equiv -x + 2a_i & \text{mod } (x - a_i)^2 + b_i^2, \\ f \equiv \lambda_i & \text{mod } x - \lambda_i. \end{cases}$$

为了满足中国剩余定理的条件, 对于 $\lambda_{2s+1}, \dots, \lambda_n$, 只保留其中相异的部分; 对于二次实多项式 $(x - a_i)^2 + b_i^2$, 若有两个多项式不互素, 则它们有公共根, 由 $b_i \neq 0$ 推出此根为虚根, 由其共轭复数也是根以及二次实多项式的约束同样有这两个多项式相等, 故可以只保留其中两两互素的多项式. 这样就由中国剩余定理构造出所需的 f . \square

习题 6.2.11 设 I 是指标集合, $\{A_i \mid i \in I\}$ 是一组两两乘积可交换的规范方阵. 证明: 存在正交方阵 P , 使得每个 $P^{-1}A_iP$ 是定理 6.9 中的准对角方阵.

证明 由习题 5.2.10 (1) 知存在复数向量 α 是所有 A_i 的公共特征向量. 不妨设 $\|\alpha\| = 1$. 当 $A_i \in \mathbb{R}^{1 \times 1}$ 时命题自动成立. 设对某个 $n \geq 1$, 命题已对大小不超过 $n \times n$ 的矩阵组成的交换族成立 (允许调换对角块顺序).

① 若 α 是实向量, 则与其对应的特征值 $\lambda_i \in \mathbb{R}$. 构造以 α 为第一列的正交方阵 P_1 , 则

$$A_i P_1 = A_i \begin{pmatrix} \alpha & * \end{pmatrix} = \begin{pmatrix} \alpha & * \end{pmatrix} \begin{pmatrix} \lambda_i & * \\ & B_i \end{pmatrix}, \quad \forall i \in I.$$

则由 $A_i = P_1 \begin{pmatrix} \lambda_i & * \\ & B_i \end{pmatrix} P_1^\top$ 及乘积可交换的性质得

$$A_i A_j = P_1 \begin{pmatrix} \lambda_i & * \\ & B_i \end{pmatrix} \begin{pmatrix} \lambda_j & * \\ & B_j \end{pmatrix} P_1^\top = P_1 \begin{pmatrix} \lambda_j & * \\ & B_j \end{pmatrix} \begin{pmatrix} \lambda_i & * \\ & B_i \end{pmatrix} P_1^\top = A_j A_i,$$

由此可得 $B_i B_j = B_j B_i$. 另外, 由 A_i 规范以及正交相似保持规范性得 $\begin{pmatrix} \lambda_i & * \\ & B_i \end{pmatrix}$ 规范, 从

$$\begin{pmatrix} \lambda_i & \\ \beta & B_i^\top \end{pmatrix} \begin{pmatrix} \lambda_i & \beta^\top \\ & B_i \end{pmatrix} = \begin{pmatrix} \lambda_i & \beta^\top \\ & B_i \end{pmatrix} \begin{pmatrix} \lambda_i & \\ \beta & B_i \end{pmatrix}$$

可得 $\beta = 0$ 且 $B_i^\top B_i = B_i B_i^\top$. 因此 $\{B_i \mid i \in I\}$ 是一组两两乘积可交换的规范方阵. 对 B_i 运用归纳假设知存在正交方阵 P_2 使得每个 $P_2^{-1}B_iP_2$ 都是定理 6.9 中的准对角方阵. 于是 $P = P_1 \begin{pmatrix} 1 & \\ & P_2 \end{pmatrix}$ 为所需的正交方阵.

② 若 $\alpha = u + iv$ 不是实向量. 若 $u = kv$, 则 $\frac{1}{k+i}\alpha$ 为公共实特征向量, 对其作单位化后化为 ① 中情形. 若 u 与 v 线性无关, 设 A_i 与 α 对应的特征值为 $\lambda_i = a + ib_i$.

• 若 $\lambda_i \in \mathbb{R}, \forall i$, 则因为实方阵的实特征值一定有对应的实特征向量 (参见习题 6.3.9 (6)) 证明中的引理 (1)), 选 v 并进行单位化就得到 A_i 的一个公共实单位特征向量, 化为 ① 中情形.

- 若存在某个 $\lambda_i \notin \mathbb{R}$. 则由习题 6.2.5 (2)(3) 可得 $\|\mathbf{u}\| = \|\mathbf{v}\|$ 且 $\mathbf{u} \perp \mathbf{v}$.

$$A_i(\mathbf{u} + i\mathbf{v}) = (a_i + b_i i)(\mathbf{u} + i\mathbf{v}) \implies \begin{cases} A_i \mathbf{u} = a_i \mathbf{u} - b_i \mathbf{v} \\ A_i \mathbf{v} = a_i \mathbf{v} + b_i \mathbf{u} \end{cases}$$

$$\iff A_i \begin{pmatrix} \mathbf{u} & \mathbf{v} \end{pmatrix} = \begin{pmatrix} \mathbf{u} & \mathbf{v} \end{pmatrix} \begin{pmatrix} a_i & b_i \\ -b_i & a_i \end{pmatrix}.$$

记 $\tilde{\mathbf{u}} = \frac{\mathbf{u}}{\|\mathbf{u}\|}$, $\tilde{\mathbf{v}} = \frac{\mathbf{v}}{\|\mathbf{v}\|}$, 则有

$$A_i \begin{pmatrix} \tilde{\mathbf{u}} & \tilde{\mathbf{v}} \end{pmatrix} = \begin{pmatrix} \tilde{\mathbf{u}} & \tilde{\mathbf{v}} \end{pmatrix} \begin{pmatrix} a_i & b_i \\ -b_i & a_i \end{pmatrix}.$$

将 $\begin{pmatrix} \tilde{\mathbf{u}} & \tilde{\mathbf{v}} \end{pmatrix}$ 扩充为正交方阵 $P_1 = \begin{pmatrix} \tilde{\mathbf{u}} & \tilde{\mathbf{v}} & * \end{pmatrix}$, 则

$$A_i P_1 = A_i \begin{pmatrix} \tilde{\mathbf{u}} & \tilde{\mathbf{v}} & * \end{pmatrix} = \begin{pmatrix} \tilde{\mathbf{u}} & \tilde{\mathbf{v}} & * \end{pmatrix} \begin{pmatrix} a_i & b_i & * \\ -b_i & a_i & * \\ & & B_i \end{pmatrix}, \quad \forall i \in I.$$

则由 $A_i = P_1 \begin{pmatrix} a_i & b_i & * \\ -b_i & a_i & * \\ & & B_i \end{pmatrix} P_1^\top$ 及乘积可交换的性质得

$$\begin{aligned} A_i A_j &= P_1 \begin{pmatrix} a_i & b_i & * \\ -b_i & a_i & * \\ & & B_i \end{pmatrix} \begin{pmatrix} a_j & b_j & * \\ -b_j & a_j & * \\ & & B_j \end{pmatrix} P_1^\top \\ &= P_1 \begin{pmatrix} a_j & b_j & * \\ -b_j & a_j & * \\ & & B_j \end{pmatrix} \begin{pmatrix} a_i & b_i & * \\ -b_i & a_i & * \\ & & B_i \end{pmatrix} P_1^\top = A_j A_i. \end{aligned}$$

由此可得 $B_i B_j = B_j B_i$. 另外, 由 A_i 规范以及正交相似保持规范性得 $\begin{pmatrix} a_i & b_i & * \\ -b_i & a_i & * \\ & & B_i \end{pmatrix}$ 规范, 而准上三角规范实方阵是准对角方阵 (参见习题 6.2.4 旧证明中用到的结论), 因此上面 “*” 为零矩阵, 进而可知 $B_i^\top B_i = B_i B_i^\top$. 故 $\{B_i \mid i \in I\}$ 是一组两两乘积可交换的规范方阵. 对 B_i 运用归纳假设知存在正交方阵 P_2 使得每个 $P_2^{-1} B_i P_2$ 都是定理 6.9 中的准对角方阵. 于是 $P = P_1 \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ & & P_2 \end{pmatrix}$ 为所需的正交方阵. \square

习题 6.3.2 设实数矩阵 $A, B \in \mathbb{R}^{m \times n}$ 满足 $A^\top A = B^\top B$. 证明: 存在正交方阵 P 使得 $PA = B$.

证明 由 $A^T A = B^T B$ 知 A 与 B 有相同的正交相抵标准形, 即存在正交方阵 P_1, P_2, Q_1, Q_2 使得 $P_1 A Q_1 = P_2 B Q_2$. 再由定理 6.10 证明过程可知 Q_1, Q_2 分别为 $A^T A, B^T B$ 作相似对角化时所用的正交方阵, 自然可取 $Q_1 = Q_2$. 于是 $(P_2^T P_1) A = B$. \square

习题 6.3.4 设实数方阵 A 的行向量两两正交, 列向量也两两正交. 证明: 存在置换方阵 P_1, P_2 , 正交方阵 Q_1, \dots, Q_k 和实数 $\lambda_1, \dots, \lambda_k$, 使得

$$A = P_1 \begin{pmatrix} \lambda_1 Q_1 & & \\ & \ddots & \\ & & \lambda_k Q_k \end{pmatrix} P_2.$$

证明 设置换方阵 P_1 使得 $P_1^T A$ 的行向量模长为降序, 再设置换方阵 P_2 使得 $P_1^T A P_2^T$ 的列向量模长为降序 (显然第二步不会破坏第一步的有序性). 记 $B = P_1^T A P_2^T$, 对 B 的非零列向量作 Gram-Schmidt 标准正交化得到正交方阵 P , 则由 B 的行 (列) 向量两两正交可知

$$B = P \begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{pmatrix}, \text{ 其中 } a_1 \geq \dots \geq a_n \geq 0 (B \text{ 的列向量可能为零向量, 因此 } a_i \text{ 可能为}$$

零). 类似地, 将上述操作改为对 B 的行向量进行可得 $B = \begin{pmatrix} b_1 & & \\ & \ddots & \\ & & b_n \end{pmatrix} Q$, 其中 Q 是正交

方阵, $b_1 \geq \dots \geq b_n \geq 0$. 则有

$$B^T B = \begin{pmatrix} a_1^2 & & \\ & \ddots & \\ & & a_n^2 \end{pmatrix} = Q^T \begin{pmatrix} b_1^2 & & \\ & \ddots & \\ & & b_n^2 \end{pmatrix} Q.$$

由 $B^T B$ 特征值的唯一性及非负性可知 b_1, \dots, b_n 是 a_1, \dots, a_n 的一个排列. 记 $\lambda_1, \dots, \lambda_k$ 为这 n 个数中所有相异的数的降序排列, 其重数依次为 n_1, \dots, n_k . 则

$$\begin{pmatrix} Q_{11} & \cdots & Q_{kk} \\ \vdots & \ddots & \vdots \\ Q_{k1} & \cdots & Q_{kk} \end{pmatrix} \begin{pmatrix} \lambda_1 I_{n_1} & & \\ & \ddots & \\ & & \lambda_k I_{n_k} \end{pmatrix} = \begin{pmatrix} \lambda_1 I_{n_1} & & \\ & \ddots & \\ & & \lambda_k I_{n_k} \end{pmatrix} \begin{pmatrix} Q_{11} & \cdots & Q_{kk} \\ \vdots & \ddots & \vdots \\ Q_{k1} & \cdots & Q_{kk} \end{pmatrix},$$

其中 $(Q_{ij})_{k \times k}$ 是 Q 的分块. 两边计算得 $\lambda_j Q_{ij} = \lambda_i Q_{ij}$, 而当 $i \neq j$ 时 $\lambda_i \neq \lambda_j$, 因此

$Q_{ij} = 0, \forall i \neq j$, 即 Q 是准对角方阵 (记为 $\text{diag}(Q_1, \dots, Q_k)$). 则

$$A = P_1 B P_2 = P_1 \begin{pmatrix} \lambda_1 I_{n_1} & & \\ & \ddots & \\ & & \lambda_k I_{n_k} \end{pmatrix} \begin{pmatrix} Q_1 & & \\ & \ddots & \\ & & Q_k \end{pmatrix} P_2 = P_1 \begin{pmatrix} \lambda_1 Q_1 & & \\ & \ddots & \\ & & \lambda_k Q_k \end{pmatrix} P_2.$$

□

习题 6.3.5 设 $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{C}$ 是 $A \in \mathbb{R}^{n \times n}$ 的所有特征值. 证明: A 与 $\text{diag}(|\lambda_1|, |\lambda_2|, \dots, |\lambda_n|)$ 正交相抵当且仅当 A 是规范方阵.

证明 \Leftarrow : 不妨设规范方阵 A 已经是定理 6.9 中的准对角方阵

$$\text{diag} \left(\begin{pmatrix} a_1 & b_1 \\ -b_1 & a_1 \end{pmatrix}, \dots, \begin{pmatrix} a_s & b_s \\ -b_s & a_s \end{pmatrix}, \lambda_{2s+1}, \dots, \lambda_n \right).$$

则

$$\begin{aligned} A^T A &= \text{diag}((a_1^2 + b_1^2)I_2, \dots, (a_s^2 + b_s^2)I_2, \lambda_{2s+1}^2, \dots, \lambda_n^2) \\ &= \text{diag}(|\lambda_1|^2 I_2, \dots, |\lambda_{2s}|^2 I_2, \lambda_{2s+1}^2, \dots, \lambda_n^2). \end{aligned}$$

于是由正交相抵标准形即得证.

\Rightarrow : 不妨设 A 就是定理 6.5 中的准上三角方阵

$$\begin{pmatrix} A_1 & * & * & * & * & * \\ & \ddots & * & * & * & * \\ & & A_s & * & * & * \\ & & & \lambda_{2s+1} & * & * \\ & & & & \ddots & * \\ & & & & & \lambda_n \end{pmatrix}.$$

再设 $A = P \text{diag}(|\lambda_1|, \dots, |\lambda_n|) Q$, 其中 P, Q 是正交方阵. 由置换可不妨设 $|\lambda_1| \geq \dots \geq |\lambda_n|$.

设 $A_i = \begin{pmatrix} a_i & b_i \\ c_i & d_i \end{pmatrix}$ ($i = 1, \dots, s$). 则由奇异值分解,

$$\text{tr}(A^T A) = \sum_{i=1}^n |\lambda_i|^2 \xrightarrow{\text{共轭虚特征值两两配对}} 2 \sum_{i=1}^s |\lambda_{2i-1}|^2 + \sum_{i=2s+1}^n \lambda_i^2.$$

又

$$\text{tr}(A^T A) = A \text{ 的所有元素平方和 } \geq \sum_{i=1}^s (a_i^2 + b_i^2 + c_i^2 + d_i^2) + \sum_{i=2s+1}^n \lambda_i^2.$$

由以上两式推出

$$2 \sum_{i=1}^s |\lambda_{2i-1}|^2 \geq \sum_{i=1}^s (a_i^2 + b_i^2 + c_i^2 + d_i^2).$$

但

$$2|\lambda_{2i-1}|^2 = 2(a_{2i-1}d_{2i-1} - b_{2i-1}c_{2i-1}) \leq a_i^2 + b_i^2 + c_i^2 + d_i^2,$$

这就迫使等号成立:

$$\operatorname{tr}(A^T A) = \sum_{i=1}^s (a_i^2 + b_i^2 + c_i^2 + d_i^2) + \sum_{i=2s+1}^n \lambda_i^2.$$

这说明 A 是准对角方阵. 再由

$$2 \sum_{i=1}^s |\lambda_{2i-1}|^2 = \sum_{i=1}^s (a_i^2 + b_i^2 + c_i^2 + d_i^2)$$

得到

$$0 = \sum_{i=1}^s (a_i^2 + b_i^2 + c_i^2 + d_i^2 - 2a_i d_i + 2b_i c_i) = \sum_{i=1}^s [(a_i - d_i)^2 + (b_i + c_i)^2] \implies \begin{cases} a_i = d_i \\ c_i = -b_i \end{cases}.$$

故每个 A_i 形如 $\begin{pmatrix} a_i & b_i \\ -b_i & a_i \end{pmatrix}$, 为规范方阵. 进而 A 是规范方阵. \square

习题 6.3.6 (2) 对于任意实数方阵 A , 存在规范的实数方阵 P_1, P_2 , 使得 $A = P_1 P_2$.

证明 设 A 的奇异值分解为 $A = P \begin{pmatrix} \Sigma & O \\ O & O \end{pmatrix} Q$, 令 $P_1 = PQ, P_2 = Q^T \begin{pmatrix} \Sigma & O \\ O & O \end{pmatrix} Q$ 即可. \square

习题 6.3.6 (3) 对于任意正交方阵 A , 存在对称的正交方阵 P_1, P_2 , 使得 $A = P_1 P_2$.

证明 由定理 6.6, 存在正交方阵 P , 使得

$$A = P \operatorname{diag} \left(\begin{pmatrix} \cos \theta_1 & \sin \theta_1 \\ -\sin \theta_1 & \cos \theta_1 \end{pmatrix}, \dots, \begin{pmatrix} \cos \theta_s & \sin \theta_s \\ -\sin \theta_s & \cos \theta_s \end{pmatrix}, I, -I \right) P^T.$$

令

$$B = \operatorname{diag} \left(\underbrace{\begin{pmatrix} 1 & \\ & -1 \end{pmatrix}, \dots, \begin{pmatrix} 1 & \\ & -1 \end{pmatrix}}_{s \uparrow}, I, I \right),$$

$$C = \text{diag} \left(\begin{pmatrix} \cos \theta_1 & \sin \theta_1 \\ \sin \theta_1 & -\cos \theta_1 \end{pmatrix}, \dots, \begin{pmatrix} \cos \theta_s & \sin \theta_s \\ \sin \theta_s & -\cos \theta_s \end{pmatrix}, I, -I \right).$$

取 $P_1 = PBP^T, P_2 = PCP^T$. 注意到 $\begin{pmatrix} 1 & \\ & -1 \end{pmatrix} \begin{pmatrix} \cos \theta_i & \sin \theta_i \\ \sin \theta_i & -\cos \theta_i \end{pmatrix} = \begin{pmatrix} \cos \theta_i & \sin \theta_i \\ -\sin \theta_i & \cos \theta_i \end{pmatrix}$, 则有 $A = P_1P_2$, 且 P_1, P_2 是对称正交方阵. \square

习题 6.3.8 (2) 设 B 是 A 的子矩阵, 则 $\|B\| \leq \|A\|$.

证明 由于与正交方阵相乘不改变原矩阵的算子范数, 所以通过置换相似可不妨设 B 在 A 的左上角. 那么

$$\|B\| = \| \begin{pmatrix} B & O \\ O & O \end{pmatrix} \| = \| \begin{pmatrix} I & O \\ O & O \end{pmatrix} A \begin{pmatrix} I & O \\ O & O \end{pmatrix} \| \leq \| \begin{pmatrix} I & O \\ O & O \end{pmatrix} \| \cdot \|A\| \cdot \| \begin{pmatrix} I & O \\ O & O \end{pmatrix} \| \leq \|A\|.$$

 \square

习题 6.3.8 (4) 设 A 是方阵, 则 A 的谱半径 $\rho(A) \leq \|A\|$.

证明 设 (λ, α) 是 A 的任意一个特征对, 则

$$A\alpha = \lambda\alpha \implies |\lambda| = \frac{\|A\alpha\|}{\|\alpha\|} \leq \max_{\alpha \neq 0} \frac{\|A\alpha\|}{\|\alpha\|} = \|A\|.$$

再由 λ 的任意性可知 $\rho(A) \leq \|A\|$. \square

习题 6.3.9 (4) 设 $p, q \in [1, +\infty]$ 满足 $\frac{1}{p} + \frac{1}{q} = 1$, 向量 $\mathbf{x} = (x_i) \in \mathbb{R}^n$ 和矩阵 $A = (a_{ij}) \in \mathbb{R}^{m \times n}$ 的 p 范数分别定义为

$$\|\mathbf{x}\|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}}, \quad \|A\|_p = \max_{\|\mathbf{x}\|_p=1} \|A\mathbf{x}\|_p.$$

证明: $\|A\|_p = \|A^T\|_q$.

证明 先证明如下引理:

引理 设 $\mathbf{x} \in \mathbb{R}^{n \times 1}, \mathbf{y} \in \mathbb{R}^{m \times 1}$, 则 $\|A\|_p = \max_{\substack{\|\mathbf{x}\|_p=1 \\ \|\mathbf{y}\|_q=1}} \mathbf{y}^T A \mathbf{x}$.

【引理的证明: 只需证 $\|A\mathbf{x}\|_p = \max_{\|\mathbf{y}\|_q=1} \mathbf{y}^T A \mathbf{x}$. 记 $\alpha = A\mathbf{x}, \beta = \mathbf{y}$, 即证 $\|\alpha\|_p = \max_{\|\beta\|_q=1} \beta^T \alpha$.

由离散形式的 Hölder 不等式有 $\beta^T \alpha \leq \|\alpha\|_p \|\beta\|_q = \|\alpha\|_p$. 设 $\alpha = \begin{pmatrix} a_1 & \cdots & a_m \end{pmatrix}^T, \beta = \begin{pmatrix} b_1 & \cdots & b_m \end{pmatrix}^T$. 则等号成立当且仅当 $\begin{pmatrix} a_1^p & \cdots & a_m^p \end{pmatrix}^T$ 与 $\begin{pmatrix} b_1^q & \cdots & b_m^q \end{pmatrix}^T$ 同向. 故能取得等号.】

因为 $\mathbf{y}^T A \mathbf{x} \in \mathbb{R}$, 所以 $\mathbf{y}^T A \mathbf{x} = (\mathbf{y}^T A \mathbf{x})^T = \mathbf{x}^T A^T \mathbf{y}$. 由引理即得 $\|A\|_p = \|A\|_q$. \square

推论 2.0.13 $\|A\|_\infty = \|A^T\|_1$.

习题 6.3.9 (5) (题干同习题 6.3.9 (4)) 证明: $\frac{\|A\|_1}{\sqrt{m}} \leq \|A\|_2 \leq \sqrt{n}\|A\|_1$.

证明 由习题 6.3.9 (3), 设 k 满足 $\sum_{i=1}^m |a_{ik}| = \max_{\|x\|_1=1} \|Ax\|_1$. 取 $x = e_k$, 则由 Cauchy-Schwarz 不等式得

$$\|A\|_2 \geq \|Ax\|_2 = \sqrt{\sum_{i=1}^m |a_{ik}|^2} \geq \frac{\sum_{i=1}^m |a_{ik}|}{\sqrt{\sum_{i=1}^m 1}} = \frac{\|A\|_1}{\sqrt{m}}.$$

由习题 6.3.8 (3) 及习题 6.3.9 (3) 得

$$\|A\|_2 \leq \|A\|_F \leq \sqrt{n \max_{1 \leq j \leq n} \sum_{i=1}^m a_{ij}^2} \leq \sqrt{n \left(\max_{1 \leq j \leq n} \sum_{i=1}^m |a_{ij}| \right)^2} = \sqrt{n}\|A\|_1.$$

□

习题 6.3.9 (6) (题干同习题 6.3.9 (4)) 证明: $\|A\|_2 = \|A^T\|_2 \leq \sqrt{\|A\|_p \cdot \|A\|_q}$.

证明 由正交相抵标准形可见 A 与 A^T 的奇异值序列相同, 故 $\|A\|_2 = \sigma_1(A) = \|A^T\|_2$.

引理 (1) 实方阵的实特征值一定有对应的实特征向量.

【引理 (1) 的证明: 设 A 是实方阵, (λ, α) 是 A 的一个特征对, $\lambda \in \mathbb{R}$. 则将 α 分解为 $\alpha = uiv$, 其中 u, v 均为实向量. 那么 $A(uiv) = \lambda(uiv) \implies Au = \lambda u, Av = \lambda v$. 取 u, v 中非零的向量即可作为所要的实特征向量.】

引理 (2) $\|A^T A\|_2 \leq \|A^T A\|_p$.

【引理 (2) 的证明: 注意到 $A^T A$ 作为实对称方阵的正交相似对角化形式即为奇异值分解 (在特征值沿主对角线降序排列的要求下), 因此若记 λ_{\max} 为 $A^T A$ 的最大特征值, 则 $\|A^T A\|_2 = \lambda_{\max}$. 由引理 (1), 设 α 是 $A^T A$ 的与 λ_{\max} 对应的实特征向量 (不妨设 $\|\alpha\|_p = 1$), 则

$$A^T A \alpha = \lambda_{\max} \alpha \implies \|A^T A \alpha\|_p = \|\lambda_{\max} \alpha\|_p = \lambda_{\max} = \|A^T A\|_2.$$

于是 $\|A^T A\|_p \geq \|A^T A \alpha\|_p = \|A^T A\|_2$.】

注意到 $A^T A$ 作为实对称方阵的正交相似对角化形式即为奇异值分解 (在特征值沿主对角线降序排列的要求下), 因此

$$\|A\|_2^2 = \sigma_1(A)^2 = \sigma_1(A^T A) = \|A^T A\|_2 \leq \|A^T A\|_p \leq \|A^T\|_p \cdot \|A\|_p \stackrel{\text{习题 6.3.9 (4)}}{=} \|A\|_p \|A\|_q.$$

□

习题 6.3.10 设 $A \in \mathbb{R}^{m \times n}$, $1 \leq k \leq \text{rank}(A)$. 证明:

$$\sigma_k(A) = \max_{X \in \mathbb{R}^{n \times k}} \frac{\sigma_k(AX)}{\sigma_1(X)}, \quad \sigma_{n+1-k}(A) = \min_{X \in \mathbb{R}^{n \times k}} \frac{\sigma_1(AX)}{\sigma_k(X)}.$$

证明 先证明如下引理.

引理 设 $A \in \mathbb{R}^{m \times n}$, 对任意 $1 \leq k \leq \text{rank}(A)$, 存在 $X \in \mathbb{R}^{n \times k}$, 使得 $\text{rank}(AX) = k$.

【引理的证明: 设置换方阵 $P \in \mathbb{R}^{n \times n}$ 使得 AP 的前 k 列线性无关. 再取 $X = P \begin{pmatrix} I_k \\ O \end{pmatrix}$, 则 $\text{rank}(AX) = k$.】

定理 (Courant–Fischer) 设 $A \in \mathbb{C}^{n \times n}$ 是 Hermite 方阵, $\lambda_k(A)$ 是 A 的第 k 大特征值. 则

$$\lambda_k(A) = \max_{\dim V=k} \min_{x \in V} R_A(x).$$

① 先证 $\sigma_k(A) = \max_{X \in \mathbb{R}^{n \times k}} \frac{\sigma_k(AX)}{\sigma_1(X)}$. 若 $\text{rank}(X) < k$, 则 $\text{rank}(AX) \leq \text{rank}(X) < k$, 从而 $\sigma_k(AX) = 0$. 由欲证的形式及引理可不妨设 X 是列满秩的且使得 $\text{rank}(AX) = k$. 在此假定下, 由

$$\sigma_k(AX) = \min_{\alpha \neq 0} \frac{\|AX\alpha\|}{\|\alpha\|}, \quad \sigma_1(X) = \max_{\alpha \neq 0} \frac{\|X\alpha\|}{\|\alpha\|}$$

可得

$$\frac{\sigma_k(AX)}{\sigma_1(X)} \leq \min_{\alpha \neq 0} \frac{\|AX\alpha\|}{\|X\alpha\|} \leq \sigma_k(A),$$

其中最后一个不等号是 Courant–Fischer 定理的推论: 对 $A \in \mathbb{R}^{m \times n}$, 有

$$\sigma_k(A) = \max_{\dim V=k} \min_{0 \neq x \in V} \frac{\|Ax\|}{\|x\|} \implies \min_{\alpha \neq 0} \frac{\|AX\alpha\|}{\|X\alpha\|} \leq \sigma_k(A).$$

只需说明存在列满秩矩阵 $X \in \mathbb{R}^{n \times k}$ 使得上式中的等号成立. 设 $A = P \begin{pmatrix} \Sigma & O \\ O & O \end{pmatrix} Q$ 是 A 的

奇异值分解, 则取 X 为 Q^T 的前 k 列, 有 $AX = P \begin{pmatrix} \Sigma & O \\ O & O \end{pmatrix} \begin{pmatrix} I_k \\ O \end{pmatrix} = P \begin{pmatrix} \Sigma \\ O \end{pmatrix}$. 由此可见

$\sigma_k(AX) = \sigma_k(A)$, 而由 $X = Q^T \begin{pmatrix} I_k \\ O \end{pmatrix}$ 可知 $\sigma_1(X) = 1$, 故等号能取到.

② 再证第二个等式. 仍设 A 的奇异值分解为 $A = P \begin{pmatrix} \Sigma & O \\ O & O \end{pmatrix} Q$, 其 Moore–Penrose 广

义逆为 $A^+ = Q^T \begin{pmatrix} \Sigma^{-1} & O \\ O & O \end{pmatrix} P^T$. 则

$$\sigma_{n+1-k}(A) = \sigma_k^{-1}(A^+) = \min_{Y \in \mathbb{R}^{m \times k}} \frac{\sigma_1(AA^+Y)}{\sigma_k(A^+Y)} \stackrel{X=A^+Y}{=} \min_{X \in \mathbb{R}^{n \times k}} \frac{\sigma_1(AX)}{\sigma_k(X)}.$$

□

习题 6.3.12 (2)(3) 给定 $A, B \in \mathbb{R}^{n \times n}$.

(2) 求正交方阵 P 使得 $\|PA - B\|_F$ 最小.

(3) 求实数 λ 和正交方阵 P 使得 $\|\lambda PA - B\|_F$ 最小.

解 (2) $\|PA - B\|_F^2 = \text{tr}((PA - B)^T(PA - B)) = \text{tr}(A^T A + B^T B - A^T P^T B - B^T P A) = \|A\|_F^2 + \|B\|_F^2 - 2 \text{tr}(B^T P A)$. 而 $\text{tr}(B^T P A) = \text{tr}(P A B^T)$, 设 $U \Sigma V$ 是 AB^T 的奇异值分解, 则

$$\text{tr}(P A B^T) = \text{tr}(P U \Sigma V) = \text{tr}(V P U \Sigma) \leq \text{tr}(\Sigma),$$

最后的不等号是由于 VPU 为正交方阵, 其对角元不超过 1. 当 $P = V^T U^T$ 时可取等号. 因此 $\|PA - B\|_F$ 的最小值为 $\sqrt{\|A\|_F^2 + \|B\|_F^2 - 2 \text{tr}(\Sigma)}$.

(3) 由所给形式不妨设 $\lambda \geq 0$. 同 (2) 有 $\|\lambda PA - B\|_F^2 = \text{tr}((\lambda PA - B)^T(\lambda PA - B)) = \text{tr}(\lambda^2 A^T A + B^T B - \lambda A^T P^T B - \lambda B^T P A) = \lambda^2 \|A\|_F^2 + \|B\|_F^2 - 2 \text{tr}(B^T P A) = \lambda^2 \|A\|_F^2 + \|B\|_F^2 - 2 \lambda \text{tr}(P A B^T)$. 设 $U \Sigma V$ 是 AB^T 的奇异值分解, 则

$$\text{tr}(P A B^T) = \text{tr}(P U \Sigma V) = \text{tr}(V P U \Sigma) \leq \text{tr}(\Sigma),$$

最后的不等号是由于 VPU 为正交方阵, 其对角元不超过 1. 当 $P = V^T U^T$ 时可取等号. 再令 $\lambda = \frac{\text{tr}(\Sigma)}{\|A\|_F^2}$ 就取得 $\|\lambda PA - B\|_F$ 的最小值 $\sqrt{\|B\|_F^2 - \frac{\text{tr}^2(\Sigma)}{\|A\|_F^2}}$. □

习题 6.3.13 设 A 是实数矩阵, $A = P \begin{pmatrix} \Sigma & O \\ O & O \end{pmatrix} Q$ 是奇异值分解, $X = Q^{-1} \begin{pmatrix} \Sigma^{-1} & O \\ O & O \end{pmatrix} P^{-1}$.

证明:

(1) X 是唯一的, 与正交方阵 P, Q 的选取无关.

(2) 若 $A = BC$, 其中 B, C 分别是列满秩和行满秩矩阵, 则 $X = C^T (CC^T)^{-1} (B^T B)^{-1} B^T$.

(3) A 的任意广义逆矩阵形如 $Y = Q^{-1} \begin{pmatrix} \Sigma^{-1} & Y_1 \\ Y_2 & Y_2 \Sigma Y_1 \end{pmatrix} P^{-1}$, 从而有 $\|Y\|_F \geq \|X\|_F$ 且

$\|Y\| \geq \|X\|$.

证明 (1) 设 $A = P_1 \begin{pmatrix} \Sigma & O \\ O & O \end{pmatrix} Q_1 = P_2 \begin{pmatrix} \Sigma & O \\ O & O \end{pmatrix} Q_2$ 是 A 的任意两种奇异值分解, 则

$$P_2^T P_1 \begin{pmatrix} \Sigma & \\ & O \end{pmatrix} = \begin{pmatrix} \Sigma & \\ & O \end{pmatrix} Q_2 Q_1^T.$$

将 $P_2^T P_1$ 对应于 $\begin{pmatrix} \Sigma & \\ & O \end{pmatrix}$ 分块为 $P_2^T P_1 = \begin{pmatrix} X_1 & X_2 \\ X_3 & X_4 \end{pmatrix}$, 将 $Q_2 Q_1^T$ 对应于 $\begin{pmatrix} \Sigma & \\ & O \end{pmatrix}$ 分块为 $Q_2 Q_1^T = \begin{pmatrix} Y_1 & Y_2 \\ Y_3 & Y_4 \end{pmatrix}$, 则

$$\begin{pmatrix} X_1 \Sigma & O \\ X_3 \Sigma & O \end{pmatrix} = \begin{pmatrix} \Sigma Y_1 & \Sigma Y_2 \\ O & O \end{pmatrix} \Rightarrow \begin{cases} X_1 \Sigma = \Sigma Y_1 \\ X_3 = O \\ Y_2 = O \end{cases}.$$

这说明 $P_2^T P_1 = \begin{pmatrix} X_1 & X_2 \\ O & X_4 \end{pmatrix}$ 是正交方阵, 通过方阵 X_1 所在行、列元素平方和相等 (都等于 X_1

阶数) 可得 $X_2 = O$. 同理可得 $Y_3 = O$. 为证明 $Q_1^{-1} \begin{pmatrix} \Sigma^{-1} & \\ & O \end{pmatrix} P_1^{-1} = Q_2^{-1} \begin{pmatrix} \Sigma^{-1} & \\ & O \end{pmatrix} P_2^{-1}$,

即证 $Q_2 Q_1^T \begin{pmatrix} \Sigma^{-1} & \\ & O \end{pmatrix} = \begin{pmatrix} \Sigma^{-1} & \\ & O \end{pmatrix} P_2^T P_1$. 由前面所得条件知这成立.

(2) 验证可知 $X := C^T (CC^T)^{-1} (B^T B)^{-1} B^T$ 满足

$$AXA = A, \quad XAX = X, \quad (AX)^T = AX, \quad (XA)^T = XA.$$

由例 (6.10) 知 X 即为 A 的 Moore-Penrose 广义逆.

(3) 设 $Y = Q^T \begin{pmatrix} Y_1 & Y_2 \\ Y_3 & Y_4 \end{pmatrix} P^T$ 是 A 的任一广义逆. 由定理 4.9 知, 这等价于 $AYA = A$

且 $YAY = Y$, 即

$$\begin{cases} \Sigma Y_1 \Sigma = \Sigma \\ Y_1 \Sigma Y_1 = Y_1 \\ Y_1 \Sigma Y_2 = Y_2 \\ Y_3 \Sigma Y_1 = Y_3 \\ Y_3 \Sigma Y_2 = Y_4 \end{cases} \Rightarrow \begin{cases} Y_1 = \Sigma^{-1} \\ Y_4 = Y_3 \Sigma Y_2 \end{cases}.$$

重新编号即得

$$Y = Q^T \begin{pmatrix} \Sigma^{-1} & Y_1 \\ Y_2 & Y_2 \Sigma Y_1 \end{pmatrix} P^T.$$

于是

$$\begin{aligned}
 \|Y\|_F &= \sqrt{\operatorname{tr}(Y^T Y)} = \sqrt{\operatorname{tr} \left(P \begin{pmatrix} \Sigma^{-1} & Y_2^T \\ Y_1^T & Y_1^T \Sigma Y_2^T \end{pmatrix} \begin{pmatrix} \Sigma^{-1} & Y_1 \\ Y_2 & Y_2 \Sigma Y_1 \end{pmatrix} P^T \right)} \\
 &= \sqrt{\operatorname{tr} \left(P^T P \begin{pmatrix} \Sigma^{-1} & Y_2^T \\ Y_1^T & Y_1^T \Sigma Y_2^T \end{pmatrix} \begin{pmatrix} \Sigma^{-1} & Y_1 \\ Y_2 & Y_2 \Sigma Y_1 \end{pmatrix} \right)} \\
 &= \sqrt{\operatorname{tr} \left(\begin{pmatrix} (\Sigma^{-1})^2 + Y_2^T Y_2 & \Sigma^{-1} Y_1 + Y_2^T Y_2 \Sigma Y_1 \\ Y_1^T \Sigma^{-1} + Y_1^T \Sigma Y_2^T Y_2 & Y_1^T Y_1 + Y_1^T \Sigma Y_2^T Y_2 \Sigma Y_1 \end{pmatrix} \right)} \\
 &= \sqrt{\operatorname{tr}((\Sigma^{-1})^2) + \operatorname{tr}(Y_2^T Y_2) + \operatorname{tr}(Y_1^T Y_1) + \operatorname{tr}((Y_2 \Sigma Y_1)^T (Y_2 \Sigma Y_1))} \\
 &\geq \sqrt{\operatorname{tr}((\Sigma^{-1})^2)} = \|X\|_F.
 \end{aligned}$$

而对于单位实向量 α ,

$$\|X\alpha\| = \|Q^{-1} \begin{pmatrix} \Sigma^{-1} & 0 \\ 0 & 0 \end{pmatrix} P^{-1} \alpha\| \stackrel{\beta=P^{-1}\alpha}{=} \left\| \begin{pmatrix} \Sigma^{-1} & 0 \\ 0 & 0 \end{pmatrix} \beta \right\| \leq \frac{1}{\sigma_r},$$

其中 $r = \operatorname{rank}(A)$, 并不妨设 $\operatorname{rank}(A) \geq 1$. 等号成立当且仅当 $\beta = e_r$ 即 $\alpha = P e_r$. 此时

$$\|Y\alpha\| = \left\| Q^{-1} \begin{pmatrix} \Sigma^{-1} & Y_1 \\ Y_2 & Y_2 \Sigma Y_1 \end{pmatrix} P^{-1} \alpha \right\| = \left\| \begin{pmatrix} \Sigma^{-1} & Y_1 \\ Y_2 & Y_2 \Sigma Y_1 \end{pmatrix} e_r \right\| \geq \frac{1}{\sigma_r} = \|X\|.$$

故 $\|Y\| \geq \|X\|$. □

习题 6.4.2 证明:

- (1) 任意 2 阶酉方阵形如 $\mu \begin{pmatrix} z & w \\ -\bar{w} & \bar{z} \end{pmatrix}$, 其中 $z, w, \mu \in \mathbb{C}$ 满足 $|z|^2 + |w|^2 = 1, |\mu| = 1$.
- (2) 任意 2 阶酉方阵形如 $\begin{pmatrix} 1 & 0 \\ 0 & e^{i\theta_1} \end{pmatrix} \begin{pmatrix} \cos \theta_2 & \sin \theta_2 \\ -\sin \theta_2 & \cos \theta_2 \end{pmatrix} \begin{pmatrix} e^{i\theta_3} & 0 \\ 0 & e^{i\theta_4} \end{pmatrix}$, 其中 $\theta_i \in \mathbb{R}, \forall i$.
- (3) 任意 2 阶规范方阵形如 $\lambda I + \mu P$, 其中 $\lambda, \mu \in \mathbb{C}$, P 是 Hermite 方阵.
- (4) 任意 2 阶规范方阵形如 $\lambda I + \mu Q$, 其中 $\lambda, \mu \in \mathbb{C}$, Q 是酉方阵.

习题 6.4.3 设 $a, b, c, d \in \mathbb{C}$. 证明: $\begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$ 与 $\begin{pmatrix} a & c \\ 0 & d \end{pmatrix}$ 酉相似当且仅当 $|b| = |c|$.

习题 6.4.5 (2) 求满足 $\operatorname{rank}(A - I) = 1$ 的所有 n 阶酉方阵 A .

解 由条件可设 $A = I + \alpha\beta^H$, $\|\alpha\| = 1$. 由 $AA^H = I$ 得

$$I + \beta\alpha^H + \alpha\beta^H + \|\beta\|^2 \alpha\alpha^H = I \implies \beta\alpha^H + \alpha\beta^H + \|\beta\|^2 \alpha\alpha^H = O.$$

下面用两种方法证明存在 $\lambda \in \mathbb{C}$ 使得 $\beta = \lambda\alpha$.

法一 不妨设 $\alpha = P e_1$, 其中 P 是 n 阶酉方阵 (若 $\alpha \neq e_1$, 则取 Householder 方阵 $H_v = I - \frac{2}{v^H v} v v^H$, 其中 $v = \alpha - e_1$). 再设 $\beta = P \tilde{\beta}$ (即取 $\tilde{\beta} = P^H \beta$), 代入上面的等式就得到

$$\tilde{\beta} e_1^H + e_1 \tilde{\beta}^H + \|\tilde{\beta}\|^2 e_1 e_1^H = O.$$

设 $\tilde{\beta} = (b_1 \ \cdots \ b_n)^T$, 通过计算上式左边方阵第一列可知, 对 $k = 2, \dots, n$, 有 $b_k = 0$. 故存在 $\lambda \in \mathbb{C}$, 使得 $\tilde{\beta} = \lambda e_1$, 即 $P^H \beta = \lambda P^H \alpha$, 也即 $\beta = \lambda \alpha$.

法二 回到最初条件 “ A 是酉方阵”:

$$A A^H = I \implies (I + \alpha \alpha^H) (I + \beta \beta^H) = I \implies \beta \alpha^H + \alpha \beta^H + \alpha \beta^H \beta \alpha^H = O,$$

$$A^H A = I \implies (I + \beta \beta^H) (I + \alpha \alpha^H) = I \implies \beta \alpha^H + \alpha \beta^H + \beta \beta^H \alpha = O,$$

对比两式可得

$$\beta \beta^H = \alpha \beta^H \beta \alpha^H \implies \beta \beta^H \beta = \alpha \beta^H \beta \alpha^H \beta \implies \beta = (\alpha^H \beta) \alpha.$$

于是 $\lambda + \bar{\lambda} + |\lambda|^2 = 0$, 即 $(\lambda + 1)(\bar{\lambda} + 1) = 1$, 也即 $|\lambda + 1| = 1$. 故 $\lambda = -1 + e^{i\theta}$, $A = I + (e^{i\theta} - 1)\alpha \alpha^H$, 其中 $\theta \in \mathbb{R}$, $\|\alpha\| = 1$. 反过来, 直接验证可知这样给出的 A 都满足 $A^H A = I$. 因此这样的 A 即为所有满足 $\text{rank}(A - I) = 1$ 的 n 阶酉方阵. \square

注 2.0.14 当 $\theta = \pi$ 时这就是 Householder 方阵.

习题 6.4.8 (3)(8) 设映射 $\rho: \mathbb{C}^{n \times n} \rightarrow \mathbb{R}^{2n \times 2n}$, $A + B i \mapsto \begin{pmatrix} A & B \\ -B & A \end{pmatrix}$, 其中 $A, B \in \mathbb{R}^{n \times n}$. 证明:

(3) $\rho(X)$ 与 $\text{diag}(X, \bar{X})$ 复相似.

(8) $\rho(X)$ 可实相似成规范方阵 $\iff X$ 可复相似成对角方阵.

证明 (3) 直接构造如下相似:

$$\begin{pmatrix} A & B \\ -B & A \end{pmatrix} \begin{pmatrix} I & iI \\ iI & I \end{pmatrix} = \begin{pmatrix} I & iI \\ iI & I \end{pmatrix} \begin{pmatrix} A + iB & \\ & A - iB \end{pmatrix}.$$

(8) 由习题 6.2.4, $\rho(X)$ 可实相似成规范方阵 $\iff \rho(X)$ 可以在 \mathbb{C} 上相似于对角方阵, 再由 (3), $\rho(X)$ 在 \mathbb{C} 上相似于 $\text{diag}(X, \bar{X})$, 于是又等价于 X 在 \mathbb{C} 上可相似对角化. \square

习题 6.4.12 设 $\lambda \in \mathbb{C}$ 是 $A \in \mathbb{C}^{n \times n}$ 的任意特征值, $x_1 \leq \dots \leq x_n$ 和 $y_1 \leq \dots \leq y_n$ 分别是 Hermite 方阵 $X = \frac{1}{2}(A + A^H)$ 和 $Y = \frac{1}{2i}(A - A^H)$ 的所有特征值. 证明:

$$x_1 \leq \text{Re}(\lambda) \leq x_n, \quad y_1 \leq \text{Im}(\lambda) \leq y_n.$$

证明 设 α 是 A 的对应于 λ 的特征向量, 则

$$A\alpha = \lambda\alpha \implies \alpha^H A \alpha = \lambda \alpha^H \alpha \implies \alpha^H A^H \alpha = \bar{\lambda} \alpha^H \alpha.$$

于是

$$\alpha^H X \alpha = \operatorname{Re}(\lambda) \alpha^H \alpha, \quad \alpha^H Y \alpha = \operatorname{Im}(\lambda) \alpha^H \alpha.$$

由 X 是 Hermite 方阵可设 $X = P \begin{pmatrix} x_1 & & \\ & \ddots & \\ & & x_n \end{pmatrix} P^H$, 其中 P 是酉方阵. 记 $\beta = P^H \alpha = (b_1 \ \cdots \ b_n)^T$, 则

$$\alpha^H X \alpha = \beta^H \begin{pmatrix} x_1 & & \\ & \ddots & \\ & & x_n \end{pmatrix} \beta = \sum_{i=1}^n x_i |b_i|^2 \in [x_1 \beta^H \beta, x_n \beta^H \beta].$$

再注意到 $\alpha^H \alpha = \beta^H \beta$ 即得证. 对 Y 同理. □

习题 6.4.14 设 $A, B \in \mathbb{R}^{n \times n}$. 证明: 若 A 与 B 酉相似, 则 A 与 B 正交相似.

证明 需要一些铺垫来引出酉方阵的 QS 分解.

定义 (1) 方阵 $A \in \mathbb{C}^{n \times n}$ 的 Descartes 分解定义为 $A = H + iK$, 其中 $H, K \in \mathbb{C}^{n \times n}$ 是 Hermite 方阵. 方阵 $H = \frac{A + A^H}{2}$ 称为 A 的 Hermite 部分, $K = \frac{A - A^H}{2i}$ 称为 A 的反 Hermite 部分.

引理 (2) 设 $A \in \mathbb{C}^{n \times n}$ 有 Descartes 分解 $A = H + iK$. 则 A 是规范方阵 $\iff HK = KH$.

【引理 (2) 的证明: 注意到

$$A^H A = (H - iK)(H + iK) = H^2 + K^2 + i(HK - KH),$$

$$A A^H = (H + iK)(H - iK) = H^2 + K^2 - i(HK - KH).$$

于是 $A A^H = A^H A \iff HK = KH.$ **】**

引理 (3) 设 $A \in \mathbb{C}^{n \times n}$ 是对称方阵. 则 A 是规范方阵 \iff 存在 n 阶正交方阵 Q 与 n 阶对角复方阵 D , 使得 $A = Q D Q^T$.

【引理 (3) 的证明: 设 $A = H + iK$ 是 A 的 Descartes 分解. 由 A 对称可得

$$H^T + iK^T = A^T = A = H + iK \implies H^T = H, K^T = K$$

又因为 $H = H^H$, 所以 H 是实对称方阵. 同理 K 也是实对称方阵. 由引理 (2), A 是规范方阵 $\iff HK = KH$. 又由习题 6.2.11, 因为 H 与 K 是乘积可交换的实对称方阵, 而实对称方阵的特征值均为实数, 所以存在正交方阵 Q 与对角实方阵 D_1, D_2 , 使得 $H = QD_1Q^T, K = QD_2Q^T$. 这样就有

$$A = H + iK = Q(D_1 + iD_2)Q^T = QDQ^T,$$

其中 $D = D_1 + iD_2$. 反之, 若 D 是对角方阵, 则 QDQ^T 是对称方阵, 从而是规范方阵.】

引理 (4) 设 $A \in \mathbb{C}^{n \times n}$ 是对称方阵. 则 A 是酉方阵 \iff 存在 n 阶正交方阵 Q 与酉对角方阵 $D = \text{diag}(e^{i\theta_1}, \dots, e^{i\theta_n})$, 使得每一个 $\theta_j \in [0, 2\pi)$ 且 $A = QDQ^T$.

【引理 (4) 的证明: 因为 A 是酉方阵, 所以由引理 (3) 有 $A = QDQ^T$, 从而 $D = Q^T A Q$ 是酉方阵. 因此 D 的元素均为模长为 1 的复数. 反过来直接验证即可.】

引理 (5) 设 V 是 n 阶对称酉方阵. 则存在多项式 p , 使得 $S = p(V)$ 是对称酉方阵, 且 $S^2 = V$.

【引理 (5) 的证明: 由引理 (4), 存在正交方阵 Q 与酉对角方阵 $D = \text{diag}(e^{i\theta_1}, \dots, e^{i\theta_n})$, 使得每一个 $\theta_j \in [0, 2\pi)$ 且 $V = QDQ^T$. 设 $E = \text{diag}(e^{i\frac{\theta_1}{2}}, \dots, e^{i\frac{\theta_n}{2}})$, 则可构造 Lagrange 插值多项式 p , 使得对每个 j 都有 $p(e^{i\theta_j}) = e^{i\frac{\theta_j}{2}}$, 所以 $p(D) = E$. 于是 $S = Qp(V)Q^T = QEQ^T$ 是对称酉方阵, 且 $S^2 = V$.】

定理 (6) (酉方阵的 QS 分解) 设 U 是 n 阶酉方阵. 则存在 n 阶正交方阵 Q 、 n 阶对称酉方阵 S 与多项式 p , 使得 $U = QS$, 且 $S = p(U^T U)$.

【定理 (6) 的证明: 因为 $U^T U$ 是对称酉方阵, 所以由引理 (5), 存在多项式 p , 使得 $S = p(U^T U)$ 是对称酉方阵, 且 $S^2 = U^T U$. 设 $Q = US^H = U\bar{S}$, 则 Q 是酉方阵且

$$Q^T Q = S^H U^T U S^H = S^H S^2 S^H = (S^H S)(S S^H) = I.$$

于是 $Q^T = Q^H$, 从而 Q 是实方阵, 进而是正交方阵, 且有 $U = (US^H)S = QS$.】

回到原命题上. 设 $A, B \in \mathbb{R}^{n \times n}$, U 是 n 阶酉方阵且使得 $A = UBU^H$. 则

$$UBU^H = A = \bar{A} = \overline{UBU^H} = \bar{U}B\bar{U}^T.$$

由 $U^H U = I$ 两边取共轭得 $U^T \bar{U} = I$. 故由上式可得 $U^T U B = B U^T U$, 即 B 与 $U^T U$ 乘积可交换. 由定理 (6) 知存在分解 $U = QS$, 其中 Q 是正交方阵, S 是对称酉方阵, 且存在多项式 p , 使得 $S = p(U^T U)$. 由 B 与 $U^T U$ 乘积可交换得 B 与 S 乘积可交换. 故

$$A = U B U^H = Q S B S^H Q^T = Q B S S^H Q^T = Q B Q^T,$$

即 A 与 B 正交相似. □

习题 7.1.7 (1) 设实数方阵 $A \neq -A^T$. 证明: 存在可逆实数方阵 P , 使得 $P^T A P$ 是上三角方阵.

(2) 设实数方阵 $A = -A^T$. 证明: 存在可逆实数方阵 P , 使得 $P^T A P = \begin{pmatrix} O & I_s & O \\ -I_s & O & O \\ O & O & O \end{pmatrix}$.

证明 (1) 将对第 (1) 问的证明建立在第 (2) 问的基础上. 当 A 为一阶方阵时结论已成立. 设对阶数不超过 $n-1$ ($n \geq 2$) 的实数方阵结论成立, 下面考虑 $A \in \mathbb{R}^{n \times n}$. 因为 $A^T \neq -A$, 所以存在 $i, j \in \{1, \dots, n\}$ 使得 $T_{ji}(1) A T_{ij}(1)$ 的主对角线元素不全为 0. 故可不妨设 $a_{11} \neq 0$. 对 $i = 2, \dots, n$, 依次通过左乘 $T_{i1} \begin{pmatrix} -a_{i1} \\ a_{11} \end{pmatrix}$ 及右乘 $T_{1i} \begin{pmatrix} -a_{1i} \\ a_{11} \end{pmatrix}$ 进行相合变换, 最后相合

成 $\begin{pmatrix} a_{11} & \\ & B \end{pmatrix}$ 的形式. 若 $B^T \neq -B$, 则由归纳假设可知存在可逆实数方阵 Q 使得 $Q^T B Q$

是上三角方阵, 从而 $P = \begin{pmatrix} 1 & \\ & Q \end{pmatrix}$ 为所求; 若 $B^T = -B$, 由 (2) 可知存在可逆实数方阵

Q 使得 $Q^T B Q = \text{diag} \left(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, O \right)$ (此处不妨设 B 的阶数至少为 2, 进

而 A 的阶数至少为 3), 设 $\begin{pmatrix} 1 & \\ & Q \end{pmatrix}^T A \begin{pmatrix} 1 & \\ & Q \end{pmatrix}$ 的第一行元素依次为 a_{11}, b_2, \dots, b_n , 取 ε

使得 $\varepsilon(a_{11} + b_3\varepsilon)(\varepsilon - b_2) \neq 0$, 通过先左乘 $T_{13}(\varepsilon)$ 、右乘 $T_{31}(\varepsilon)$, 再左乘 $T_{21} \begin{pmatrix} \varepsilon \\ -a_{11} + b_3\varepsilon \end{pmatrix}$,

右乘 $T_{12} \begin{pmatrix} \varepsilon \\ -a_{11} + b_3\varepsilon \end{pmatrix}$ 进行相合变换. 注意到此变换仅改变主对角线上第二个元素 (变为 $\frac{\varepsilon(\varepsilon - b_2)}{a_{11} + b_3\varepsilon}$), 此时得到的方阵右下角为 $n-1$ 阶非反对称方阵, 由归纳假设同样得证.

(2) 因为 A 是反对称实方阵, 由定理 6.8 可不妨设正交方阵 P_1 使

$$P_1^T A P_1 = \text{diag} \left(\begin{pmatrix} 0 & b_1 \\ -b_1 & 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 & b_s \\ -b_s & 0 \end{pmatrix}, O \right),$$

且其中 $b_i > 0$ ($i = 1, \dots, s$) (否则通过 $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ 置换相似即可). 设

$$P_2 = \text{diag} \left(\frac{1}{\sqrt{b_1}}, \frac{1}{\sqrt{b_1}}, \dots, \frac{1}{\sqrt{b_s}}, \frac{1}{\sqrt{b_s}}, I \right),$$

则

$$P_2^T P_1^T A P_1 P_2 = \text{diag} \left(\underbrace{\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}}_{s \uparrow}, O \right).$$

取置换方阵 P_3 , 使得 P_3^T 左乘 $P_2^T P_1^T A P_1 P_2$ 的作用为将 $(2k, 2k-1)$ 位置上的元素移至 $(s+k, k)$ 位置 ($k = 1, \dots, s$) 而其他元素保持不动. 于是 $P = P_1 P_2 P_3$ 即为所求可逆实数方阵. \square

习题 7.2.2 设 $A \in \mathbb{R}^{n \times n}$ 是正定的对称方阵. 分别求

$$\int_{\mathbf{x}^T A \mathbf{x} \leq 1} dx_1 \cdots dx_n \quad \text{和} \quad \int_{\mathbb{R}^n} e^{-\mathbf{x}^T A \mathbf{x}} dx_1 \cdots dx_n,$$

并将其表示成 A 的函数形式.

解 ① 由 A 是正定实对称方阵可设 $A = P^T \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} P$, 其中 P 为正交方阵, $\lambda_1, \dots, \lambda_n \in$

$\mathbb{R}_{>0}$. 记 $D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$, 则

$$\begin{aligned} \int_{\mathbf{x}^T A \mathbf{x} \leq 1} dx_1 \cdots dx_n &\stackrel{\mathbf{y}=P\mathbf{x}}{=} \int_{\mathbf{y}^T D \mathbf{y} \leq 1} dy_1 \cdots dy_n \stackrel{z_i = \sqrt{\lambda_i} y_i}{=} \frac{1}{\sqrt{\lambda_1 \cdots \lambda_n}} \int_{\sum_{i=1}^n z_i^2 \leq 1} dz_1 \cdots dz_n \\ &= \frac{1}{\sqrt{\det(A)}} \cdot n \text{ 维单位球体积} = \begin{cases} \frac{\pi^k a^{2k}}{k! \sqrt{\det(A)}}, & n = 2k, \\ \frac{2^k \pi^{k-1} a^{2k-1}}{(2k-1)!! \sqrt{\det(A)}}, & n = 2k-1. \end{cases} \end{aligned}$$

② 同①中换元可得

$$\begin{aligned} \int_{\mathbb{R}^n} e^{-\mathbf{x}^T A \mathbf{x}} dx_1 \cdots dx_n &= \frac{1}{\sqrt{\det(A)}} \int_{\mathbb{R}^n} e^{-\sum_{i=1}^n z_i^2} dz_1 \cdots dz_n = \frac{1}{\sqrt{\det(A)}} \left(\int_{\mathbb{R}} e^{-z^2} dz \right)^n \\ &= \sqrt{\frac{\pi^n}{\det(A)}}. \end{aligned}$$

□

习题 7.2.3 设 $0 \leq \theta_1 \leq \theta_2 \leq \theta_3 \leq \pi$. 证明:

$$\begin{pmatrix} 1 & \cos \theta_1 & \cos \theta_2 \\ \cos \theta_1 & 1 & \cos \theta_3 \\ \cos \theta_2 & \cos \theta_3 & 1 \end{pmatrix} \text{ 是正定的 } \iff \theta_3 < \theta_1 + \theta_2 < 2\pi - \theta_3.$$

证明 记题中所给矩阵为 G . 则 G 正定 $\iff G$ 是 \mathbb{R}^3 的标准内积在一组基 $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ 下的度量矩阵. 由 G 的对角元均为 1 知 $\mathbf{u}, \mathbf{v}, \mathbf{w}$ 均为单位向量, 进而 \mathbf{u} 与 \mathbf{v} 、 \mathbf{v} 与 \mathbf{w} 、 \mathbf{w} 与 \mathbf{u} 的夹角分别为 θ_1 、 θ_3 、 θ_2 . 于是原命题等价于证明:

单位向量组 $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\} \subset \mathbb{R}^3$ 是 \mathbb{R}^3 的基 $\iff \theta_3 < \theta_1 + \theta_2 < 2\pi - \theta_3$, 其中 $\theta_1, \theta_2, \theta_3$ 分别为 \mathbf{u} 与 \mathbf{v} 、 \mathbf{w} 与 \mathbf{u} 、 \mathbf{v} 与 \mathbf{w} 的夹角.

\Rightarrow : 由三面角公式,

$$\cos \theta_3 = \cos \alpha \sin \theta_1 \sin \theta_2 + \cos \theta_1 \cos \theta_2,$$

其中 α 为平面 $\text{Span}(\mathbf{u}, \mathbf{v})$ 与平面 $\text{Span}(\mathbf{w}, \mathbf{u})$ 形成的二面角. 由 $\cos \alpha \in (-1, 1]$ 及 $\sin \theta_1 \sin \theta_2 > 0$ 可得

$$-\sin \theta_1 \sin \theta_2 + \cos \theta_1 \cos \theta_2 < \cos \theta_3 \leq \sin \theta_1 \sin \theta_2 + \cos \theta_1 \cos \theta_2,$$

也即

$$\cos(\theta_1 + \theta_2) < \cos \theta_3 \leq \cos(\theta_1 - \theta_2).$$

- 若 $\theta_1 + \theta_2 \leq \theta_3$, 由 $\theta_3 \in (0, \pi]$ 可得 $\cos(\theta_1 + \theta_2) \geq \cos \theta_3$, 矛盾. 故 $\theta_1 + \theta_2 > \theta_3$.
- 若 $\theta_1 + \theta_2 \geq 2\pi - \theta_3$, 由 $2\pi > \theta_1 + \theta_2 \geq 2\pi - \theta_3 > \pi$ 得

$$\cos(\theta_1 + \theta_2) \geq \cos(2\pi - \theta_3) = \cos \theta_3,$$

矛盾. 故 $\theta_1 + \theta_2 < 2\pi - \theta_3$.

\Leftarrow : 由 $\theta_1 + \theta_2 > \theta_3$ 及 $\theta_3 \geq \theta_2 \geq \theta_1 \geq 0$ 知 $\theta_i > 0$, 因此 $\mathbf{u}, \mathbf{v}, \mathbf{w}$ 两两不共线. 由 $\theta_1 + \theta_2 + \theta_3 < 2\pi$ 知 $\mathbf{u}, \mathbf{v}, \mathbf{w}$ 不共面. 于是 $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ 线性无关, 构成 \mathbb{R}^3 的基. □

习题 8.1.2 (6) 判断下列集合是否是 \mathbb{R} 上的线性空间 (涉及到的加法和数乘运算都是通常意义下的加法和数乘运算): 收敛半径为 1 的实系数幂级数 $\sum_{n=0}^{\infty} a_n x^n$ 的全体.

解 不是, 因为该集合不存在加法单位元. □

习题 8.1.3 (5) 作为 \mathbb{F} 上的线性空间, 下列 U 是否 V 的子空间?

$$U = \{A \in V \mid A^H = A\}, V = \mathbb{C}^{n \times n}, \mathbb{F} = \mathbb{C}.$$

解 不是, 因为 U 对数乘不封闭. □

习题 8.1.5 证明: 当 $m \neq n$ 时, \mathbb{F} 上的线性空间 \mathbb{F}^m 与 \mathbb{F}^n 不同构.

证明 不妨设 $m < n$. 若存在同构 $f: \mathbb{F}^m \rightarrow \mathbb{F}^n$, 则 $f^{-1}(\mathbf{e}_1), \dots, f^{-1}(\mathbf{e}_n) \in \mathbb{F}^m$ 线性相关, 再由同构保加法和数乘得 $\mathbf{e}_1, \dots, \mathbf{e}_n \in \mathbb{F}^n$ 线性相关, 但这是不可能的. □

习题 8.2.2 (3) 判断以下 \mathbb{R} 上的线性空间 $\mathbb{R}[x]$ 的子集 S 是否是线性相关的:

$$S = \{x^i(1-x)^j \mid 0 \leq i < j \leq n\}, \text{ 其中 } n \text{ 是给定的正整数.}$$

解 ① 若 $n = 1$: $\{1-x\}$ 线性无关.

② 若 $n \geq 2$: 注意到 $x^r + x^{r-1}(1-x) = x^{r-1}$, 从而 $x^r(1-x)^s + x^{r-1}(1-x)^{s+1} = x^{r-1}(1-x)^s$, 故 S 线性相关. □

习题 8.2.2 (4) 判断以下 \mathbb{R} 上的线性空间 $\mathbb{R}[x]$ 的子集 S 是否是线性相关的:

$$S = \{(x-a_0)^n, (x-a_1)^n, \dots, (x-a_n)^n\}, \text{ 其中 } a_0, a_1, \dots, a_n \in \mathbb{R} \text{ 两两不同.}$$

解 线性无关. 设 $f(x) = \lambda_0(x-a_0)^n + \lambda_1(x-a_1)^n + \dots + \lambda_n(x-a_n)^n$ 为零函数, 对其求前 n 阶导数得到

$$\begin{pmatrix} 1 & 1 & \cdots & 1 \\ x-a_0 & x-a_1 & \cdots & x-a_n \\ \vdots & \vdots & \ddots & \vdots \\ (x-a_0)^n & (x-a_1)^n & \cdots & (x-a_n)^n \end{pmatrix} \begin{pmatrix} \lambda_0 \\ \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

由 a_0, a_1, \dots, a_n 两两不同及 Vandermonde 方阵性质得 $\lambda_0 = \lambda_1 = \dots = \lambda_n = 0$. 故 S 线性无关. □

习题 8.2.3 (1) 判断以下 \mathbb{R} 上的线性空间 $\mathcal{C}[0, 2\pi]$ 的子集 S 是否是线性相关的:

$$S = \{\sin(nx) \mid n \in \mathbb{N}^*\}.$$

解 线性无关. 对任意 $n \in \mathbb{N}^*$, 设 $f(x) = \lambda_1 \sin(x) + \dots + \lambda_{2n} \sin(2nx)$ 为零函数, 对其求第 $0, 2, \dots, 4n-2$ 阶导数得到

$$\begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & 4 & \cdots & 4n^2 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 4^{2n-1} & \cdots & (4n^2)^{2n-1} \end{pmatrix} \begin{pmatrix} \lambda_1 \sin(x) \\ \lambda_2 \sin(2x) \\ \vdots \\ \lambda_{2n} \sin(2nx) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

由 Vandermonde 方阵性质得 $\lambda_1 \sin(x) = \cdots = \lambda_{2n} \sin(2nx) = 0$, 从而 $\lambda_1 = \cdots = \lambda_{2n} = 0$. 由 n 的任意性知 S 线性无关. \square

习题 8.2.3 (3) 判断以下 \mathbb{R} 上的线性空间 $\mathcal{C}[0, 2\pi]$ 的子集 S 是否是线性相关的:

$$S = \{\sin(mx) \cos(nx) \mid m, n \in \mathbb{N}^*\}.$$

解 线性相关. 由 $\cos(3x) + \cos x = 2 \cos(2x) \cos x$ 得 $\sin x \cos(3x) + \sin x \cos x = \sin(2x) \cos(2x)$. \square

习题 8.2.3 (4) 判断以下 \mathbb{R} 上的线性空间 $\mathcal{C}[0, 2\pi]$ 的子集 S 是否是线性相关的:

$$S = \{\sin^m x \cos^n x \mid m, n \in \mathbb{N}^*\}.$$

解 线性相关. $\sin x \cos x = \sin x \cos x (\sin^2 x + \cos^2 x) = \sin^3 x \cos x + \sin x \cos^3 x$. \square

习题 8.2.7 设线性空间 V 中的向量组 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 满足每个 α_k 都不是 $\alpha_1, \alpha_2, \cdots, \alpha_{k-1}$ 的线性组合, $k = 1, 2, \cdots, n$. 证明: $\alpha_1, \alpha_2, \cdots, \alpha_n$ 是线性无关的.

证明 设 $\lambda_1 \alpha_1 + \lambda_2 \alpha_2 + \cdots + \lambda_n \alpha_n = \mathbf{0}$. 若 $\lambda_n \neq 0$, 则 α_n 可由 $\alpha_1, \cdots, \alpha_{n-1}$ 线性表出, 与题设矛盾, 故 $\lambda_n = 0$. 依此类推可得 $\lambda_n = \lambda_{n-1} = \cdots = \lambda_1 = 0$ (需注意题设中 $k = 1$ 的情形蕴含着 $\alpha_1 \neq \mathbf{0}$), 即 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 是线性无关的. \square

习题 8.2.9 设线性空间 V 的子集 $S_1 = \{\alpha_1, \alpha_2, \cdots, \alpha_n\}$, $S_2 = \{\alpha_1 + \alpha_2, \alpha_2 + \alpha_3, \cdots, \alpha_{n-1} + \alpha_n, \alpha_n + \alpha_1\}$. 证明或否定:

(1) 若 S_1 是线性相关的, 则 S_2 是线性相关的.

(2) 若 S_1 是线性无关的, 则 S_2 是线性无关的.

解 (1) 正确. 设 T_1 为 S_1 的极大线性无关组, 则 $|T_1| < n$. 因为 $S_2 \subset \text{Span}(S_1) \subset \text{Span}(T_1)$, 若 S_2 线性无关, 由 Steinitz 替换定理知 $|S_2| \leq |T_1| < n$, 矛盾. 故 S_2 线性相关.

(2) 当 n 为奇数时正确, 当 n 为偶数时错误.

① 若 n 为奇数. 设 $\lambda_1(\alpha_1 + \alpha_2) + \lambda_2(\alpha_2 + \alpha_3) + \cdots + \lambda_{n-1}(\alpha_{n-1} + \alpha_n) + \lambda_n(\alpha_n + \alpha_1) = \mathbf{0}$, 则合并同类项后 $(\lambda_n + \lambda_1)\alpha_1 + (\lambda_1 + \lambda_2)\alpha_2 + \cdots + (\lambda_{n-1} + \lambda_n)\alpha_n = \mathbf{0}$. 由 S_1 线性无关可知

$$\underbrace{\begin{pmatrix} 1 & & & 1 \\ & 1 & & \\ & & \ddots & \ddots \\ & & & 1 & 1 \end{pmatrix}}_{\text{记为 } A \in \mathbb{F}^{n \times n}} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

因为 $\det(A) = 1 + (-1)^{n-1} = 2 \neq 0$, 所以 A 可逆, 上述方程组只有零解 $\lambda_1 = \cdots = \lambda_n = 0$. 故 S_2 线性无关.

② 若 n 为偶数. 由 $[(\alpha_1 + \alpha_2) - (\alpha_2 + \alpha_3)] + \cdots + [(\alpha_{n-1} + \alpha_n) - (\alpha_n + \alpha_1)] = \mathbf{0}$ 知 S_2 线性相关. \square

习题 8.3.1 设 $T \subset S \subset V$ 满足 $\text{Span}(T) = \text{Span}(S)$, 并且不存在 $\hat{T} \subsetneq T$ 满足 $\text{Span}(\hat{T}) = \text{Span}(S)$. 证明: T 是 S 的极大线性无关组.

证明 若 T 不是 S 的极大线性无关组, 则由 $S \subset \text{Span}(S) = \text{Span}(T)$ 及定理 8.9 知 T 线性相关. 设 \hat{T} 是 T 的一个极大线性无关组, 则 $\hat{T} \subsetneq T$. 由 $T \subset \text{Span}(\hat{T})$ 知 $\text{Span}(S) = \text{Span}(T) \subset \text{Span}(\hat{T})$, 而 $\hat{T} \subset S$, 故 $\text{Span}(\hat{T}) = \text{Span}(S)$, 这与题设矛盾. 故 T 是 S 的极大线性无关组. \square

习题 8.3.5 (1) 证明: 若 $\text{Span}(S_1) = \text{Span}(S_2)$, 则 $\text{rank}(S_1) = \text{rank}(S_2)$.

证明 设 S_i 的极大线性无关组为 T_i ($i = 1, 2$). 由 $S_i \subset \text{Span}(T_i)$ 得 $\text{Span}(S_i) \subset \text{Span}(T_i)$, 再由定理 8.9 知 T_i 是 $\text{Span}(S_i)$ 的极大线性无关组, 从而 $\text{rank}(S_1) = \text{rank}(\text{Span}(S_1)) = \text{rank}(\text{Span}(S_2)) = \text{rank}(S_2)$. \square

习题 8.3.5 (2) 证明: 若 $S_1 \subset S_2$ 且 $\text{rank}(S_1) = \text{rank}(S_2) < \infty$, 则 $\text{Span}(S_1) = \text{Span}(S_2)$.

证明 设 $\alpha_1, \cdots, \alpha_r$ 与 β_1, \cdots, β_r 分别是 S_1 与 S_2 的极大线性无关组. 由 $S_1 \subset S_2$ 知 $\alpha_1, \cdots, \alpha_r$ 可由 β_1, \cdots, β_r 线性表出, 由 Steinitz 替换定理知 $\alpha_1, \cdots, \alpha_r$ 与 β_1, \cdots, β_r 等价, 故 $\text{Span}(S_1) = \text{Span}(S_2)$. \square

习题 8.3.5 (3) 举例: $S_1 \subset S_2$ 且 $\text{rank}(S_1) = \text{rank}(S_2)$, 但 $\text{Span}(S_1) \neq \text{Span}(S_2)$.

解 $S_1 = \{\mathbf{e}_2, \cdots, \mathbf{e}_n, \cdots\}, S_2 = \{\mathbf{e}_1, \cdots, \mathbf{e}_n, \cdots\}$. \square

习题 8.3.7 设 $A \in \mathbb{F}^{m \times n}, B \in \mathbb{F}^{n \times p}, \text{rank}(AB) = \text{rank}(A)$. 证明:

(1) 存在 $X \in \mathbb{F}^{p \times n}$, 使得 $ABX = A$.

(2) 对于任意 $Y \in \mathbb{F}^{q \times m}$, 都有 $\text{rank}(YAB) = \text{rank}(YA)$.

证明 (1) 设 T_1, T_2 分别为 A, AB 的列向量组的极大线性无关组, 则由矩阵乘法意义可知 T_2 可由 T_1 线性表出, 由 $\text{rank}(AB) = \text{rank}(A)$ 可知 $|T_2| = |T_1|$, 因此由 Steinitz 替换定理得 T_1 与 T_2 等价, 从而 A 的列向量组可由 AB 的列向量组线性表出, 即存在 $X \in \mathbb{F}^{p \times n}$, 使得 $ABX = A$.

(2) 由 (1) 知 A 与 AB 的列向量组等价, 而矩阵 Y 左乘的作用是行变换, 因此作用后 YA 与 YAB 的列向量组仍然等价, 可相互线性表出, 即 $\text{rank}(YAB) = \text{rank}(YA)$. \square

习题 8.3.9 设 $A \in \mathbb{F}^{m \times n}$, $B \in \mathbb{F}^{m \times p}$, $\text{rank}(A, B) = \text{rank}(A)$. 证明: 存在 $X \in \mathbb{F}^{n \times p}$, 使得 $AX = B$.

证明 设 T_1, T_2 分别为 $A, \begin{pmatrix} A & B \end{pmatrix}$ 的列向量组的极大线性无关组, 则由 $\text{rank}(A, B) = \text{rank}(A)$ 知 $|T_1| = |T_2|$. 又 T_1 可由 T_2 线性表出, 由 Steinitz 替换定理知 T_1 与 T_2 等价, 从而 A 与 $\begin{pmatrix} A & B \end{pmatrix}$ 的列向量组可相互线性表出, 即存在 $X \in \mathbb{F}^{n \times p}$, 使得 $AX = B$. \square

习题 8.3.10 设 $M = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \mathbb{F}^{m \times n}$, $A \in \mathbb{F}^{p \times q}$. 证明: $\text{rank}(M) \geq \text{rank}(A)$, 等号成立的充分必要条件是存在 $X \in \mathbb{F}^{q \times (n-q)}$ 和 $Y \in \mathbb{F}^{(m-p) \times p}$, 使得 $B = AX$, $C = YA$, $D = YAX$.

证明 两次运用例 (8.12) 可得 $\text{rank}(M) \geq \text{rank} \begin{pmatrix} A & B \end{pmatrix} \geq \text{rank}(A)$.

\Rightarrow : 两次运用习题 8.3.9 结论: 由 $\text{rank}(M) = \text{rank} \begin{pmatrix} A & B \end{pmatrix}$ 知, 存在 $Y \in \mathbb{F}^{(m-p) \times p}$ 使得 $Y \begin{pmatrix} A & B \end{pmatrix} = \begin{pmatrix} C & D \end{pmatrix}$; 由 $\text{rank} \begin{pmatrix} A & B \end{pmatrix} = \text{rank}(A)$ 知, 存在 $X \in \mathbb{F}^{q \times (n-q)}$, 使得 $B = AX$. 综合即得 $B = AX, C = YA, D = YAX$.

\Leftarrow : 由初等变换, $\text{rank} \begin{pmatrix} A & AX \\ YA & YAX \end{pmatrix} = \text{rank}(A)$. \square

习题 8.3.11 设 $A, B \in \mathbb{F}^{m \times n}$. 证明: $|\text{rank}(A) - \text{rank}(B)| \leq \text{rank}(A+B) \leq \text{rank}(A) + \text{rank}(B)$.

证明 ① 先证右半不等式. 设 A, B 的列向量组的极大线性无关组分别为 T_1, T_2 , 则由 $A+B$ 的列向量组 $\subset \text{Span}(T_1 \cup T_2)$ 得 $\text{rank}(A+B) \leq \text{rank}(T_1 \cup T_2) \leq |T_1| + |T_2| = \text{rank}(A) + \text{rank}(B)$.

② 由①, $\text{rank}(A) = \text{rank}((A+B) + (-B)) \leq \text{rank}(A+B) + \text{rank}(-B)$, 即 $\text{rank}(A) - \text{rank}(B) \leq \text{rank}(A+B)$, 再将 A, B 互换即得左半不等式. \square

习题 8.4.7 作为有理数域 \mathbb{Q} 上的线性空间, 实数域 \mathbb{R} 与复数域 \mathbb{C} 是否同构?

证明 同构. 我们证明更一般的结论: \mathbb{R} 与 \mathbb{R}^n 作为 \mathbb{Q} 上的线性空间同构.

① 先证明 \mathbb{R} 是 \mathbb{Q} 上的不可数无穷维线性空间. 用反证法, 若 \mathbb{R} 作为 \mathbb{Q} 上的线性空间维数至多可数, 设 r_1, r_2, r_3, \dots 是其一组 Hamel 基, 则由任意 $r \in \mathbb{R}$ 可由 r_1, r_2, r_3, \dots 有限线性表出知 $|\mathbb{R}| \leq |\mathbb{Q}^{\mathbb{N}}|$. 又由对角线方法可知 $\mathbb{Q}^{\mathbb{N}}$ 可数, 但这与 \mathbb{R} 不可数矛盾.

② 由①, 设 $r_\lambda (\lambda \in I)$ 是 \mathbb{R} 在 \mathbb{Q} 上的一组基, 其中指标集 I 不可数. 则 $\bigcup_{i=1}^n \{r_\lambda e_i \mid \lambda \in I\}$ 是 \mathbb{R}^n 在 \mathbb{Q} 上的一组基. 由 $|\mathbb{R}| = |\mathbb{R}^n|$ (可参见习题 8.4.8 证明 1 中的 ④) 得 $|\mathbb{R}^n| \leq |\mathbb{R}^n| = |\mathbb{R}|$,

而 $|\mathbb{R}| \leq |\mathbb{R}^n|$, 由 **Cantor-Schröder-Bernstein 定理** 就得到 $|\mathbb{R}| = |\mathbb{R}^n|$. 于是 $\aleph_0 < \dim_{\mathbb{Q}}(\mathbb{R}^n) \leq |\mathbb{R}^n| = |\mathbb{R}| = \aleph_1$. 由连续统假设, $\dim_{\mathbb{Q}}(\mathbb{R}^n) = \aleph_1 = \dim_{\mathbb{Q}}(\mathbb{R})$. 由定理 8.12, \mathbb{R} 与 \mathbb{R}^n 作为 \mathbb{Q} 上的线性空间同构. \square

习题 8.4.8 作为 \mathbb{R} 上的线性空间, 实系数幂级数 $\sum_{n=0}^{\infty} a_n x^n$ 的全体 $\mathbb{R}[[x]]$ 与实系数多项式的全体 $\mathbb{R}[x]$ 是否同构?

证明 1 不同构. 我们证明 $\dim(\mathbb{R}[[x]]) = \aleph_1$, 而 $\dim(\mathbb{R}[x]) = \aleph_0$.

① 先证明 $\mathbb{N}^{\mathbb{N}}$ 与区间 $(0, 1]$ 等势. 我们对 $(0, 1]$ 中的每一个实数采用二进制表示, 并处理为只有有限个 0 的形式 (如 $\frac{1}{2}$ 用 0.01 表示). 在这种约定下, $(0, 1]$ 中的实数的数位 1 都仅在小数点后出现.

• 对任意的 $x \in (0, 1]$, 通过以下方法映射到 $\mathbf{x} \in \mathbb{N}^{\mathbb{N}}$: 将 x 从左至右第一个数位 1 与小数点间相隔的位数作为 \mathbf{x} 的第一个分量, 然后依次将 x 的下一个 1 与上一个 1 间相隔的位数作为 \mathbf{x} 的下一个分量. 例如: $x = 0.100101 \cdots$ 对应 $\mathbf{x} = (0, 2, 1, \cdots)$.

• 容易看出, 上面定义的映射反过来也能将任意 $\mathbf{x} \in \mathbb{N}^{\mathbb{N}}$ 唯一地对应于 $x \in (0, 1]$. 例如: $\mathbf{x} = (2, 0, 2, 3, \cdots)$ 对应 $x = 0.00110010001 \cdots$.

于是我们证明了 $\mathbb{N}^{\mathbb{N}}$ 与区间 $(0, 1]$ 等势.

② 由 ① 及熟知的 $(0, 1]$ 与 \mathbb{R} 等势, 就得到 $\mathbb{N}^{\mathbb{N}}$ 与 \mathbb{R} 等势.

③ 再证明 $\mathbb{R}^{\mathbb{N}}$ 与 $\mathbb{N}^{\mathbb{N}}$ 等势. 对任意 $(x_1, x_2, x_3, \cdots) \in \mathbb{R}^{\mathbb{N}}$, 由 ② 可将每个 $x_i (i = 1, 2, 3, \cdots)$ 与 $\mathbf{x}_i = (x_{i1}, x_{i2}, x_{i3}, \cdots) \in \mathbb{N}^{\mathbb{N}}$ 一一对应. 再将每个 \mathbf{x}_i 的分量按行排列, 利用对角线方法从左上角开始沿各对角线逐个“列举”. 这样就得到 $\mathbb{R}^{\mathbb{N}}$ 与 $\mathbb{N}^{\mathbb{N}}$ 之间的双射.

④ 结合 ② ③ 就得到 $\mathbb{R}^{\mathbb{N}}$ 与 \mathbb{R} 等势. 故 $\dim(\mathbb{R}[[x]]) \leq |\mathbb{R}[[x]]| = |\mathbb{R}^{\mathbb{N}}| = |\mathbb{R}| = \aleph_1$.

⑤ 记 $\alpha_{\lambda} := \sum_{n=0}^{\infty} (\lambda x)^n$. 我们断言 $\{\alpha_{\lambda} \mid \lambda \in (0, 1)\} \subset \mathbb{R}[[x]]$ 线性无关. 否则, 则存在 $0 < \lambda_1 < \cdots < \lambda_n < 1$ 与不全为 0 的 $\mu_1, \cdots, \mu_n \in \mathbb{R}$, 使得 $\mu_1 \alpha_{\lambda_1} + \cdots + \mu_n \alpha_{\lambda_n} = 0$. 注意到由 $\lambda_i \in (0, 1)$ 可知 $\alpha_{\lambda_i} = \frac{1}{1 - \lambda_i x}$, 因而 $\frac{\mu_1}{1 - \lambda_1 x} + \cdots + \frac{\mu_n}{1 - \lambda_n x} = 0$. 不妨设 $\mu_1 \neq 0$, 则令 $x \rightarrow \frac{1}{\lambda_1}$ 可得 LHS $\rightarrow \infty$, 与 RHS 恒为 0 矛盾. 故 $\{\alpha_{\lambda} \mid \lambda \in (0, 1)\}$ 线性无关.

⑥ 由 ⑤ 可知 $\dim(\mathbb{R}[[x]]) \geq \#\{\alpha_{\lambda} \mid \lambda \in (0, 1)\} = |\mathbb{R}| = \aleph_1$.

⑦ 结合 ④ ⑥ 即得 $\dim(\mathbb{R}[[x]]) = \aleph_1$. 由 $1, x, x^2, \cdots$ 是 $\mathbb{R}[x]$ 的一组基知 $\dim(\mathbb{R}[x]) = \aleph_0$. 根据定理 8.12, 从 $\dim(\mathbb{R}[[x]]) \neq \dim(\mathbb{R}[x])$ 得 $\mathbb{R}[[x]]$ 与 $\mathbb{R}[x]$ 不同构. \square

证明 2 不同构. 用反证法. 已知 $1, x, x^2, \cdots$ 是 $\mathbb{R}[x]$ 的一组基, 若 $\mathbb{R}[x]$ 与 $\mathbb{R}[[x]]$ 同构, 则

$\mathbb{R}[[x]]$ 的基也可数, 不妨设为 $t_1 = \sum_{n=0}^{\infty} a_{1n}x^n, t_2 = \sum_{n=0}^{\infty} a_{2n}x^n, \dots$. 可以构造幂级数 $\sum_{n=0}^{\infty} b_n x^n$, 使得 $b_0 + b_1x$ 不能由 $a_{10} + a_{11}x$ 线性表出, $b_0 + b_1x + b_2x^2$ 不能由 $a_{10} + a_{11}x + a_{12}x^2$ 与 $b_{10} + b_{11}x + b_{12}x^2$ 线性表出, \dots , $\sum_{k=0}^n b_k x^k$ 不能由 $\sum_{k=0}^n a_{ik}x^k$ ($i = 1, \dots, n$) 线性表出. 由于 t_1, t_2, \dots 是 $\mathbb{R}[[x]]$ 的一组基, 可设 N 是 $\sum_{n=0}^{\infty} b_n x^n$ 的坐标中最大非 0 分量对应的基的下标. 则由前述构造可知, 当 $n \geq N$ 时, $\sum_{n=0}^{\infty} b_n x^n$ 不能由 t_1, t_2, \dots, t_N 线性表出, 矛盾. \square

习题 8.4.9 设 V 是 $\mathbb{F}^{n \times n}$ 的子空间, $\dim V \geq kn + 1$, $|\mathbb{F}| > n$. 证明: 存在 $A \in V$, 使得 $\text{rank}(A) \geq k + 1$.

证明 先证明一个引理.

引理 设 $\begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix} \in \mathbb{F}^{n \times n}$, $A_1 \in \mathbb{F}^{r \times r}$ 满足 $\text{rank} \begin{pmatrix} \lambda I_r + A_1 & A_2 \\ A_3 & A_4 \end{pmatrix} \leq r, \forall \lambda \in \mathbb{F}$. 若 $|\mathbb{F}| > r$, 则 $A_4 = O$.

【引理的证明: 假设 $A_4 \neq O$, 则 A_4 有 $s \geq 1$ 阶可逆子矩阵 B_4 . 设 $B = \begin{pmatrix} A_1 & B_2 \\ B_3 & B_4 \end{pmatrix}$ 是 A 的 $r + s$ 阶子矩阵. 设

$$f(x) = \det \begin{pmatrix} xI_r + A_1 & B_2 \\ B_3 & B_4 \end{pmatrix} \stackrel[\text{例 (2.19)}]{\text{Schur 公式}} \det(B_4) \det(xI_r + A_1 - B_2 B_4^{-1} B_3).$$

由于 $\deg(f) = r$, $|\mathbb{F}| > r$, 故存在 $\lambda \in \mathbb{F}$ 使得 $f(\lambda) \neq 0$, 与 $\text{rank} \begin{pmatrix} \lambda I_r + A_1 & A_2 \\ A_3 & A_4 \end{pmatrix} \leq r$ 矛盾.】

对 k 归纳. 当 $k = 1$ 时结论显然成立.

设结论对 $k - 1$ ($k \geq 1$) 成立, $\dim V \geq kn + 1$, 则 $m := \max_{A \in V} \text{rank}(A) \geq k$. 若 $m = k + 1$ 则结论已经成立. 下设 $m = k$, 任取 $1 \leq i \leq n$.

设 $U_i = \{A \in V \mid A\mathbf{e}_i = \mathbf{0}\}$, 则 $\dim U_i \geq (k - 1)n + 1$. 由归纳假设, 存在 $B_i \in U_i$ 使得 $\text{rank}(B_i) = k$. 设 $B_i = P_i \begin{pmatrix} I_k & \\ & O \end{pmatrix} Q_i$, 其中 $P_i, Q_i \in \mathbb{F}^{n \times n}$ 可逆. 由 $B_i \mathbf{e}_i = \mathbf{0}$ 即

$\begin{pmatrix} I_k & \\ & O \end{pmatrix} Q_i \mathbf{e}_i = \mathbf{0}$ 得 $Q_i \mathbf{e}_i \in \text{Span}(\mathbf{e}_{k+1}, \dots, \mathbf{e}_n)$. 对任意 $A \in V$, $\text{rank}(A + \lambda B_i) \leq k, \forall \lambda \in \mathbb{F}$.

设 $A = P_i \begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix} Q_i$, 其中 $A_1 \in \mathbb{F}^{k \times k}$. 由引理, $A_4 = O$, 从而 $A\mathbf{e}_i \in \text{Span}(P_i \mathbf{e}_1, \dots, P_i \mathbf{e}_k)$.

令 i 取遍 $1, \dots, n$ 就得到 $\dim V \leq kn$, 与假设矛盾. \square

习题 8.4.10 设 $\mathbb{F}^{n \times n}$ 的子空间 V 中任意两个方阵乘积可交换. 证明: $\dim V \leq \frac{n^2}{4} + 1$.

证明 由两两乘积可交换, 存在 \mathbb{F} 的扩域 \mathbb{K} 上可逆方阵 P 使得对任意 $X \in V$, $P^{-1}XP$ 都是上三角方阵 (习题 5.2.10 (3)). 故不妨设 V 中方阵都是上三角方阵.

对 n 归纳. 当 $n = 1$ 时结论成立.

设结论对 $n - 1$ 成立. 设 V_1 是 V 中形如 $A = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ & 0 & \cdots & 0 \\ & & \ddots & \vdots \\ & & & 0 \end{pmatrix}$ 的方阵构成的子空间,

V_2 是 V 中形如 $B = \begin{pmatrix} 0 & \cdots & 0 & b_1 \\ & \ddots & \vdots & \vdots \\ & & 0 & b_{n-1} \\ & & & b_n \end{pmatrix}$ 的方阵构成的子空间. 根据归纳假设,

$$\dim V \leq \dim V_i + \left\lfloor \frac{(n-1)^2}{4} \right\rfloor + 1, \quad i = 1, 2.$$

由 $AB = BA = O$, 得 $\dim V_1 + \dim V_2 \leq n$. 因此,

$$\dim V \leq \left\lfloor \frac{n}{2} \right\rfloor + \left\lfloor \frac{(n-1)^2}{4} \right\rfloor + 1 \leq \left\lfloor \frac{n^2 + 1}{4} \right\rfloor + 1 \leq \frac{n^2}{4} + 1.$$

归纳即得结论成立. □

习题 8.5.3 分别求 $V_1 \cap V_2$ 和 $V_1 + V_2$ 的基, 其中 V_1, V_2 分别是

- (1) $\{\cos^n x \mid n \in \mathbb{N}^*\}$ 和 $\{\sin^n x \mid n \in \mathbb{N}^*\}$ 生成的 $\mathcal{C}[0, 2\pi]$ 的子空间.
- (2) $\{\cos(nx) \mid n \in \mathbb{N}^*\}$ 和 $\{\sin(nx) \mid n \in \mathbb{N}^*\}$ 生成的 $\mathcal{C}[0, 2\pi]$ 的子空间.

解 (1) ① 对任意 $\lambda_1 \cos x + \lambda_2 \cos^2 x + \cdots + \lambda_n \cos^n x = \mu_1 \sin x + \mu_2 \sin^2 x + \cdots + \mu_n \sin^n x \in V_1 \cap V_2$, 在 $x \in [0, \frac{\pi}{2}]$ 上用 $\sqrt{1 - \sin^2 x}$ 代替 $\cos x$ 并将 $\sin x$ 看成新的变元, 从而等号两边都必须是 $\sin x$ 的多项式, 因此 $\lambda_1 = \lambda_3 = \cdots = \lambda_{2k+1} = \cdots = 0$. 故将该交空间中元素记为 $\lambda_2 \cos^2 x + \lambda_4 \cos^4 x + \cdots + \lambda_{2k} \cos^{2k} x$. 代入 $x = 0$ 得到 $\lambda_2 + \lambda_4 + \cdots + \lambda_{2k} = \mu_1 \sin 0 + \mu_2 \sin 0 + \cdots + \mu_n \sin 0 = 0$. 于是又可将该元素记为

$$\lambda_4 (\cos^4 x - \cos^2 x) + \lambda_6 (\cos^6 x - \cos^2 x) + \cdots + \lambda_{2k} (\cos^{2k} x - \cos^2 x).$$

对任意 $k \geq 2$, $\cos^{2k} x - \cos^2 x = (1 - \sin^2 x)^k - 1 + \sin^2 x \in V_1 \cap V_2$, 又由次数可知 $\cos^{2k} x - \cos^2 x$ ($k = 2, 3, 4, \cdots$) 线性无关, 故它们是 $V_1 \cap V_2$ 的基.

② $1, \cos x, \cos^2 x, \cos^3 x, \dots, \sin x, \sin^3 x, \sin^5 x, \dots$ 是 $V_1 + V_2$ 的基. 容易看出 $\cos x$ 与 $\sin x$ 的正整数幂次都可由它们生成. 下证它们线性无关: 若不然, 存在非零多项式 f 与 g 使得 $f(\cos x) = g(\cos x) \sin x$. 在 $x \in [0, \frac{\pi}{2}]$ 上用 $\sqrt{1 - \cos^2 x}$ 代替 $\sin x$ 并将 $\cos x$ 看成新的变元, 于是推出 $f = g = 0$, 矛盾.

(2) ① 对 $m, n \in \mathbb{N}^*$, $\int_0^{2\pi} \sin(mx) \cos(nx) dx = \frac{1}{2} \int_0^{2\pi} [\sin(m+n)x + \sin(m-n)x] dx = 0$. 设 $f(x) = \sum_{k=0}^n \lambda_k \sin(m_k x) = \sum_{k=0}^n \mu_k \cos(n_k x) \in V_1 \cap V_2$, 其中 $m_k, n_k \in \mathbb{N}^*$. 则

$$\int_0^{2\pi} f^2(x) dx = \int_0^{2\pi} \left(\sum_{k=0}^n \lambda_k \sin(m_k x) \right) \left(\sum_{k=0}^n \mu_k \cos(n_k x) \right) dx = 0 \implies f = 0.$$

故 $V_1 \cap V_2 = \{0\}$, \emptyset 是其基.

② 由习题 8.2.3 (1) (2) 知 $\{\sin(nx) \mid n \in \mathbb{N}^*\}$ 与 $\{\cos(nx) \mid n \in \mathbb{N}^*\}$ 作为 $\mathcal{C}[0, 2\pi]$ 的子集都线性无关, 再结合 ① 可知 $\{\sin(nx) \mid n \in \mathbb{N}^*\} \cup \{\cos(nx) \mid n \in \mathbb{N}^*\}$ 是 $V_1 + V_2$ 的基. \square

习题 8.5.4 设 $A = \begin{pmatrix} & I_{n-1} \\ 0 & \end{pmatrix}$, $V_1 = \{X \in \mathbb{R}^{n \times n} \mid AX = XA\}$, $V_2 = \{X \in \mathbb{R}^{n \times n} \mid A^T X = X A^T\}$.

证明: V_1, V_2 都是 $\mathbb{R}^{n \times n}$ 的子空间. 并分别求 $V_1, V_2, V_1 \cap V_2, V_1 + V_2$ 的维数.

证明 注意到 AX 对 X 的作用是每行向上移一行而尾行清零, XA 对 X 的作用是每列向右移一行而首列清零. 由此可知 $(2, 1), (3, 1), \dots, (n, 1)$ 位置元素所在对角线全为 0, 而其余每条对角线上元素相同. 故 $\dim(V_1) = n$. 又由 $A^T X = X A^T \iff X^T A = A X^T$ 可知 $\dim(V_2) = \dim(V_1) = n$. 因为 $V_1 \cap V_2 = \{\lambda I \mid \lambda \in \mathbb{R}\}$, 所以 $\dim(V_1 \cap V_2) = 1$. 由维数定理, $\dim(V_1 + V_2) = \dim(V_1) + \dim(V_2) - \dim(V_1 \cap V_2) = 2n - 1$. \square

习题 8.5.6 设 V_1, V_2 都是线性空间 V 的子空间. 证明: 若 $V_1 \cup V_2 = V_1 + V_2$, 则 $V_1 \subset V_2$ 或 $V_2 \subset V_1$.

证明 若 $V_1 \cup V_2 = V_1 + V_2$ 且 $V_1 \not\subset V_2$, 则存在 $\alpha \in V_1 \setminus V_2$. 对任意 $\beta \in V_2$, 因为 $\alpha + \beta \in V_1 + V_2 = V_1 \cup V_2$, 所以 $\alpha + \beta \in V_1$ 或 $\alpha + \beta \in V_2$. 若 $\alpha + \beta \in V_2$, 则 $\alpha = (\alpha + \beta) - \beta \in V_2$, 矛盾. 故 $\alpha + \beta \in V_1$. 从而 $\beta = (\alpha + \beta) - \alpha \in V_1$, 由 β 的任意性知 $V_2 \subset V_1$. \square

习题 8.5.7 (3) 设 V_1, V_2, W 都是线性空间 V 的子空间. 证明:

$$(V_1 \cap W) + (V_2 \cap W) = ((V_1 \cap W) + V_2) \cap W = (V_1 + (V_2 \cap W)) \cap W.$$

证明 先给出和空间的等价定义.

引理 $V_1 + V_2 = \{v_1 + v_2 \mid v_1 \in V_1, v_2 \in V_2\}$.

由此引理, $(V_1 \cap W) + (V_2 \cap W) = \{\alpha + \beta \mid \alpha \in V_1 \cap W, \beta \in V_2 \cap W\}$. 对 $\alpha \in V_1 \cap W$ 与 $\beta \in V_2 \cap W$, 有 $\alpha + \beta \in (V_1 \cap W) + V_2$, 而 $\alpha, \beta \in W \implies \alpha + \beta \in W$. 故 $(V_1 \cap W) + (V_2 \cap W) \subset ((V_1 \cap W) + V_2) \cap W$. 反过来, 对任意 $v \in ((V_1 \cap W) + V_2) \cap W$, 由 $v \in (V_1 \cap W) + V_2$ 可知 $v = \alpha + \beta$, 其中 $\alpha \in V_1 \cap W, \beta \in V_2$. 由 $v, \alpha \in W$ 得 $\beta = v - \alpha \in W$, 于是 $\beta \in V_2 \cap W$. 故 $((V_1 \cap W) + V_2) \cap W \subset (V_1 \cap W) + (V_2 \cap W)$. 综上可得第一个等号, 第二个等号同理. \square

习题 8.5.8 (3) 设 V_1, V_2, W 都是线性空间 V 的子空间. 证明:

$$((V_1 + W) \cap V_2) + W = (V_1 \cap (V_2 + W)) + W = (V_1 + W) \cap (V_2 + W).$$

证明 由习题 8.5.7 (3) 证明中的引理 (和空间的等价定义) 易得 $W = W \cap (V_2 + W)$. 利用习题 8.5.7 (3) 第一式与第三式相等可得

$$\begin{aligned} (V_1 + W) \cap (V_2 + W) &= (V_1 + (W \cap (V_2 + W))) \cap (V_2 + W) \\ &= (V_1 + (W \cap (V_2 + W))) \cap (V_2 + W) \\ &= (V_1 \cap (V_2 + W)) + (W \cap (V_2 + W)) \\ &= (V_1 \cap (V_2 + W)) + W. \end{aligned}$$

这就证明了第二个等号, 第一个等号同理. \square

习题 8.5.9 设 V_1, V_2, V_3 都是线性空间 V 的有限维子空间.

(1) 举例: $\dim(V_1 + V_2 + V_3) \neq \dim V_1 + \dim V_2 + \dim V_3 - \dim V_1 \cap V_2 - \dim V_1 \cap V_3 - \dim V_2 \cap V_3 + \dim V_1 \cap V_2 \cap V_3$.

(2) 推广定理 8.15, 给出 $\dim(V_1 + V_2 + V_3)$ 的正确公式.

解 (1) $V = \mathbb{R}^2$, $V_1 = \text{Span}((1, 0))$, $V_2 = \text{Span}((0, 1))$, $V_3 = \text{Span}((1, 1))$. 则 $V_1 + V_2 + V_3 = V$, 而 $V_1 \cap V_2 = V_1 \cap V_3 = V_2 \cap V_3 = V_1 \cap V_2 \cap V_3 = \mathbf{0}$. 代入维数验证可知等式不成立.

(2) 运用两个子空间的维数定理进行归纳, 可得 s 个子空间的维数定理:

$$\sum_{i=1}^s \dim(V_i) = \dim\left(\sum_{i=1}^s V_i\right) + \dim(V_2 \cap V_1) + \dim\left(V_3 \cap \sum_{i=1}^2 V_i\right) + \cdots + \dim\left(V_s \cap \sum_{i=1}^{s-1} V_i\right).$$

\square

习题 8.6.3 (1) 设 $A \in \mathbb{R}^{m \times n}$, \mathbb{R}^n 的子空间 V_1 由 A 的行向量生成, V_2 是线性方程组 $Ax = \mathbf{0}$ 的解空间. 证明: $\mathbb{R}^n = V_1 \oplus V_2$.

证明 由定理 4.8(2), $\dim(V_2) = n - \text{rank}(A) = n - \dim(V_1)$, 而由矩阵乘法可知 $V_1 \cap V_2 = \{\alpha \in \mathbb{R}^{n \times 1} \mid \alpha^T \alpha = 0\} = \{\mathbf{0}\}$, 故由定理 8.17 知 $V_1 + V_2$ 是直和. 又由维数定理知 $\dim(V_1 + V_2) = \dim(V_1) + \dim(V_2) - \dim(V_1 \cap V_2) = n$, 故 $V_1 \oplus V_2 = V_1 + V_2 = \mathbb{R}^n$. \square

习题 8.6.4 证明: $\sum_{i=1}^k V_i$ 是直和的充分必要条件是 $\left(\sum_{j=1}^{i-1} V_j\right) \cap V_i = \{\mathbf{0}\}, \forall i$.

证明 利用定理 8.17 及多个子空间的维数定理 (参见习题 8.5.9 (2)) 有

$$\begin{aligned} \sum_{i=1}^k V_i \text{ 是直和} &\iff \dim\left(\sum_{i=1}^k V_i\right) = \sum_{i=1}^k \dim(V_i) \\ &\iff \dim\left(V_i \cap \sum_{j=1}^{i-1} V_j\right) = 0, \forall i \\ &\iff V_i \cap \sum_{j=1}^{i-1} V_j = \{\mathbf{0}\}, \forall i. \end{aligned}$$

习题 8.6.6 求实线性空间 $V = \{f \in \mathbb{R}[x] \mid f(1) = 0\}$ 的子空间 $U = \{f \in V \mid f(-1) = 0\}$ 的一个补空间.

解 任意满足 $f(1) = 0$ 的 $f \in \mathbb{R}[x]$ 可表成

$$f(x) = (x-1) \sum_{k=0}^n a_k (x+1)^k = a_0(x-1) + (x-1) \sum_{k=1}^n a_k (x+1)^k$$

的形式, 且右式写法唯一. 因此 $W = \{k(x-1) \mid k \in \mathbb{R}\}$ 是 U 的一个补空间. \square

习题 8.6.9 (2)(3) 设 $A_i \in \mathbb{R}^{n \times n}$, U_i 是 A_i 的行向量生成的 \mathbb{R}^n 的子空间, W_i 是线性方程组 $A_i \mathbf{x} = \mathbf{0}$ 在 \mathbb{R}^n 中的解空间, $i = 1, 2$.

(2) $U_1 \cap U_2$ 是否一定是 $W_1 + W_2$ 的补空间?

(3) $U_1 + U_2$ 是否一定是 $W_1 \cap W_2$ 的补空间?

解 (2) 是. 将对第 (2) 问的证明建立在第 (3) 问的基础上. 同习题 8.6.3 (1) 可知 $(U_1 \cap U_2) \cap (W_1 + W_2) = \{\mathbf{0}\}$. 为证 $\mathbb{R}^n = (U_1 \cap U_2) \oplus (W_1 + W_2)$, 只需证 $\dim((U_1 \cap U_2) + (W_1 + W_2)) = n$, 也即证 $\dim(U_1 \cap U_2) + \dim(W_1 + W_2) = n$. 由维数定理,

$$\dim(U_1 + U_2) = \dim(U_1) + \dim(U_2) - \dim(U_1 \cap U_2), \quad (1)$$

$$\dim(W_1 + W_2) = \dim(W_1) + \dim(W_2) - \dim(W_1 \cap W_2). \quad (2)$$

由第 (3) 问结论,

$$\dim(U_1 + U_2) + \dim(W_1 \cap W_2) = n. \quad (3)$$

由 (1)+(2)-(3) 并移项得

$$\dim(W_1 + W_2) + \dim(U_1 \cap U_2) + n = \dim(U_1) + \dim(U_2) + \dim(W_1) + \dim(W_2). \quad (4)$$

而由定理 4.8(2) 知

$$\dim(U_1) + \dim(W_1) = n, \quad (5)$$

$$\dim(U_2) + \dim(W_2) = n, \quad (6)$$

将 (5)(6) 两式代入 (4) 中即得 $\dim(U_1 \cap U_2) + \dim(W_1 + W_2) = n$.

(3) 是. 一方面, $W_1 \cap W_2$ 中的向量与 U_1, U_2 中向量均正交, 从而也与 $U_1 + U_2$ 中向量正交; 另一方面, 与 $U_1 + U_2$ 中任意向量均正交的向量自然也与 U_1, U_2 中向量均正交, 从而是

$W_1 \cap W_2$ 中元素. 设 $\alpha_1, \dots, \alpha_r \in \mathbb{R}^n$ 是 $W_1 \cap W_2$ 的基, 取 $A = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_r \\ O \end{pmatrix} \in \mathbb{R}^{n \times n}$, 则 $W_1 \cap W_2$

是 A 的行向量生成的 \mathbb{R}^n 的子空间, $U_1 + U_2$ 是线性方程组 $Ax = 0$ 在 \mathbb{R}^n 中的解空间. 同习题 8.6.3 (1) 可知, $\mathbb{R}^n = (W_1 \cap W_2) \oplus (U_1 + U_2)$. \square

习题 8.7.3 设 U, W 都是线性空间 V 的子空间. 证明: $(U + W)/W$ 与 $U/(U \cap W)$ 同构.

证明 先证明线性空间的任意子空间都有补空间.

引理 (1) 设 V 是非零线性空间. 若 $I \subset S \subset V$ 满足 I 线性无关、 $\text{Span}(S) = V$, 则存在 V 的基 B 满足 $I \subset B \subset S$.

【引理 (1) 的证明: 在集合

$$X = \{A \mid I \subset A \subset S \text{ 且 } A \text{ 是线性无关的}\}$$

上定义偏序 $A \prec B \iff A \subset B$. 若 $Y = \{I_k \mid k \in K\}$ 是 X 中的一个链 (全序子集), 则并集 $U = \bigcup_{k \in K} I_k$ 线性无关且满足 $I \subset U \subset S$, 从而 $U \in X$, 即 X 中任意链都有在 X 中的上界. 根据 Zorn 引理, 存在 X 关于偏序 \prec 的一个极大元素 B . 若存在 $s \in S$ 不能被 B 中向量线性表出, 则 $B \cup \{s\} \subset S$ 线性无关, 与 B 是 X 中极大元矛盾, 因此 B 是 V 的基.】

引理 (2) 线性空间的任意子空间都有补空间.

【引理 (2) 的证明: 设 W 是线性空间 V 的子空间. 由引理 (1), 存在 W 的基 B_0 与 V 的基 $B \supset B_0$. 因此 $V = \text{Span}(B_0) \oplus \text{Span}(B \setminus B_0) = W \oplus \text{Span}(B \setminus B_0)$.】

由引理 (2) 可设 $U = (U \cap W) \oplus A$, $W = (U \cap W) \oplus B$, 则 $U/(U \cap W) \cong A$, 从而只需证 $A \oplus W = U + W$. 由 $A \cap W$ 为零空间知 $A + W$ 是直和. 设 T_1 是 $U \cap W$ 的基, T_2 是 A 的基, T_3 是 B 的基, 则 $A \oplus W = \text{Span}(T_1 \cup T_2 \cup T_3) = U + W$. \square

习题 8.7.4 设 V 是 \mathbb{R}^n 的 r 维子空间, $r < n$. 证明: 存在 $A \in \mathbb{R}^{(n-r) \times n}$ 使得 $V = \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} = \mathbf{0}\}$.

证明 设 $\alpha_1, \dots, \alpha_r \in \mathbb{R}^n$ 是 V 的基, 构造 $B = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_r \end{pmatrix} \in \mathbb{R}^{r \times n}$. 设 U 是线性方程组 $B\mathbf{x} = \mathbf{0}$ 的解空间, 则由习题 8.6.3 (1) 可知 $\mathbb{R}^n = V \oplus U$, $\dim(U) = n - \dim(V) = n - r$. 设 $\beta_1, \dots, \beta_{n-r}$ 是 U 的基, 取 $A = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_{n-r} \end{pmatrix} \in \mathbb{R}^{(n-r) \times n}$, 则线性方程组 $A\mathbf{x} = \mathbf{0}$ 的解空间维数为 r , 而 $\alpha_1, \dots, \alpha_r$ 均为该解空间中元素, 故 $V = \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} = \mathbf{0}\}$. \square

习题 8.7.5 旧 设 $A \in \mathbb{F}^{m \times n}$, U 是 A 的行向量生成的 \mathbb{F}^n 的子空间, $W = \{\mathbf{x} \in \mathbb{F}^n \mid A\mathbf{x} = \mathbf{0}\}$. 构造 $\mathbb{F}^n/W \rightarrow U$ 的同构映射.

解 设 $A = P \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} Q$, 其中 $P \in \mathbb{F}^{m \times m}$, $Q \in \mathbb{F}^{n \times n}$ 可逆. 又设 $Q = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = (\beta_1 \ \dots \ \beta_n)$.

注意到 A 的行空间即 $\begin{pmatrix} I_r & O \\ O & O \end{pmatrix} Q = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_r \\ O \end{pmatrix}$ 的行空间, 故 $U = \text{Span}(\alpha_1, \dots, \alpha_r)$. 由

$$A\mathbf{x} = \mathbf{0} \iff \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} Q\mathbf{x} = \mathbf{0} \iff Q\mathbf{x} \text{ 的前 } r \text{ 个分量均为 } 0 \text{ 知 } W = \text{Span}(\beta_{r+1}, \dots, \beta_n).$$

设 $V = \text{Span}(\beta_1, \dots, \beta_r)$, 则 $\mathbb{F}^n = V \oplus W$. 于是任意 $\gamma \in \mathbb{F}^n$ 可唯一表成 $\gamma = \mathbf{v} + \mathbf{w}$ 的形式, 其中 $\mathbf{v} \in V$, $\mathbf{w} \in W$. 由此可知映射

$$\rho: \mathbb{F}^n/W \rightarrow U, \quad \left[\sum_{i=1}^r \lambda_i \beta_i \right] \mapsto \sum_{i=1}^r \lambda_i \alpha_i$$

是良定的双射. 易见 ρ 保加法及数乘运算, 故 ρ 是 $\mathbb{F}^n/W \rightarrow U$ 的同构映射. \square

习题 8.7.7 设 I 是指标集合, V_i 是线性空间 V 的子空间, W_i 是 V_i 的子空间, $\forall i \in I$. 若 $U = \bigoplus_{i \in I} V_i$, $W = \bigoplus_{i \in I} W_i$, 则 U/W 与 $\prod_{i \in I} (V_i/W_i)$ 是否一定同构?

解 不一定. 反例如下: 取 $I = \mathbb{N}$, $V_i = \{ax^i \mid a \in \mathbb{R}\}$, $W_i = \{0\}$, 则 $U = \mathbb{R}[x]$, W 是零空间. 由 $\prod_{i \in I} (V_i/W_i) = \prod_{i \in I} V_i$ 与 $\mathbb{R}[[x]]$ 同构知它与 U/W 不同构. \square

习题 9.1.7 (3) 求满足 $\mathcal{A}(XY) = \mathcal{A}(YX)$ 的所有线性变换 $\mathcal{A} \in \mathcal{L}(\mathbb{F}^{n \times n})$.

解 以下默认 $|\mathbb{F}| > n$. 由习题 5.1.8 (2), 若非纯量方阵 $M \in \mathbb{F}^{n \times n}$ 满足 $\text{tr}(M) = 0$, 则存在可逆方阵 $X, Y \in \mathbb{F}^{n \times n}$, 使得 $M = XY - YX$, 因此有 $\mathcal{A}(M) = \mathcal{A}(XY - YX) = 0$. 对任意满足将迹为 0 的矩阵映为零矩阵的线性变换, 都有

$$\mathcal{A}(M_1 + M_2) = \text{tr}(M_1)\mathcal{A}(I_n) + \text{tr}(M_2)\mathcal{A}(I_n) = \mathcal{A}(M_1) + \mathcal{A}(M_2), \quad \forall M_1, M_2 \in \mathbb{F}^{n \times n},$$

$$\mathcal{A}(\lambda M) = \text{tr}(\lambda M)\mathcal{A}(I_n) = \lambda \text{tr}(M)\mathcal{A}(I_n) = \lambda \mathcal{A}(M), \quad \forall M \in \mathbb{F}^{n \times n}, \lambda \in \mathbb{F}.$$

故所有将迹为 0 的矩阵映为零矩阵的线性变换即为所求. \square

习题 9.2.5 设 $V = \mathbb{F}^{n \times n}$, $\mathcal{A} \in \mathcal{L}(V): X \mapsto aX + bX^T$, $a, b \in \mathbb{F}$.

(1) 若 \mathcal{A} 可逆, a, b 应满足的充分必要条件是什么?

(2) 求 \mathcal{A} 的最小多项式.

解 (1) 设 $A \in \mathbb{F}^{n^2 \times n^2}$ 是 \mathcal{A} 在 $\mathbb{F}^{n \times n}$ 的标准基下的矩阵表示, 则 A 的主对角线上有 n 个 $a+b$ (其余元素为 0), 这 n 个元素所在行、列其他元素均为 0. 其余 $n^2 - n$ 列可两两配对, 使得每组的两个列向量在相同两行处为 (a, b) 与 (b, a) , 而这两行的其他元素均为 0. 由定理 9.7,

$$\mathcal{A} \text{ 可逆} \iff A \text{ 可逆} \iff \begin{cases} a+b \neq 0 \\ (a, b) \text{ 与 } (b, a) \text{ 线性无关} \end{cases} \iff a \pm b \neq 0.$$

(2) ① 若 $b = 0$, 则 \mathcal{A} 的最小多项式为 $x - a$.

② 若 $b \neq 0$, 则 \mathcal{A} 的最小多项式次数至少为 2. 由 $(\mathcal{A} - a\mathcal{I})^2 = b^2\mathcal{I}$ 知 \mathcal{A} 的最小多项式为 $x^2 - 2ax + a^2 - b^2$. \square

习题 9.2.6 设 V 是 \mathbb{F} 上的有限维线性空间, $\mathcal{A} \in \mathcal{L}(V)$. 证明:

(1) 存在非零多项式 $p(x) \in \mathbb{F}[x]$, 使得 $p(\mathcal{A}) = 0$.

(2) \mathcal{A} 是单射 $\iff \mathcal{A}$ 是满射 $\iff \mathcal{A}$ 可逆.

(3) 当 V 是无限维时, (1) 和 (2) 是否仍然成立?

证明 (1) 设 $\dim(V) = n$, 则 $\dim(\mathcal{L}(V)) = n^2$. 由 $\mathcal{I}, \mathcal{A}, \mathcal{A}^2, \dots, \mathcal{A}^{n^2}$ 线性相关知存在 \mathcal{A} 的化零多项式 $p(x)$.

(2) ① \mathcal{A} 单射 $\implies \mathcal{A}$ 可逆: 设 $\alpha_1, \dots, \alpha_n$ 是 V 的基, 若 $\mathcal{A}\alpha_1, \dots, \mathcal{A}\alpha_n$ 线性相关, 则存在不全为 0 的 $\lambda_1, \dots, \lambda_n$ 使得 $\lambda_1\mathcal{A}\alpha_1 + \dots + \lambda_n\mathcal{A}\alpha_n = 0$, 从而 $\mathcal{A}(\lambda_1\alpha_1 + \dots + \lambda_n\alpha_n) = 0$, 由 \mathcal{A} 是单射知 $\lambda_1\alpha_1 + \dots + \lambda_n\alpha_n = 0$, 但这与 $\alpha_1, \dots, \alpha_n$ 线性无关矛盾. 故 $\mathcal{A}\alpha_1, \dots, \mathcal{A}\alpha_n$ 线性无关. 由定理 9.1(6) 知 \mathcal{A} 是满射, 从而可逆.

② \mathcal{A} 满射 $\implies \mathcal{A}$ 可逆: 由 \mathcal{A} 是满射可设 $\mathcal{A}\alpha_1, \dots, \mathcal{A}\alpha_n$ 是 V 的基, 则由定理 9.1(4) 知 $\alpha_1, \dots, \alpha_n$ 也是 V 的基. 若 \mathcal{A} 不是单射, 则存在 $\mathbf{0} \neq v = \mu_1\alpha_1 + \dots + \mu_n\alpha_n$, 使得 $\mathcal{A}(v) = \mu_1\mathcal{A}\alpha_1 + \dots + \mu_n\mathcal{A}\alpha_n = \mathbf{0}$, 于是 $\mu_1 = \dots = \mu_n = 0$, 但这与 $v \neq \mathbf{0}$ 矛盾.

(3) 不一定成立. 反例如下:

① 微分变换 $\mathcal{D}: \mathbb{F}[x] \rightarrow \mathbb{F}[x], p(x) \mapsto p'(x)$ 不存在化零多项式.

② 线性变换 $\mathcal{M}: \mathbb{F}[x] \rightarrow \mathbb{F}[x], f(x) \mapsto xf(x)$ 是单射, 但不是满射. \square

习题 9.2.8 (3)(4) 设 $V = \mathbb{F}[x]$, $\text{char } \mathbb{F} = 0$, $\mathcal{D}: V \rightarrow V, p(x) \mapsto p'(x)$ 是微分变换, $\mathcal{S}: V \rightarrow V, p(x) \mapsto \int_0^x p(t) dt$ 是积分变换.

(3) 设 $\mathcal{A} \in \mathcal{L}(V)$ 满足 $\mathcal{A} \circ \mathcal{D} = \mathcal{D} \circ \mathcal{A}$. 证明: 存在数列 $\{a_k\}_{k \in \mathbb{N}}$, 使得 $\mathcal{A}(f) = \sum_{k=0}^{\deg(f)} a_k \mathcal{D}^k(f)$.

(4) 设 $\mathcal{A} \in \mathcal{L}(V)$ 满足 $\mathcal{A} \circ \mathcal{S} = \mathcal{S} \circ \mathcal{A}$. 证明: 存在 $p \in \mathbb{F}[x]$, 使得 $\mathcal{A} = p(\mathcal{S})$.

证明 (3) 由 $\mathcal{D} \circ \mathcal{A}(1) = \mathcal{A} \circ \mathcal{D}(1) = \mathcal{A}(0) = 0$ 得 $\mathcal{A}(1)$ 为常数, 记为 λ . 由 $\mathcal{D} \circ \mathcal{A}(x) = \mathcal{A} \circ \mathcal{D}(x) = \mathcal{A}(1) = \lambda$ 得 $\mathcal{A}(x) = \lambda x + \mu$, 其中 $\mu \in \mathbb{F}$ 为常数, 再由 $\text{char } \mathbb{F} = 0$ 知 $\mu = 0$, $\mathcal{A}(x) = \lambda x$. 类似

可得 $\mathcal{A}(x^k) = \lambda x^k$ ($k \in \mathbb{N}$). 故 $\mathcal{A} = \lambda \mathcal{I}$. 取 $a_k = \begin{cases} \lambda, & k=0, \\ 0, & k \in \mathbb{N}^* \end{cases}$ 就有 $\mathcal{A}(f) = \sum_{k=0}^{\deg(f)} a_k \mathcal{D}^k(f)$.

(4) 先取 $p \in \mathbb{F}[x]$ 使得 $\mathcal{A}(1) = p(\mathcal{S})(1)$. 对此 p , 有 $\mathcal{A}(x^k) = \mathcal{A} \circ \mathcal{S}(kx^{k-1}) = k\mathcal{S} \circ \mathcal{A}(x^{k-1}) = \dots = k!\mathcal{S}^k \circ \mathcal{A}(1) = k!\mathcal{S}^k \circ p(\mathcal{S})(1) = k!p(\mathcal{S}) \circ \mathcal{S}^k(1) = p(\mathcal{S})(x^k)$. 故此 p 即为所求. \square

习题 9.3.2 设 S, T 都是 \mathbb{F} 上的有限维线性空间 V 的基, P 是从 S 到 T 的过渡矩阵, S^*, T^* 分别是 S, T 的对偶基. 求从 S^* 到 T^* 的过渡矩阵.

解 设 $S = \{\alpha_1, \dots, \alpha_n\}$, $T = \{\beta_1, \dots, \beta_n\}$, 从 S^* 到 T^* 的过渡矩阵为 Q . 则由

$$\begin{pmatrix} \beta_1 & \dots & \beta_n \end{pmatrix} = \begin{pmatrix} \alpha_1 & \dots & \alpha_n \end{pmatrix} P, \quad \begin{pmatrix} \beta_1^* & \dots & \beta_n^* \end{pmatrix} = \begin{pmatrix} \alpha_1^* & \dots & \alpha_n^* \end{pmatrix} Q$$

可得

$$Q^T \begin{pmatrix} \alpha_1^* \\ \vdots \\ \alpha_n^* \end{pmatrix} \begin{pmatrix} \alpha_1 & \dots & \alpha_n \end{pmatrix} P = \begin{pmatrix} \beta_1^* \\ \vdots \\ \beta_n^* \end{pmatrix} \begin{pmatrix} \beta_1 & \dots & \beta_n \end{pmatrix}.$$

再由 $\alpha_i^*(\alpha_j) = \delta_{ij} = \beta_i^*(\beta_j)$ 可得 $\left(\alpha_i^*(\alpha_j)\right)_{n \times n} = \left(\beta_i^*(\beta_j)\right)_{n \times n} = I$. 于是上式化简为 $Q^T P = I$, 故从 S^* 到 T^* 的过渡矩阵 $Q = P^{-T}$. \square

注 2.0.15 以上是形式上的矩阵写法, 本质上即多重求和.

习题 9.3.3 旧 设 $V = \mathbb{R}[x]$. 对于任意 $a \in \mathbb{R}$, 定义 $\pi_a \in V^* : f(x) \mapsto f(a)$. 记 $S = \{\pi_a \mid a \in \mathbb{R}\}$. 证明: $\text{Span}(S) \neq V^*$.

证明 考虑 $\pi_0 \circ \mathcal{D} \in V^* : f \mapsto f'(0)$. 若存在两两不同的 $a_1, \dots, a_n \in \mathbb{R}$ 与 $\lambda_1, \dots, \lambda_n \in \mathbb{R}$, 使得 $\pi_0 \circ \mathcal{D} = \sum_{i=1}^n \lambda_i \pi_{a_i}$, 则对任意 $p \in \mathbb{R}[x]$, 都有 $\sum_{i=1}^n \lambda_i p(a_i) = p'(0)$. 依次取 $p = x^{3k}$ ($k = 0, 1, \dots, n-1$), 得到 $\sum_{i=1}^n \lambda_i a_i^{3k} = 0$, 即

$$\begin{pmatrix} 1 & \cdots & 1 \\ a_1^3 & \cdots & a_n^3 \\ a_1^6 & \cdots & a_n^6 \\ \vdots & \ddots & \vdots \\ a_1^{3(n-1)} & \cdots & a_n^{3(n-1)} \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

由 a_1^3, \dots, a_n^3 两两不同, 利用 Vandermonde 方阵性质可得 $\lambda_1 = \dots = \lambda_n = 0$. 但 $\pi_0 \circ \mathcal{D}$ 显然不是零函数. 这说明 $\pi_0 \circ \mathcal{D} \notin \text{Span}(S)$, 从而 $\text{Span}(S) \neq V^*$. \square

习题 9.3.4 设 V 是 \mathbb{F} 上的有限维线性空间, $S = \{e_1, \dots, e_n\}$ 是 V 的基, $S^* = \{e_1^*, \dots, e_n^*\}$ 是 S 的对偶基, $\mathcal{A} \in \mathcal{L}(V)$, A 是 \mathcal{A} 在 S 下的矩阵表示.

(1) 设 $\alpha \in V$, 定义 $p_\alpha : V^* \rightarrow \mathbb{F}, f \mapsto f(\alpha)$. 证明: $p_\alpha \in V^{**}$. 求 p_α 在 $(S^*, 1)$ 下的矩阵表示.

(2) 定义自然映射 $\tau : \alpha \mapsto p_\alpha$. 证明: $\tau \in \mathcal{L}(V, V^{**})$ 并且 τ 是同构映射.

(3) 证明: $\text{tr}(A) = \sum_{i=1}^n e_i^*(\mathcal{A}e_i)$, 并且 $\text{tr}(A)$ 与 S 的选取无关. 从而可以定义 $\text{tr}(\mathcal{A}) = \text{tr}(A)$.

证明 (1) 对 $p_{\alpha_i} : V^* \rightarrow \mathbb{F}, f \mapsto f(\alpha_i)$ ($i = 1, 2$), 有

$$\begin{aligned} (p_{\alpha_1} + p_{\alpha_2})(f + g) &= p_{\alpha_1}(f + g) + p_{\alpha_2}(f + g) \\ &= (f + g)(\alpha_1) + (f + g)(\alpha_2) \\ &= f(\alpha_1) + f(\alpha_2) + g(\alpha_1) + g(\alpha_2) \\ &= (p_{\alpha_1} + p_{\alpha_2})(f) + (p_{\alpha_1} + p_{\alpha_2})(g), \quad \forall f, g \in V^*. \end{aligned}$$

对任意 $p_\alpha : V^* \rightarrow \mathbb{F}, f \mapsto f(\alpha)$, 有

$$p_\alpha(\lambda f) = (\lambda f)(\alpha) = \lambda f(\alpha) = \lambda p_\alpha(f), \quad \forall \lambda \in \mathbb{F}, \forall f \in V^*.$$

故 $p_\alpha \in V^{**}$, 其在 $(S^*, 1)$ 下的矩阵表示为 $(e_1^*(\alpha) \cdots e_n^*(\alpha))$.

(2) 因为

$$\tau(\alpha + \beta) = p_{\alpha+\beta} = p_\alpha + p_\beta = \tau(\alpha) + \tau(\beta), \quad \forall \alpha, \beta \in V,$$

$$\tau(\lambda\alpha) = p_{\lambda\alpha} = \lambda p_\alpha = \lambda \tau(\alpha), \quad \forall \lambda \in \mathbb{F}, \forall \alpha \in V,$$

所以 $\tau \in \mathcal{L}(V, V^{**})$. 对于 $\alpha = x_1 e_1 + \cdots + x_n e_n$, 其像 $\tau(\alpha) = p_\alpha$ 在 $(S^*, 1)$ 下的坐标为 $(x_1 \cdots x_n)^\top$, 由此可见 τ 是双射, 从而是同构映射.

(3) 设 $A = (a_{ij})_{n \times n}$, 则

$$\sum_{i=1}^n e_i^*(\mathcal{A}(e_i)) = \sum_{i=1}^n e_i^*(Ae_i) = \sum_{i=1}^n e_i^*(A \text{ 的第 } i \text{ 列}) = \sum_{i=1}^n a_{ii} = \text{tr}(A).$$

设 \tilde{S} 是 V 的另一组基, \tilde{A} 是 \mathcal{A} 在 \tilde{S} 下的矩阵表示, 则由定理 9.3 可知 \tilde{A} 与 A 相似, 从而 $\text{tr}(\tilde{A}) = \text{tr}(A)$, 即 $\text{tr}(A)$ 与 S 的选取无关. \square

习题 9.3.5 对于下列线性变换 $\mathcal{A} \in \mathcal{L}(V)$, 求 $\text{tr } \mathcal{A}$, 其中 $P, Q \in \mathbb{F}^{n \times n}$ 是给定的.

(1) $V = \mathbb{F}^{n \times n}$, $\mathcal{A}(X) = PXQ$.

(2) $V = \{X \in \mathbb{F}^{n \times n} \mid \text{tr}(X) = 0\}$, $\mathcal{A}(X) = PXP^{-1}$.

(3) $V = \{X \in \mathbb{F}^{n \times n} \mid X^\top = X\}$, $\mathcal{A}(X) = PXP^\top$.

(4) $V = \{X \in \mathbb{F}^{n \times n} \mid X^\top = -X\}$, $\mathcal{A}(X) = PXP^\top$.

解 (1) 用 vec 表示将矩阵列向量竖直叠放成新向量的算子. 设 $Q = (q_{ij})_{n \times n} = (\mathbf{q}_1 \cdots \mathbf{q}_n)$, $X = (\mathbf{x}_1 \cdots \mathbf{x}_n)$, 则 PXQ 的第 k 列是

$$PX\mathbf{q}_k = P \sum_{i=1}^n q_{ik} \mathbf{x}_i = (q_{1k}P \cdots q_{nk}P) \text{vec } X = (\mathbf{q}_k^\top \otimes P) \text{vec } X.$$

把这些向量竖直地叠放在一起, 就得到

$$\text{vec}(PXQ) = \begin{pmatrix} \mathbf{q}_1^\top \otimes P \\ \vdots \\ \mathbf{q}_n^\top \otimes P \end{pmatrix} \text{vec } X = (Q^\top \otimes P) \text{vec } X.$$

因此 \mathcal{A} 在 $E_{11}, \cdots, E_{n1}, \cdots, E_{1n}, \cdots, E_{nn}$ 下的矩阵表示为 $Q^\top \otimes P$. 由习题 9.3.4 (3) 知 $\text{tr}(\mathcal{A}) = \text{tr}(Q^\top \otimes P) = \text{tr}(Q^\top) \text{tr}(P) = \text{tr}(P) \text{tr}(Q)$ (第二个等号由相似上三角化易得).

(2) 选定 V 的基为 $E_{11} - E_{nn}, E_{21}, \dots, E_{n1}, E_{12}, E_{22} - E_{nn}, \dots, E_{n2}, \dots, E_{1n}, \dots, E_{1,n-1}$. 设 \mathcal{A} 在这组基下的矩阵表示为 $A \in \mathbb{F}^{(n^2-1) \times (n^2-1)}$. 结合第 (1) 问可注意到

$A = (P^{-\top}) \otimes P$ 的第 $1, n+2, 2n+3, \dots, n^2$ 列依次减去第 n^2 列后
所得 n^2 阶方阵的第 $n^2 - 1$ 个顺序主子矩阵.

若记 $P^{-\top} \otimes P$ 最后一列的第 $1, n+2, 2n+3, \dots, n^2$ 个元素为 $a_1, a_2, a_3, \dots, a_n$, 则

$$\sum_{i=1}^n a_i = \operatorname{tr}(PE_{nn}P^{-1}) = \operatorname{tr}(E_{nn}) = 1.$$

因此

$$\operatorname{tr}(\mathcal{A}) = \operatorname{tr}(A) = \operatorname{tr}(P^{-\top} \otimes P) - \sum_{i=1}^n a_i = \operatorname{tr}(P) \operatorname{tr}(P^{-1}) - 1.$$

(3) 记 $P = (p_{ij})_{n \times n}$, 取 V 的基 $\{E_{ii} \mid 1 \leq i \leq n\} \cup \{E_{ij} + E_{ji} \mid 1 \leq i < j \leq n\}$, 其对偶基满足

$$E_{ii}^*(X) = \operatorname{tr}(E_{ii}X), \quad (E_{ij} + E_{ji})^*(X) = \operatorname{tr}(E_{ij}X).$$

因此,

$$\begin{aligned} \operatorname{tr}(\mathcal{A}) &= \sum_{i=1}^n E_{ii}^*(\mathcal{A}E_{ii}) + \sum_{1 \leq i < j \leq n} (E_{ij} + E_{ji})^*(\mathcal{A}(E_{ij} + E_{ji})) \\ &= \sum_{i=1}^n \operatorname{tr}(E_{ii}PE_{ii}P^{\top}) + \sum_{1 \leq i < j \leq n} \operatorname{tr}(E_{ij}P(E_{ij} + E_{ji})P^{\top}) \\ &= \sum_{i=1}^n p_{ii}^2 + \sum_{1 \leq i < j \leq n} (p_{ji}p_{ij} + p_{jj}p_{ii}) \\ &= \frac{1}{2} \sum_{i,j=1}^n (p_{ji}p_{ij} + p_{jj}p_{ii}) \\ &= \frac{\operatorname{tr}(P^2) + (\operatorname{tr}(P))^2}{2}. \end{aligned}$$

(4) 记 $P = (p_{ij})_{n \times n}$, 取 V 的基 $\{E_{ij} - E_{ji} \mid 1 \leq i < j \leq n\}$, 其对偶基满足

$$(E_{ij} - E_{ji})^*(X) = -\frac{1}{2} \operatorname{tr}(E_{ij}X).$$

因此,

$$\begin{aligned}\operatorname{tr}(\mathcal{A}) &= \sum_{1 \leq i < j \leq n} (E_{ij} - E_{ji})^* (\mathcal{A}(E_{ij} - E_{ji})) \\ &= -\frac{1}{2} \sum_{1 \leq i < j \leq n} \operatorname{tr} (E_{ij} P (E_{ij} - E_{ji}) P^T) \\ &= -\frac{1}{2} \sum_{1 \leq i < j \leq n} (p_{ji} p_{ij} - p_{jj} p_{ii}) \\ &= \frac{(\operatorname{tr}(P))^2 - \operatorname{tr}(P^2)}{4}.\end{aligned}$$

□

注 2.0.16 本题 (3)(4) 问均默认 $\operatorname{char} \mathbb{F} \neq 2$.

习题 9.3.6 设 V 是 \mathbb{F} 上的线性空间, S 是 V 的任意非空子集.

$$\operatorname{Ann}(S) = \{f \in V^* \mid f(x) = 0, \forall x \in S\}$$

称为 S 的零化子. 证明:

- (1) $\operatorname{Ann}(S)$ 是 V^* 的子空间, 并且 $\operatorname{Ann}(S) = \operatorname{Ann}(\operatorname{Span}(S))$.
- (2) 当 V 是有限维时, $\dim \operatorname{Ann}(S) = \dim V - \operatorname{rank}(S)$.
- (3) 对于 V 的任意子空间 V_1, V_2 , 有

$$\operatorname{Ann}(V_1 \cap V_2) = \operatorname{Ann}(V_1) + \operatorname{Ann}(V_2), \quad \operatorname{Ann}(V_1 + V_2) = \operatorname{Ann}(V_1) \cap \operatorname{Ann}(V_2).$$

- (4) 若 $V = V_1 \bigoplus V_2$, 则 $V^* = \operatorname{Ann}(V_1) \bigoplus \operatorname{Ann}(V_2)$.

证明 (1) 因为对任意 $\lambda \in \mathbb{F}$ 与 $f, g \in \operatorname{Ann}(S)$ 都有 $f + g \in \operatorname{Ann}(S)$, $\lambda f \in \operatorname{Ann}(S)$, 所以 $\operatorname{Ann}(S) \leq V^*$. 由定义知 $\operatorname{Ann}(\operatorname{Span}(S)) \subset \operatorname{Ann}(S)$. 又对任意 $f \in \operatorname{Ann}(S)$, 对任意 $x = \lambda_1 \alpha_1 + \cdots + \lambda_k \alpha_k \in \operatorname{Span}(S)$, 其中 $\lambda_1, \cdots, \lambda_k \in \mathbb{F}$, $\alpha_1, \cdots, \alpha_k \in S$, 都有 $f(x) = \lambda_1 f(\alpha_1) + \cdots + \lambda_k f(\alpha_k) = 0$, 即 $f(x) = 0, \forall x \in \operatorname{Span}(S)$. 故 $\operatorname{Ann}(S) \subset \operatorname{Ann}(\operatorname{Span}(S))$, 从而 $\operatorname{Ann}(S) = \operatorname{Ann}(\operatorname{Span}(S))$.

(2) **法一** 设 e_1, \cdots, e_n 是 V 的基, e_1^*, \cdots, e_n^* 是其对偶基. 设 S 的极大线性无关组的坐标为 $\alpha_1, \cdots, \alpha_r$, $A = (\alpha_1 \cdots \alpha_r)$. 对于 $f = \lambda_1 e_1^* + \cdots + \lambda_n e_n^* \in V^*$,

$$\begin{aligned}f \in \operatorname{Ann}(S) &\iff \underbrace{\begin{pmatrix} \lambda_1 e_1^* & \cdots & \lambda_n e_n^* \end{pmatrix} \begin{pmatrix} e_1 & \cdots & e_n \end{pmatrix}}_{(\lambda_1 \cdots \lambda_n)} \underbrace{\begin{pmatrix} \alpha_1 & \cdots & \alpha_r \end{pmatrix}}_A = O \\ &\iff \begin{pmatrix} \lambda_1 & \cdots & \lambda_n \end{pmatrix} A = O.\end{aligned}$$

由定理 4.8(2) 即得 $\dim \operatorname{Ann}(S) = n - \operatorname{rank}(A) = \dim V - \operatorname{rank}(S)$.

法二 设 e_1, \dots, e_r 为 $\text{Span}(S)$ 的基, 将其扩充为 V 的基 $e_1, \dots, e_r, e_{r+1}, \dots, e_n$, 并设其对偶基为 e_1^*, \dots, e_n^* . 若 $f \in \text{Ann}(\text{Span}(S))$, 则 f 的前 r 个坐标分量均为 0 (否则 f 作用在该坐标对应的向量上非 0). 反过来, 对于任意 $f = \lambda_{r+1}e_{r+1}^* + \dots + \lambda_n e_n^*$, 有

$$f(\mu_1 e_1 + \dots + \mu_r e_r) = 0, \quad \forall \mu_1, \dots, \mu_r \in \mathbb{F}.$$

因此 $f \in \text{Ann}(\text{Span}(S))$. 再由第 (1) 问知 $\dim \text{Ann}(S) = \dim \text{Ann}(\text{Span}(S)) = n - r = \dim V - \text{rank}(S)$.

法三 由习题 8.7.3 证明中的引理 (2), 可设 $V = \text{Span}(S) \oplus T$. 构造映射

$$\rho: T^* \rightarrow \text{Ann}(\text{Span}(S)), \quad f \mapsto \bar{f},$$

其中 \bar{f} 是 f 在 V 上的零延拓, 即 $\bar{f}(x) = \begin{cases} f(x), & x \in T, \\ 0, & x \in \text{Span}(S). \end{cases}$ 容易验证 ρ 是 $T^* \rightarrow \text{Ann}(\text{Span}(S))$ 的同构映射. 因此

$$\dim \text{Ann}(S) = \dim \text{Ann}(\text{Span}(S)) = \dim(T^*) \xrightarrow{T \text{ 有限维}} \dim T = \dim V - \text{rank}(S).$$

(3) **等式一** 若 $f \in \text{Ann}(V_1 \cap V_2)$. 设 e_1, \dots, e_r 是 $V_1 \cap V_2$ 的基, 将其扩充为 $V_1 + V_2$ 的基 $e_1, \dots, e_r, e_{r+1}, \dots, e_n$, 并设其对偶基为 e_1^*, \dots, e_n^* . 同第 (2) 问知 $f \in \text{Span}(e_{r+1}^*, \dots, e_n^*) = \text{Span}(\text{Ann}(V_1) \cup \text{Ann}(V_2)) = \text{Ann}(V_1) + \text{Ann}(V_2)$. 因此 $\text{Ann}(V_1 \cap V_2) \subset \text{Ann}(V_1) + \text{Ann}(V_2)$.

若 $f \in \text{Ann}(V_1) + \text{Ann}(V_2)$, 则存在 $g \in \text{Ann}(V_1)$ 与 $h \in \text{Ann}(V_2)$ 使得 $f = g + h$, 从而对任意 $x \in V_1 \cap V_2$ 有 $f(x) = g(x) + h(x) = 0$. 故 $\text{Ann}(V_1) + \text{Ann}(V_2) \subset \text{Ann}(V_1 \cap V_2)$.

综上所述可得 $\text{Ann}(V_1 \cap V_2) = \text{Ann}(V_1) + \text{Ann}(V_2)$.

等式二 若 $f \in \text{Ann}(V_1) \cap \text{Ann}(V_2)$, 则对任意 $x = x_1 + x_2 \in V_1 + V_2$, 其中 $x \in V_1, x_2 \in V_2$, 有 $f(x) = f(x_1) + f(x_2) = 0$, 即 $f \in \text{Ann}(V_1 + V_2)$. 故 $\text{Ann}(V_1) \cap \text{Ann}(V_2) \subset \text{Ann}(V_1 + V_2)$.

若 $f \in \text{Ann}(V_1 + V_2)$. 由 $V_i \subset V_1 + V_2$ ($i = 1, 2$) 可知 $f \in \text{Ann}(V_i)$ ($i = 1, 2$), 从而 $f \in \text{Ann}(V_1) \cap \text{Ann}(V_2)$. 故 $\text{Ann}(V_1 + V_2) \subset \text{Ann}(V_1) \cap \text{Ann}(V_2)$.

综上所述可得 $\text{Ann}(V_1 + V_2) = \text{Ann}(V_1) \cap \text{Ann}(V_2)$.

(4) 由第 (3) 问, $V = V_1 \oplus V_2 \implies \text{Ann}(V_1) + \text{Ann}(V_2) = \text{Ann}(V_1 \cap V_2) = \text{Ann}(\{\mathbf{0}\}) = V^*$, 再结合 $\text{Ann}(V_1) \cap \text{Ann}(V_2) = \text{Ann}(V_1 + V_2) = \text{Ann}(V) = \{\mathbf{0}\}$ 得 $V^* = \text{Ann}(V_1) \oplus \text{Ann}(V_2)$. □

注 2.0.17 下面是关于习题 9.3.6 中“零化子”的几个实例:

(1) 设 $V = \mathbb{R}[x]$, $S = \{f \in V \mid f \text{ 是 } x^2 \text{ 的倍式}\} \subset V$, $\varphi \in V^*: p \mapsto p'(0)$. 则 $\varphi \in \text{Ann}(S)$.

(2) 设 e_1, e_2, e_3, e_4, e_5 是 \mathbb{R}^5 的标准基, $e_1^*, e_2^*, e_3^*, e_4^*, e_5^*$ 是其对偶基. 设

$$S = \text{Span}(e_1, e_2) = \{(x_1, x_2, 0, 0, 0) \in \mathbb{R}^5 \mid x_1, x_2 \in \mathbb{R}\},$$

则 $\text{Ann}(S) = \text{Span}(e_3^*, e_4^*, e_5^*)$.

习题 9.4.2 设 $V = \mathbb{F}^{3 \times 3}$, $\mathcal{A} \in \mathcal{L}(V) : X \mapsto X + X^\top$. 分别求 $\text{Ker } \mathcal{A}$ 和 $\text{Im } \mathcal{A}$ 的一个基.

解 ① 若 $\text{char } \mathbb{F} \neq 2$. 由 $\text{Ker } \mathcal{A} = \{X \mid X^\top = -X\}$ 及 $\text{Im } \mathcal{A} = \{X \mid X^\top = X\}$ 知 $\text{Ker } \mathcal{A}$ 的一个基为 $E_{21} - E_{12}, E_{31} - E_{13}, E_{32} - E_{23}$, $\text{Im } \mathcal{A}$ 的一个基为 $E_{11}, E_{21} + E_{12}, E_{22}, E_{31} + E_{13}, E_{32} + E_{23}, E_{33}$.

② 若 $\text{char } \mathbb{F} = 2$. 由 $\text{Ker } \mathcal{A} = \{X \mid X^\top = -X\}$ 知 $\text{Ker } \mathcal{A}$ 的一个基为 $E_{11}, E_{21} + E_{12}, E_{22}, E_{31} + E_{13}, E_{32} + E_{23}, E_{33}$. 由 $\text{Im } \mathcal{A} = \{X \mid X^\top = X \text{ 且 } X \text{ 的对角元均为 } 0\}$ 知 $\text{Im } \mathcal{A}$ 的一个基为 $E_{21} + E_{12}, E_{31} + E_{13}, E_{32} + E_{23}$. \square

习题 9.4.3 设 $\mathcal{A} \in \mathcal{L}(U, V)$, $W = \text{Ker } \mathcal{A}$, U_1, U_2 都是 U 的有限维子空间, $U_1 \subset U_2$. 证明:

$$\dim(U_2 \cap W) - \dim(U_1 \cap W) \leq \dim U_2 - \dim U_1,$$

并求等号成立的充分必要条件.

证明 设 \mathcal{A}_i 是 \mathcal{A} 在 (U_i, V) 上的限制映射 ($i = 1, 2$). 由同态基本定理,

$$\dim(U_i \cap W) = \dim(\text{Ker } \mathcal{A}_i) = \dim U_i - \dim(\text{Im } \mathcal{A}_i), \quad i = 1, 2.$$

由 $\text{Im } \mathcal{A}_1 \subset \text{Im } \mathcal{A}_2$ 得 $\dim(\text{Im } \mathcal{A}_1) \leq \dim(\text{Im } \mathcal{A}_2)$, 结合上面两式得

$$\dim U_1 - \dim(U_1 \cap W) \leq \dim U_2 - \dim(U_2 \cap W),$$

移项即得证. 等号成立当且仅当 $\dim(\text{Im } \mathcal{A}_1) = \dim(\text{Im } \mathcal{A}_2)$, 由 U_1, U_2 有限维又等价于 $\text{Im } \mathcal{A}_1 = \text{Im } \mathcal{A}_2$. \square

习题 9.4.4 (2)(3) 设 $\mathcal{A} \in \mathcal{L}(U, V)$, $\mathcal{B} \in \mathcal{L}(V, W)$. 证明:

(2) 若 $\text{Ker } \mathcal{A} = \text{Ker } \mathcal{B}\mathcal{A}$, 则存在 $\mathcal{C} \in \mathcal{L}(W, V)$, 使得 $\mathcal{A} = \mathcal{C}\mathcal{B}\mathcal{A}$.

(3) 若 $\text{Im } \mathcal{B}\mathcal{A} = \text{Im } \mathcal{B}$, 则存在 $\mathcal{C} \in \mathcal{L}(V, U)$, 使得 $\mathcal{B} = \mathcal{B}\mathcal{A}\mathcal{C}$.

证明 (2) 设 \mathcal{B}_1 是 \mathcal{B} 在 $(\text{Im } \mathcal{A}, W)$ 上的限制映射. 若 $\text{Ker } \mathcal{B}_1 \neq \{0\}$, 则存在 $x \in U$, 使得 $\mathcal{A}x \neq 0$ 且 $\mathcal{B}_1(\mathcal{A}x) = \mathcal{B}\mathcal{A}x = 0$, 由 $\text{Ker } \mathcal{A} = \text{Ker } \mathcal{B}\mathcal{A}$ 得 $\mathcal{A}x = 0$, 矛盾, 因此 $\text{Ker } \mathcal{B}_1 = \{0\}$, 从而 \mathcal{B}_1 是单射. 构造 $\mathcal{C} \in \mathcal{L}(W, V)$ 满足 $\mathcal{C}x = \begin{cases} \mathcal{B}_1^{-1}x, & x \in \text{Im } \mathcal{B}_1 \\ 0, & x \in S \end{cases}$, 其中 $S \oplus \text{Im } \mathcal{B}_1 = W$. 由直和容易验证 \mathcal{C} 的确为线性映射, 且 $\mathcal{A} = \mathcal{C}\mathcal{B}\mathcal{A}$.

(3) 设 $\{\beta_i\}_{i \in I}$ 是 V 的基, 因为 $\text{Im } \mathcal{B}\mathcal{A} = \text{Im } \mathcal{B}$, 所以存在 $\{\alpha_i\}_{i \in I}$ 使得 $\mathcal{A}\alpha_i = \beta_i$. 由定理 9.2 知存在 $\mathcal{C} \in \mathcal{L}(V, U)$ 满足 $\mathcal{C}\beta_i = \alpha_i$. 容易验证 $\mathcal{B} = \mathcal{B}\mathcal{A}\mathcal{C}$. \square

注 2.0.18 第 (2) 问中构造 \mathcal{C} 时用到了零化子与直和的关系: $(S \oplus T)^* = \text{Ann } S \oplus \text{Ann } T$.

习题 9.4.5 设 $\mathcal{A} \in \mathcal{L}(U, V)$ 且 $\mathcal{A} \neq \mathcal{O}$. 证明:

- (1) 存在 $\mathcal{B} \in \mathcal{L}(V, U)$, 满足 $\mathcal{A}\mathcal{B}\mathcal{A} = \mathcal{A}$ 且 $\mathcal{B}\mathcal{A}\mathcal{B} = \mathcal{B}$. \mathcal{B} 称为 \mathcal{A} 的一个广义逆映射.
- (2) \mathcal{B} 是唯一的 $\iff \mathcal{A}$ 是可逆映射.

证明 (1) 设 $U = \text{Ker } \mathcal{A} \oplus U_1$, $V = \text{Im } \mathcal{A} \oplus V_1$, \mathcal{A}_1 是 \mathcal{A} 在 $(U_1, \text{Im } \mathcal{A})$ 上的限制映射. 构造 $\mathcal{B} \in \mathcal{L}(V, U)$ 满足 $\mathcal{B}x = \begin{cases} \mathcal{A}_1^{-1}x, & x \in \text{Im } \mathcal{A} \\ \mathbf{0}, & x \in V_1 \end{cases}$. 可验证 \mathcal{B} 的确为线性映射 (参见注 2.0.18).

对任意 $x = \alpha_1 + \alpha_2 \in U$, 其中 $\alpha_1 \in U_1, \alpha_2 \in \text{Ker } \mathcal{A}$, 有 $\mathcal{A}\mathcal{B}\mathcal{A}x = \mathcal{A}\mathcal{B}\mathcal{A}\alpha_1 = \mathcal{A}\alpha_1 = \mathcal{A}x$, 即 $\mathcal{A}\mathcal{B}\mathcal{A} = \mathcal{A}$. 对任意 $y = \beta_1 + \beta_2 \in V$, 其中 $\beta_1 \in \text{Im } \mathcal{A}, \beta_2 \in V_1$, 有 $\mathcal{B}\mathcal{A}\mathcal{B}y = \mathcal{B}\beta_1 = \mathcal{B}y$, 即 $\mathcal{B}\mathcal{A}\mathcal{B} = \mathcal{B}$.

(2) \Leftarrow : 若 \mathcal{A} 可逆, 则 $\mathcal{B}\mathcal{A} = \mathcal{I}_U$ 且 $\mathcal{A}\mathcal{B} = \mathcal{I}_V$, \mathcal{B} 只能为 \mathcal{A}^{-1} .

\Rightarrow : 先证明线性空间的非平凡子空间的补空间一定不唯一.

引理 设 V 是线性空间, $\{\mathbf{0}\} \neq U \subsetneq V$, 则 U 的补空间不唯一.

【引理的证明: 设 $V = U \oplus W$, S 是 U 的基, T 是 W 的基, 则 S, T 均非空. 任取 $s \in S$ 与 $t \in T$, 则由 (外) 直和的定义可验证 $V = U \oplus \text{Span}((T \setminus \{t\}) \cup \{s + t\})$.**】**

用反证法. 假设 \mathcal{A} 不可逆. 由 $\mathcal{A} \neq \mathcal{O}$ 知 $\text{Ker } \mathcal{A} \neq U$ 且 $\text{Im } \mathcal{A} \neq \{\mathbf{0}\}$.

① 若 \mathcal{A} 不是单射, 则 $\text{Ker } \mathcal{A} \neq \{\mathbf{0}\}$. 由引理知 (1) 中 U_1 不唯一, 从而 \mathcal{B} 不唯一 (设 $x \notin U_1$ 是 $\text{Ker } \mathcal{A}$ 另一个补空间中的向量, $x = \gamma_1 + \gamma_2$, 其中 $\gamma_1 \in U_1, \mathbf{0} \neq \gamma_2 \in \text{Ker } \mathcal{A}$, 则 $\mathcal{B}\mathcal{A}x = \mathcal{B}\mathcal{A}\gamma_1 \neq x$, 而由新的补空间定义的映射将 x 映为 x), 矛盾.

② 若 \mathcal{A} 不是满射, 则 $\text{Im } \mathcal{A} \neq V$. 由引理知 (1) 中 V_1 不唯一, 从而 \mathcal{B} 不唯一 (设 $x \notin V_1$ 是 $\text{Im } \mathcal{A}$ 另一个补空间中的向量, $x = \gamma_1 + \gamma_2$, 其中 $\mathbf{0} \neq \gamma_1 \in \text{Im } \mathcal{A}, \gamma_2 \in V_1$, 则 $\mathcal{B}x = \mathcal{B}\gamma_1 \neq \mathbf{0}$, 而由新的补空间定义的映射将 x 映为 $\mathbf{0}$), 矛盾.

故假设不成立, \mathcal{A} 可逆. \square

习题 9.4.6 设 $\mathcal{A} \in \mathcal{L}(U, V), \mathcal{B} \in \mathcal{L}(V, U)$ 满足 $\mathcal{A}\mathcal{B}\mathcal{A} = \mathcal{A}$ 且 $\mathcal{B}\mathcal{A}\mathcal{B} = \mathcal{B}$. 证明:

(1) $U = \text{Ker } \mathcal{A} \oplus \text{Im } \mathcal{B}, V = \text{Im } \mathcal{A} \oplus \text{Ker } \mathcal{B}$.

(2) \mathcal{A} 在 $(\text{Im } \mathcal{B}, \text{Im } \mathcal{A})$ 上的限制映射与 \mathcal{B} 在 $(\text{Im } \mathcal{A}, \text{Im } \mathcal{B})$ 上的限制映射互为逆映射.

证明 (1) ① 设 $x \in \text{Ker } \mathcal{A} \cap \text{Im } \mathcal{B}$, 则由 $\mathcal{B}\mathcal{A}\mathcal{B} = \mathcal{B}$ 得 $x = \mathcal{B}\mathcal{A}x = \mathbf{0}$, 即 $\text{Ker } \mathcal{A} \cap \text{Im } \mathcal{B} = \{\mathbf{0}\}$. 对任意 $x \in U$, 由 $\mathcal{A}\mathcal{B}\mathcal{A} = \mathcal{A}$ 得 $x - \mathcal{B}\mathcal{A}x \in \text{Ker } \mathcal{A}$, 即 $x \in \text{Ker } \mathcal{A} + \text{Im } \mathcal{B}$. 综上, $U = \text{Ker } \mathcal{A} \oplus \text{Im } \mathcal{B}$.

② 设 $x \in \text{Im } \mathcal{A} \cap \text{Ker } \mathcal{B}$, 则由 $\mathcal{A}\mathcal{B}\mathcal{A} = \mathcal{A}$ 得 $x = \mathcal{A}\mathcal{B}x = \mathbf{0}$, 即 $\text{Im } \mathcal{A} \cap \text{Ker } \mathcal{B} = \{\mathbf{0}\}$. 对任意 $x \in V$, 由 $\mathcal{B}\mathcal{A}\mathcal{B} = \mathcal{B}$ 得 $x - \mathcal{A}\mathcal{B}x \in \text{Ker } \mathcal{B}$, 即 $x \in \text{Im } \mathcal{A} + \text{Ker } \mathcal{B}$. 综上, $V = \text{Im } \mathcal{A} \oplus \text{Ker } \mathcal{B}$.

(2) 设 \mathcal{A}_1 为 \mathcal{A} 在 $(\text{Im } \mathcal{B}, \text{Im } \mathcal{A})$ 上的限制映射, \mathcal{B}_1 为 \mathcal{B} 在 $(\text{Im } \mathcal{A}, \text{Im } \mathcal{B})$ 上的限制映射. 对任意 $x = \mathcal{B}_1 y \in \text{Im } \mathcal{B}$, 由 $\mathcal{B}\mathcal{A}\mathcal{B} = \mathcal{B}$ 得 $\mathcal{B}_1 \mathcal{A}_1 x = x$; 对任意 $y = \mathcal{A}_1 x \in \text{Im } \mathcal{A}$, 由 $\mathcal{A}\mathcal{B}\mathcal{A} = \mathcal{A}$ 得 $\mathcal{A}_1 \mathcal{B}_1 y = y$. 故 \mathcal{A}_1 与 \mathcal{B}_1 互为逆映射. \square

习题 9.4.7 设 $\mathcal{A}, \mathcal{B} \in \mathcal{L}(U, V)$ 满足 $\text{Im}(\mathcal{A} + \mathcal{B}) = \text{Im } \mathcal{A} \oplus \text{Im } \mathcal{B}$. 证明: 存在 U 的子空间 U_1, U_2, U_3 , 使得 \mathcal{A} 在 $(U_1, \text{Im } \mathcal{A})$ 上的限制映射和 \mathcal{B} 在 $(U_2, \text{Im } \mathcal{B})$ 上的限制映射都是可逆映射, 并且

$$U = U_1 \oplus U_2 \oplus U_3, \quad \text{Ker } \mathcal{A} = U_2 \oplus U_3, \quad \text{Ker } \mathcal{B} = U_1 \oplus U_3.$$

证明 设 $U_3 = \text{Ker } \mathcal{A} \cap \text{Ker } \mathcal{B}$, $\text{Ker } \mathcal{A} = U_2 \oplus U_3$, $\text{Ker } \mathcal{B} = U_1 \oplus U_3$. 由 $U_2 \cap U_3 = \{\mathbf{0}\}$ 及 $U_1 \cap (U_2 + U_3) = \{\mathbf{0}\}$ 知 $U_1 + U_2 + U_3$ 是直和. 只需再证 $U = \text{Ker } \mathcal{A} + \text{Ker } \mathcal{B}$. 由 $\text{Im}(\mathcal{A} + \mathcal{B}) = \text{Im } \mathcal{A} \oplus \text{Im } \mathcal{B}$, 对任意 $\alpha \in U$, 存在 $\beta \in U$, 使得 $\mathcal{A}\alpha = (\mathcal{A} + \mathcal{B})\beta$, 即 $\mathcal{A}(\alpha - \beta) = \mathcal{B}\beta = \mathbf{0}$ (最后的等号由 $\text{Im } \mathcal{A} + \text{Im } \mathcal{B}$ 是直和可得). 故 $\alpha - \beta \in \text{Ker } \mathcal{A}$ 且 $\beta \in \text{Ker } \mathcal{B}$, 即 $\alpha \in \text{Ker } \mathcal{A} + \text{Ker } \mathcal{B}$, $U = \text{Ker } \mathcal{A} + \text{Ker } \mathcal{B}$.

由 $U_1 \cap \text{Ker } \mathcal{A} = \{\mathbf{0}\}$ 得 \mathcal{A} 在 $(U_1, \text{Im } \mathcal{A})$ 上是可逆的; 由 $U_2 \cap \text{Ker } \mathcal{B} = \{\mathbf{0}\}$ 得 \mathcal{B} 在 $(U_2, \text{Im } \mathcal{B})$ 上是可逆的. \square

习题 9.4.8 设 $\mathcal{A}, \mathcal{B} \in \mathcal{L}(V)$ 满足 $\mathcal{A}^2 = \mathcal{A}$, $\mathcal{B}^2 = \mathcal{B}$. 证明:

(1) $\text{Im } \mathcal{A} = \text{Im } \mathcal{B} \iff \mathcal{A}\mathcal{B} = \mathcal{B}$ 且 $\mathcal{B}\mathcal{A} = \mathcal{A}$.

(2) $\text{Ker } \mathcal{A} = \text{Ker } \mathcal{B} \iff \mathcal{A}\mathcal{B} = \mathcal{A}$ 且 $\mathcal{B}\mathcal{A} = \mathcal{B}$.

(3) 设 V 是有限维的. $\text{rank}(\mathcal{A}) = \text{rank}(\mathcal{B}) \iff$ 存在可逆映射 $\mathcal{C} \in \mathcal{L}(V)$, 使得 $\mathcal{A}\mathcal{C} = \mathcal{C}\mathcal{B}$.

证明 (1) \Leftarrow : 对任意 $y = \mathcal{A}x \in \text{Im } \mathcal{A}$, 由 $\mathcal{A}x = \mathcal{B}\mathcal{A}x$ 得 $\mathcal{B}(\mathcal{A}x) = y$, 即 $\text{Im } \mathcal{A} \subset \text{Im } \mathcal{B}$. 同理, 利用 $\mathcal{A}\mathcal{B} = \mathcal{B}$ 可得 $\text{Im } \mathcal{B} \subset \text{Im } \mathcal{A}$. 故 $\text{Im } \mathcal{A} = \text{Im } \mathcal{B}$.

\Rightarrow : 由例 (9.15) 可知 \mathcal{A} 在 $\text{Im } \mathcal{A} = \text{Im } \mathcal{B}$ 上的限制映射是恒等变换, 故 $\mathcal{A}\mathcal{B} = \mathcal{B}$. 同理可证 $\mathcal{B}\mathcal{A} = \mathcal{A}$.

(2) \Leftarrow : 若 $x \in \text{Ker } \mathcal{B}$, 则 $\mathcal{A}x = \mathcal{A}\mathcal{B}x = \mathbf{0}$, 即 $\text{Ker } \mathcal{B} \subset \text{Ker } \mathcal{A}$. 同理可证 $\text{Ker } \mathcal{A} \subset \text{Ker } \mathcal{B}$. 故 $\text{Ker } \mathcal{A} = \text{Ker } \mathcal{B}$.

\Rightarrow : 由 $\mathcal{A}^2 = \mathcal{A}$ 可知, 对任意 $\alpha \in V$, $\alpha - \mathcal{A}\alpha \in \text{Ker } \mathcal{A} = \text{Ker } \mathcal{B}$, 即 $\mathcal{B}\alpha = \mathcal{B}\mathcal{A}\alpha$, 故 $\mathcal{B}\mathcal{A} = \mathcal{B}$. 同理可证 $\mathcal{A}\mathcal{B} = \mathcal{A}$.

(3) 设 $\dim V = n$. 由例 (9.15) 知 $V = \text{Im } \mathcal{A} \oplus \text{Ker } \mathcal{A} = \text{Im } \mathcal{B} \oplus \text{Ker } \mathcal{B}$, 故可取 $\text{Im } \mathcal{A}$ 的基 $\alpha_1, \dots, \alpha_r$, $\text{Ker } \mathcal{A}$ 的基 $\alpha_{r+1}, \dots, \alpha_n$, $\text{Im } \mathcal{B}$ 的基 β_1, \dots, β_r , $\text{Ker } \mathcal{B}$ 的基 $\beta_{r+1}, \dots, \beta_n$. 由定理 9.2, 存在 $\mathcal{C} \in \mathcal{L}(V) : \beta_i \mapsto \alpha_i, \forall i$.

$$\begin{aligned} \mathcal{A}\mathcal{C}\beta_i &= \mathcal{A}\alpha_i, & \mathcal{C}\mathcal{B}\beta_i &\stackrel{\mathcal{B}^2=\mathcal{B}}{\underset{\beta_i \in \text{Im } \mathcal{B}}{=}} \mathcal{C}\beta_i = \mathcal{A}\alpha_i, & i = 1, \dots, r. \\ \mathcal{A}\mathcal{C}\beta_i &= \mathcal{A}\alpha_i \stackrel{\alpha_i \in \text{Ker } \mathcal{A}}{=} \mathbf{0}, & \mathcal{C}\mathcal{B}\beta_i &\stackrel{\beta_i \in \text{Ker } \mathcal{B}}{=} \mathcal{C}(\mathbf{0}) = \mathbf{0}, & i = r+1, \dots, n. \end{aligned}$$

故 $\mathcal{A}\mathcal{C} = \mathcal{C}\mathcal{B}$. □

习题 9.4.9 设 $\mathcal{A} \in \mathcal{L}(V)$, \mathcal{B} 是 \mathcal{A} 在 $\text{Im } \mathcal{A}$ 上的限制映射. 证明:

- (1) $V = \text{Im } \mathcal{A} + \text{Ker } \mathcal{A} \iff \mathcal{B}$ 是满射.
- (2) $V = \text{Im } \mathcal{A} \oplus \text{Ker } \mathcal{A} \iff \mathcal{B}$ 是可逆映射.

证明 (1) \Leftarrow : \mathcal{B} 是满射即 $\text{Im } \mathcal{A}^2 = \text{Im } \mathcal{A}$, 因此对任意 $\alpha \in V$, 存在 $\beta \in \text{Im } \mathcal{A}$, 使得 $\mathcal{A}\alpha = \mathcal{B}\beta$, 即 $\alpha - \beta \in \text{Ker } \mathcal{A}$. 故 $V = \text{Im } \mathcal{A} + \text{Ker } \mathcal{A}$.

\Rightarrow : 对任意 $\alpha = \beta + \gamma \in V$, 其中 $\beta \in \text{Im } \mathcal{A}, \gamma \in \text{Ker } \mathcal{A}$, 有 $\mathcal{A}\alpha = \mathcal{B}\beta$. 故 \mathcal{B} 是满射.

- (2) 由 (1) 及 $\text{Im } \mathcal{A} \cap \text{Ker } \mathcal{A} = \{\mathbf{0}\} \iff \mathcal{B}$ 是单射即得. □

习题 9.4.10 设 $\mathcal{A} \in \mathcal{L}(V)$ 满足 $\mathcal{A}^m = \mathcal{O}$, 其中 m 是给定的正整数. \mathcal{A} 称为幂零变换. 证明: 存在 V 的子空间 U , 使得 $V = \bigoplus_{i=1}^m \mathcal{A}^{i-1}(U)$.

证明 对 m 归纳. 当 $m = 1$ 时, 结论显然成立. 假设结论对 $m-1$ 成立. 设 \mathcal{B} 是 \mathcal{A} 在 $\text{Im } \mathcal{A}$ 上的限制映射, 则 $\mathcal{B}^{m-1} = \mathcal{O}$. 因此, 存在 $U_1 \leq \text{Im } \mathcal{A}$ 使得 $\text{Im } \mathcal{A} = \bigoplus_{i=1}^{m-1} \mathcal{A}^{i-1}(U_1)$. 设 $\{\mathcal{A}\alpha_i \mid i \in I\}$ 是 U_1 的基, $U_0 = \text{Span}(\{\alpha_i \mid i \in I\})$, 则 $\{\alpha_i \mid i \in I\}$ 线性无关, $\mathcal{A}(U_0) = U_1$.

• $U_0 \cap \text{Im } \mathcal{A} = \{\mathbf{0}\}$. 因为 $\mathcal{A}^{m-1}(U_1) = \mathcal{B}^{m-1}(U_1) = \{\mathbf{0}\}$, 所以 $\mathcal{A}(\text{Im } \mathcal{A}) = \bigoplus_{i=2}^{m-1} \mathcal{A}^{i-1}(U_1)$. 由两个直和可知 $\mathcal{A}(U_0) \cap \mathcal{A}(\text{Im } \mathcal{A}) = \{\mathbf{0}\}$, 从而 $(U_0 \cap \text{Im } \mathcal{A}) \subset \text{Ker } \mathcal{A}$. 因为 $U_0 \cap \text{Ker } \mathcal{A} = \{\mathbf{0}\}$, 所以 $U_0 \cap \text{Im } \mathcal{A} = \{\mathbf{0}\}$.

• $V = \left(\bigoplus_{i=1}^{m-1} \mathcal{A}^{i-1}(U_0) \right) + \text{Ker } \mathcal{A}$. 因为 $\text{Im } \mathcal{A} = \bigoplus_{i=1}^{m-1} \mathcal{A}^{i-1}(U_1)$, 所以对任意 $\alpha \in V$, 存在 $\beta_i \in \mathcal{A}^{i-1}(U_1)$ ($i = 1, \dots, m-1$), 使得 $\mathcal{A}\alpha = \sum_{i=1}^{m-1} \beta_i$. 因为 $\beta_1 \in U_1 = \mathcal{A}(U_0)$, 所以存在 $\gamma_1 \in U_0$ 使得 $\mathcal{A}\gamma_1 = \beta_1$; 当 $i \geq 2$ 时, $\beta_i \in \mathcal{A}^{i-1}(U_1) \subset \mathcal{A}^{i-1}(\text{Im } \mathcal{A}) = \text{Im } \mathcal{A}^i \subset \text{Im } \mathcal{A}^2$,

因此存在 $\gamma_i \in \text{Im } \mathcal{A}$ ($i = 2, \dots, m-1$), 使得 $\mathcal{A}\gamma_i = \beta_i$. 故 $\mathcal{A}\left(\alpha - \gamma_1 - \sum_{i=2}^{m-1} \gamma_i\right) = \mathbf{0}$, 即 $\alpha - \gamma_1 - \sum_{i=2}^{m-1} \gamma_i \in \text{Ker } \mathcal{A}$. 于是 $V = U_0 + \text{Im } \mathcal{A} + \text{Ker } \mathcal{A} = U_0 + \bigoplus_{i=1}^{m-1} \mathcal{A}^i(U_0) + \text{Ker } \mathcal{A}$. 注意到 $\mathcal{A}^{m-1}(U_0) = \mathcal{A}^{m-2}(U_1) = \mathcal{B}^{m-2}(U_1) \subset \text{Ker } \mathcal{A}$, 以及上一点证明的 $U_0 \cap \text{Im } \mathcal{A} = \{\mathbf{0}\}$, 就有 $V = \left(\bigoplus_{i=1}^{m-1} \mathcal{A}^{i-1}(U_0)\right) + \text{Ker } \mathcal{A}$.

设 $W_0 \leq \text{Ker } \mathcal{A}$ 使得 $V = \left(\bigoplus_{i=1}^m \mathcal{A}^{i-1}(U_0)\right) \oplus W_0$, 取 $U = U_0 \oplus W_0$ (由 U_1 与 U_0 关系可知 $U_0 \cap \text{Ker } \mathcal{A} = \{\mathbf{0}\}$, 从而验证直和), 则由 $\mathcal{A}(U_0 \oplus W_0) = \mathcal{A}(U_0)$ 可得

$$\begin{aligned} V &= W_0 \oplus U_0 \oplus \mathcal{A}(U_0) \oplus \dots \oplus \mathcal{A}^{m-1}(U_0) \\ &= (U_0 \oplus W_0) \oplus \mathcal{A}(U_0 \oplus W_0) \oplus \dots \oplus \mathcal{A}^{m-1}(U_0 \oplus W_0) \\ &= \bigoplus_{i=1}^m \mathcal{A}^{i-1}(U). \end{aligned}$$

□

注 2.0.19 关于幂零变换另见习题 9.7.5 与注 2.0.21.

习题 9.5.2 设 V 是 \mathbb{F} 上的线性空间, $\mathcal{A} \in \mathcal{L}(V)$, V 的子空间都是 \mathcal{A} -不变的. 证明: $\mathcal{A} = a\mathcal{I}$, $a \in \mathbb{F}$.

证明 考虑 V 的 \mathcal{A} -不变子空间 $\text{Span}(\alpha_i)$, 其中 $\alpha_i \in V$, 则存在 $\lambda_i \in \mathbb{F}$, 使得 $\mathcal{A}\alpha_i = \lambda_i\alpha_i$. 于是对任意 $\alpha_1, \alpha_2 \in V$, 存在 $\mu \in \mathbb{F}$, 使得 $\mathcal{A}(\alpha_1 + \alpha_2) = \lambda_1\alpha_1 + \lambda_2\alpha_2 = \mu(\alpha_1 + \alpha_2)$, 由此可知 $\lambda_1 = \lambda_2 =: a$. 故 $\mathcal{A} = a\mathcal{I}$. □

习题 9.5.6 设 V 是 \mathbb{F} 上的线性空间, $\mathcal{A}, \mathcal{B} \in \mathcal{L}(V)$ 满足 $\mathcal{A}\mathcal{B} - \mathcal{B}\mathcal{A} = \mathcal{A}$, n 是正整数, $p \in \mathbb{F}[x]$.

- (1) 证明: $\text{Ker } \mathcal{A}^n$ 和 $\text{Im } \mathcal{A}^n$ 都是 \mathcal{B} -不变的.
- (2) 设 $\text{char } \mathbb{F} = 0$, $\dim V = n$. 证明: $\mathcal{A}^n = \mathcal{O}$, $\text{Ker } p(\mathcal{A})$ 和 $\text{Im } p(\mathcal{A})$ 都是 \mathcal{B} -不变的.
- (3) 当 V 是无限维时, $\text{Ker } p(\mathcal{A})$ 和 $\text{Im } p(\mathcal{A})$ 是否一定是 \mathcal{B} -不变的?
- (4) $\text{Ker } \mathcal{B}^n$ 和 $\text{Im } \mathcal{B}^n$ 是否一定是 \mathcal{A} -不变的?

解 (1) 先证明 $\mathcal{A}^k\mathcal{B} - \mathcal{B}\mathcal{A}^k = k\mathcal{A}^k$. 当 $k=1$ 时即已知. 设结论对 $k-1$ 成立, 则

$$\begin{aligned} \mathcal{A}^k\mathcal{B} &= \mathcal{A}(\mathcal{A}^{k-1}\mathcal{B}) = \mathcal{A}(\mathcal{B}\mathcal{A}^{k-1} + (k-1)\mathcal{A}^{k-1}) = (\mathcal{B}\mathcal{A} + \mathcal{A})\mathcal{A}^{k-1} + (k-1)\mathcal{A}^k \\ &= \mathcal{B}\mathcal{A}^k + k\mathcal{A}^k. \end{aligned}$$

归纳即得 $\mathcal{A}^k \mathcal{B} - \mathcal{B} \mathcal{A}^k = k \mathcal{A}^k$. 于是

$$\mathcal{A}^n \alpha = \mathbf{0} \implies \mathcal{A}^n (\mathcal{B} \alpha) = (\mathcal{B} + n\mathcal{I})(\mathcal{A}^n \alpha) = \mathbf{0},$$

$$\alpha = \mathcal{A}^n \beta \implies \mathcal{B} \alpha = \mathcal{B} \mathcal{A}^n \beta = \mathcal{A}^n (\mathcal{B} - n\mathcal{I}) \beta,$$

即 $\text{Ker } \mathcal{A}^n$ 和 $\text{Im } \mathcal{A}^n$ 都是 \mathcal{B} -不变的.

(2) 由 (1) 中的 $\mathcal{A}^k \mathcal{B} - \mathcal{B} \mathcal{A}^k = k \mathcal{A}^k$ 可得对任意 $f \in \mathbb{F}[x]$, $f(\mathcal{A})\mathcal{B} - \mathcal{B}f(\mathcal{A}) = \mathcal{A}f'(\mathcal{A})$. 特别地, 考虑 \mathcal{A} 的最小多项式 $d_{\mathcal{A}}(x)$, 可知 $xd'_{\mathcal{A}}(x)$ 也是 \mathcal{A} 的化零多项式, 从而 $d_{\mathcal{A}}(x) \mid xd'_{\mathcal{A}}(x)$. 由两边次数相等知存在整数 $m \leq n$, 使得 $d_{\mathcal{A}}(x) = mxd'_{\mathcal{A}}(x)$, 即 $d_{\mathcal{A}}(x) = x^m$. 由此可知 $\mathcal{A}^n = \mathcal{O}$.

对任意 $p \in \mathbb{F}[x]$, 由 $d_{\mathcal{A}}(x)$ 的根均为 0 可设 $p(x) = x^k q(x)$, 其中 $q(x)$ 满足 $q(\mathcal{A})$ 是可逆变换, 则 $\text{Ker } p(\mathcal{A}) = \text{Ker } \mathcal{A}^k$, $\text{Im } p(\mathcal{A}) = \text{Im } \mathcal{A}^k$. 再由 (1) 即得证.

(3) 不一定, 反例如下. 设 $V = \mathcal{C}^\infty(\mathbb{R})$, $\mathcal{A} : f(x) \mapsto f'(x)$, $\mathcal{B} : f(x) \mapsto xf'(x)$. 则由 $(xf'(x))' - x(f'(x))' = f'(x)$ 可知 $\mathcal{A}\mathcal{B} - \mathcal{B}\mathcal{A} = \mathcal{A}$.

• $\text{Ker}(\mathcal{A} + \mathcal{I}) = \{f \in \mathcal{C}^\infty(\mathbb{R}) \mid f'(x) + f(x) = 0\} = \{ce^{-x} \mid c \in \mathbb{R}\}$. 但当 $c \neq 0$ 时, $\mathcal{B}(ce^{-x}) = -cxe^{-x} \notin \text{Ker}(\mathcal{A} + \mathcal{I})$.

•

(4) 不一定, 反例如下. 取 $V = \mathbb{F}^3$, $A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ 分别为 \mathcal{A}, \mathcal{B} 的

矩阵表示, 这满足 $\mathcal{A}\mathcal{B} - \mathcal{B}\mathcal{A} = \mathcal{A}$. 计算可知:

• $\text{Ker } \mathcal{B} = \text{Span}(\mathbf{e}_3)$, 但 $\mathcal{A}(\text{Ker } \mathcal{B}) = \text{Span}(\mathbf{e}_2)$, 即 $\text{Ker } \mathcal{B}$ 不是 \mathcal{A} -不变的.

• $\text{Im } \mathcal{B} = \text{Span}(\mathbf{e}_1, \mathbf{e}_2)$, 但 $\mathcal{A}(\text{Im } \mathcal{B}) = \text{Span}(\mathbf{e}_3)$, 即 $\text{Im } \mathcal{B}$ 不是 \mathcal{A} -不变的. \square

习题 9.5.7 (2) 设 V 是 \mathbb{F} 上的线性空间, $\mathcal{A} \in \mathcal{L}(V)$, W 是 V 的 \mathcal{A} -不变子空间, $\mathcal{B} \in \mathcal{L}(V/W) : [\alpha] \rightarrow [\mathcal{A}\alpha]$. 证明: 对于任意 \mathcal{B} -不变子空间 \tilde{U} , 存在 \mathcal{A} -不变子空间 $U \supset W$, 使得 $U/W = \tilde{U}$.

证明 取 $U = \{\alpha \in V \mid [\alpha] \in \tilde{U}\}$, 容易验证 $U \leq V$ 且 $W \subset U$. 对任意 $\alpha \in U$, $[\alpha] \in \tilde{U}$, 从而 $\mathcal{B}([\alpha]) = [\mathcal{A}\alpha] \in \tilde{U}$, $\mathcal{A}\alpha \in U$. 故 U 是 \mathcal{A} -不变的. \square

习题 9.5.8 设 V 是 \mathbb{F} 上的线性空间, $\mathcal{A}, \mathcal{B} \in \mathcal{L}(V)$, U 是 V 的 \mathcal{A} -不变子空间.

(1) 是否一定存在 $p \in \mathbb{F}[x]$, 使得 $U = \text{Ker } p(\mathcal{A})$?

(2) 是否一定存在 $p \in \mathbb{F}[x]$, 使得 $U = \text{Im } p(\mathcal{A})$?

(3) 是否一定存在 \mathcal{A} -不变子空间 W , 使得 $V = U \oplus W$?

(4) 若 \mathcal{A} 可逆, $\mathcal{B} = \mathcal{A}^{-1}$, 则 U 是否一定是 \mathcal{B} -不变的?

解 (1) 不一定. 取 $V = \mathbb{F}^2$, $\mathcal{A} = \mathcal{O}$, $U = \text{Span}(\mathbf{e}_1)$, 则 $p(\mathcal{A}) = \mathcal{O}$ 或 \mathcal{I} , $\text{Ker } p(\mathcal{A}) = \mathbb{F}^2$ 或 $\{\mathbf{0}\}$.

(2) 不一定. 取 $V = \mathbb{F}^2$, $\mathcal{A} = \mathcal{O}$, $U = \text{Span}(\mathbf{e}_1)$, 则 $p(\mathcal{A}) = \mathcal{O}$ 或 \mathcal{I} , $\text{Im } p(\mathcal{A}) = \{\mathbf{0}\}$ 或 \mathbb{F}^2 .

(3) 不一定. 取 $V = \mathbb{F}[x]$, \mathcal{A} 是微分变换, $U = \mathbb{F}_n[x]$, 则 U 的任一补空间 W 中一定含有次数为 n 的多项式, 从而 W 不是 \mathcal{A} -不变的.

(4) 不一定. 取 $V = \{x^{-n}f(x) \mid n \in \mathbb{N}, f \in \mathbb{F}[x]\}$, $\mathcal{A}: f(x) \mapsto xf(x)$, $U = \mathbb{F}[x]$, 则有 $1 \in U$ 但 $\mathcal{B}(1) = \frac{1}{x} \notin U$. \square

注 2.0.20 第 (3) 问说明在一般情形下, \mathcal{A} -不变子空间的补空间不一定是 \mathcal{A} -不变子空间, 但当 \mathcal{A} 是实内积空间 V 上的正交变换, 且 U 是有限维 \mathcal{A} -不变子空间时, 就有 $V = U \oplus U^\perp$ 且 U^\perp 是 \mathcal{A} -不变的 (定理 10.12); 第 (4) 问的反例只能从无限维线性空间中找, 因为有限维时 \mathcal{A}^{-1} 可表成 \mathcal{A} 的多项式 (参见习题 5.4.10 证明中的 ①), 从而 U 一定是 \mathcal{B} -不变的.

习题 9.6.4 设 V 是 \mathbb{F} 上的线性空间, $\mathcal{A} \in \mathcal{L}(V)$ 有最小多项式 $d_{\mathcal{A}} = \prod_{i=1}^k (x - \lambda_i)$, 其中

$\lambda_1, \dots, \lambda_k \in \mathbb{F}$ 两两不同. 证明: 任意 $\alpha \in V$ 可以表示为 $\alpha = \sum_{i=1}^k \alpha_i$ 的形式, 其中 $\alpha_i =$

$$f_i(\mathcal{A})\alpha \in \text{Ker}(\lambda_i \mathcal{I} - \mathcal{A}), f_i(x) = \prod_{j \neq i} \frac{x - \lambda_j}{\lambda_i - \lambda_j}.$$

证明 因为 $(\mathcal{A} - \lambda \mathcal{I})f_i(\mathcal{A})\alpha = d_{\mathcal{A}}(\mathcal{A}) \prod_{j \neq i} (\lambda_i - \lambda_j)^{-1} = \mathcal{O}$, 所以 $f_i(\mathcal{A})\alpha \in \text{Ker}(\lambda \mathcal{I} - \mathcal{A})$. 利用

Lagrange 插值多项式 可知, $\sum_{i=1}^k \prod_{j \neq i} \frac{x - \lambda_j}{\lambda_i - \lambda_j} = 1$, 从而 $\sum_{i=1}^k \alpha_i = \alpha$. \square

习题 9.6.5 设 V 是 \mathbb{C} 上的线性空间, $\mathcal{A} \in \mathcal{L}(V)$ 满足 $\mathcal{A}^n = \mathcal{I}$. 证明: 任意 $\alpha \in V$ 可以表示为 $\alpha = \sum_{k=0}^{n-1} \alpha_k$ 的形式, 其中 $\alpha_k = \frac{1}{n} \sum_{j=0}^{n-1} \omega^{-kj} \mathcal{A}^j \alpha \in \text{Ker}(\omega^k \mathcal{I} - \mathcal{A})$, $\omega = \cos \frac{2\pi}{n} + i \sin \frac{2\pi}{n}$.

证明 由习题 9.6.4 证明的前半部分可知将 $d_{\mathcal{A}}$ 改为 \mathcal{A} 的任一化零多项式结论仍成立. 注意到 \mathcal{A} 的化零多项式 $x^n - 1 = \prod_{i=0}^{n-1} (x - \omega^i)$, 因此只需证 $\prod_{j \neq k} \frac{x - \omega^j}{\omega^k - \omega^j} = \frac{1}{n} \sum_{j=0}^{n-1} \omega^{-kj} x^j$. 由

$$\prod_{j \neq k} (x - \omega^j) = \frac{x^n - 1}{x - \omega^k} = \frac{x^n - (\omega^k)^n}{x - \omega^k} = \sum_{j=0}^{n-1} (\omega^k)^{n-1-j} x^j = \sum_{j=0}^{n-1} \omega^{-k-kj} x^j \text{ 可得}$$

$$\prod_{j \neq k} \frac{x - \omega^j}{\omega^k - \omega^j} = \frac{\sum_{j=0}^{n-1} \omega^{-k-kj} x^j}{\sum_{j=0}^{n-1} \omega^{-k}} = \frac{1}{n} \sum_{j=0}^{n-1} \omega^{-kj} x^j.$$

\square

习题 9.6.6 设 V 是 \mathbb{F} 上的线性空间, $\mathcal{A} \in \mathcal{L}(V)$ 有最小多项式 $d_{\mathcal{A}} = \prod_{i=1}^k (x - \lambda_i)^{m_i}$, 其中 $\lambda_1, \dots, \lambda_k \in \mathbb{F}$ 两两不同. 证明:

- (1) $W_i = \text{Ker}(\lambda_i \mathcal{I} - \mathcal{A})^{m_i}$ 是 λ_i 对应的根子空间.
- (2) $(x - \lambda_i)^{m_i}$ 是 \mathcal{A} 在 W_i 上的限制映射的最小多项式.
- (3) 当 W_i 是有限维时, $m_i \leq \dim W_i \leq m_i \dim V_i$, 其中 $V_i = \text{Ker}(\lambda_i \mathcal{I} - \mathcal{A})$.

证明 (1) 由 9.16, $V = \bigoplus_{j=1}^k W_j$. 由定理 9.13, 对任意 $m \geq m_i$, $\text{Ker}(\lambda_i \mathcal{I} - \mathcal{A})^m$ 也是 $\bigoplus_{j \neq i} W_j$ 的补空间. 由例 (8.22), 由 $W_i \subset \text{Ker}(\lambda_i \mathcal{I} - \mathcal{A})^m$ 得 $W_i = \text{Ker}(\lambda_i \mathcal{I} - \mathcal{A})^m$. 故 W_i 是 λ_i 对应的根子空间.

(2) 由 W_i 定义知 $(x - \lambda_i)^{m_i}$ 是 \mathcal{A} 在 W_i 上的限制映射的化零多项式, 从而最小多项式 $d_i(x) = (x - \lambda_i)^m$, 其中 $m \leq m_i$. 若 $m < m_i$, 由定理 9.16 中的直和分解可知 $\frac{d_{\mathcal{A}}(x) \cdot d_i(x)}{(x - \lambda_i)^{m_i}}$ 也是 \mathcal{A} 的化零多项式, 与 $d_{\mathcal{A}}$ 的最小性矛盾. 故 $m = m_i$, 即 $(x - \lambda_i)^{m_i}$ 是 \mathcal{A} 在 W_i 上的限制映射的最小多项式.

(3) ① 由 W_i 是有限维的及 (2) 中 \mathcal{A} 在 W_i 上的限制映射的最小多项式次数为 m_i 知 \mathcal{A} 在 W_i 上的限制映射在某个基下的矩阵表示 A 阶数至少为 m_i . 由定理 9.4 即得 $\dim W_i \geq m_i$.

② 先证明一个不等式.

引理 设 $\dim V = n < \infty$, $\mathcal{A}, \mathcal{B} \in \mathcal{L}(V)$, 则 $\text{null}(\mathcal{A}\mathcal{B}) \leq \text{null}(\mathcal{A}) + \text{null}(\mathcal{B})$.

【引理的证明: 由 Sylvester 秩不等式 (例 (4.9), 或在例 (9.15) 中取 $\mathcal{B} = \mathcal{I}$), $\text{rank}(\mathcal{A}) + \text{rank}(\mathcal{B}) \leq \text{rank}(\mathcal{A}\mathcal{B}) + n$. 由矩阵表示可知 $\text{null}(\mathcal{A}) = n - \text{rank}(\mathcal{A})$, $\text{null}(\mathcal{B}) = n - \text{rank}(\mathcal{B})$. 联立即得证.】

由引理中的不等式, $\dim W_i = \text{null}(\lambda_i \mathcal{I} - \mathcal{A})^{m_i} \leq \text{null}(\lambda_i \mathcal{I} - \mathcal{A}) + \text{null}(\lambda_i \mathcal{I} - \mathcal{A})^{m_i-1} \leq \dots \leq m_i \text{null}(\lambda_i \mathcal{I} - \mathcal{A}) = m_i \dim V_i$. □

习题 9.6.7 设 $V = \mathbb{F}[x]$, $\mathcal{A} \in \mathcal{L}(V) : f(x) \mapsto xf'(x)$.

- (1) 求 \mathcal{A} 的所有特征值以及相应的特征子空间和根子空间.
- (2) 求 \mathcal{A} 关于任意非零向量 $\alpha \in V$ 的最小多项式 $d_{\mathcal{A}, \alpha}$.

证明 (1) 解 $xf'(x) = \lambda f(x)$ 得 $f(x) = cx^\lambda$, 其中 $c \in \mathbb{F}$. 而 $f(x) \in \mathbb{F}[x]$, 故任意 $\lambda \in \mathbb{N}$ 都是 \mathcal{A} 的特征值, λ 对应的特征子空间是 $\{cx^\lambda \mid c \in \mathbb{F}\}$. 任取 f 不是 λ 对应的特征向量, 则 f 的完全展开式中含有形如 bx^μ ($b \neq 0, \mu \neq \lambda$) 项. 因为 $\mathcal{A}(bx^\mu) - \lambda(bx^\mu) = b(\mu - \lambda)x^\mu \neq 0$, 所以归纳

可知对任意 $n \in \mathbb{N}$, $(\mathcal{A} - \lambda \mathcal{I})^n(bx^\mu) \neq 0$, 从而 $f \notin \bigcup_{k=1}^n \text{Ker}(\lambda \mathcal{I} - \mathcal{A})^k$, 故 $f \notin \bigcup_{k=1}^{\infty} \text{Ker}(\lambda \mathcal{I} - \mathcal{A})^k$. 因此 λ 对应的特征子空间也是 $\{cx^\lambda \mid c \in \mathbb{F}\}$.

(2) 设 $\alpha = \sum_{i=0}^n a_i x^i$, 记 $I = \{0 \leq i \leq n \mid a_i \neq 0\}$, 由 $\alpha \neq 0$ 知 I 非空. 对任意 $i \in I$, 由 (1) 知 \mathcal{A} 关于 $a_i x^i$ 的最小多项式为 $x - i$. 故 $f(x) = \prod_{i \in I} (x - i)$ 是 \mathcal{A} 关于 $a_i x^i$ ($\forall i$) 的一个化零多项式, 从而是 \mathcal{A} 关于 α 的一个化零多项式, 再由 $x - i \nmid \frac{f(x)}{x - i}$ ($\forall i$) 知 $f(x)$ 是 \mathcal{A} 关于 α 的最小多项式. \square

习题 9.6.9 设 V 是 \mathbb{F} 上的线性空间, $\mathcal{A} \in \mathcal{L}(V)$, S 是 \mathcal{A} 的所有特征值的集合, W_λ 是 λ 对应的根子空间. 证明: 若 $V = \bigoplus_{\lambda \in S} W_\lambda$, 则对于任意 $\alpha \in V$, $d_{\mathcal{A}, \alpha}$ 存在并且可以在 $\mathbb{F}[x]$ 中分解为一次因式的乘积.

证明 由 $V = \bigoplus_{\lambda \in S} W_\lambda$, 对任意 $\alpha \in V$, 存在两两不同的 $\lambda_1, \dots, \lambda_k \in S$ 和 $\alpha_i \in W_{\lambda_i}$, 使得 $\alpha = \alpha_1 + \dots + \alpha_k$. 由根子空间的定义, 对任意 $1 \leq i \leq k$, 存在 $n_i \in \mathbb{N}$ 使得 $\alpha_i \in \text{Ker}(\lambda \mathcal{I} - \mathcal{A})^{n_i}$, 即 $(\mathcal{A} - \lambda \mathcal{I})^{n_i} \alpha_i = \mathbf{0}$. 设 $f(x) = \prod_{i=1}^k (x - \lambda_i)^{n_i}$, 则 $f(\mathcal{A})\alpha_i = \mathbf{0}, \forall i$. 故 $f(\mathcal{A})\alpha = f(\mathcal{A})\left(\sum_{i=1}^k \alpha_i\right) = \mathbf{0}$, 即 f 是 \mathcal{A} 关于 α 的一个化零多项式, 进而存在 \mathcal{A} 关于 α 的最小多项式 $d_{\mathcal{A}, \alpha}$. 由 $d_{\mathcal{A}, \alpha} \mid f$ 可知 $d_{\mathcal{A}, \alpha}$ 可以在 $\mathbb{F}[x]$ 中分解为一次因式的乘积. \square

习题 9.7.5 设 V 是 \mathbb{F} 上的线性空间, $\mathcal{A} \in \mathcal{L}(V)$ 是幂零变换. 对 $\deg(d_{\mathcal{A}})$ 应用数学归纳法证明: V 可以分解为若干 \mathcal{A} -循环子空间的直和.

证明

注 2.0.21 结合习题 9.4.10 与习题 9.7.5 可得线性空间 V 的两种分解方式:

$$V = \bigoplus_{i=1}^m \mathcal{A}^{i-1}(U) = \bigoplus_{i \in I} \mathbb{F}[\mathcal{A}]\alpha_i,$$

其中 m 是 \mathcal{A} 的幂零指数, I 可能为不可数指标集 (这在下图中也能体现).

	$\mathbb{F}[\mathcal{A}]\alpha_1$	$\mathbb{F}[\mathcal{A}]\alpha_2$	$\mathbb{F}[\mathcal{A}]\alpha_3$	\cdots
U	α_1	α_2	α_3	\cdots
$\mathcal{A}(U)$	$\mathcal{A}\alpha_1$	$\mathcal{A}\alpha_2$	$\mathcal{A}\alpha_3$	\cdots
$\mathcal{A}^2(U)$	$\mathcal{A}^2\alpha_1$	$\mathcal{A}^2\alpha_2$	$\mathcal{A}^2\alpha_3$	\cdots
\vdots	\vdots	\vdots	\vdots	
$\mathcal{A}^{m-1}(U)$	$\mathcal{A}^{m-1}\alpha_1$	$\mathcal{A}^{m-1}\alpha_2$	$\mathcal{A}^{m-1}\alpha_3$	\cdots

习题 10.1.3 证明: 下列方阵 $A = (a_{ij}) \in \mathbb{R}^{n \times n}$ 是正定方阵, $1 \leq i, j \leq n$.

$$(1) a_{ij} = \frac{1}{i+j} \quad (2) a_{ij} = \frac{1}{(i+j)^2} \quad (3) a_{ij} = \frac{(i+j)!}{i!j!}$$

证明 由定理 10.1(1), 只需构造实内积使得 A 为其在给定基下的度量矩阵.

(1) A 是 $\mathbb{R}_n[x]$ 上内积 $\rho(f, g) = \int_0^1 xf(x)g(x)dx$ 在基 $1, x, x^2, \dots, x^{n-1}$ 下的度量矩阵.

(2) A 是 $\mathbb{R}_n[x]$ 上内积 $\rho(f, g) = \int_0^1 x|\ln x|f(x)g(x)dx$ 在基 $1, x, x^2, \dots, x^{n-1}$ 下的度量矩阵.

(3) A 是 $\mathbb{R}_n[x]$ 上内积 $\rho(f, g) = \int_0^{+\infty} x^2 f(x)g(x)e^{-x}dx$ 在基 $1, \frac{x}{2!}, \frac{x^2}{3!}, \dots, \frac{x^{n-1}}{n!}$ 下的度量矩阵. \square

习题 10.1.4 设 $V = \mathbb{R}^{m \times n}$, $S \in \mathbb{R}^{m \times m}$. 证明: $\rho(X, Y) = \text{tr}(X^T SY)$ 是 V 上的内积 $\iff S$ 是正定的.

证明 \Rightarrow : 对任意 $0 \neq x \in \mathbb{R}^m$, 取 $X \in \mathbb{R}^{m \times n}$ 为以 x 为第一列、其余元素均为 0 的矩阵. 由内积的正定性、 $X \neq O$ 得 $\text{tr}(X^T SX) = x^T Sx > 0$, 为证 S 正定只需证 S 对称. 由内积的对称性得 $\text{tr}(X^T SY) = \text{tr}(Y^T SX) = \text{tr}(X^T S^T Y)$, 即 $\text{tr}(X^T (S - S^T)Y) = 0, \forall X, Y \in \mathbb{R}^{m \times n}$. 依次取 $X = E_{1i}, Y = E_{j1}$ ($1 \leq i, j \leq n$) 可得 $S - S^T = O$, 即 $S = S^T$.

\Leftarrow : 设 $S = P^T P$, 其中 $P \in \mathbb{R}^{m \times m}$ 可逆. 则对任意 $X \in \mathbb{R}^{m \times n}$,

$$\text{tr}(X^T SX) = \text{tr}(X^T P^T P X) = P X \text{ 所有元素平方和 } \geq 0,$$

等号成立当且仅当 $PX = O$ 即 $X = O$. 又对称性和双线性显见, 故 ρ 是 V 上的内积. \square

习题 10.1.5 (2) 设 $w(x) \in C[0, 1]$. 若 $\rho(f, g) = \int_0^1 f(x)g(x)w(x)dx$ 是 $V = \mathbb{R}_n[x]$ 上的内积, 是否一定有 $w(x) \geq 0, \forall x \in [0, 1]$?

解 不一定. 可构造 $[0, 1]$ 上不恒非负的 $2n$ 次多项式 $w(x)$ 使 $w(x) - 1$ 与任意 $Q(x) \in \mathbb{R}_{2n-1}[x]$ 正交, 从而对任意 $x \in \mathbb{R}_n[x]$ 都有 $\int_0^1 f^2(x)w(x)dx = \int_0^1 f^2(x)dx$, 即 ρ 满足正定性.

取 $R(x) = x^{2n}(x-1)^{2n}, X_{2n}(x) = \frac{d^{2n}}{dx^{2n}}R(x)$. 由高阶分部积分公式,

$$\begin{aligned} \int_0^1 X_{2n}(x)Q(x)dx &= \int_0^1 R^{(2n)}(x)Q(x)dx = [Q(x)R^{(2n-1)}(x) - Q'(x)R^{(2n-2)}(x) + \dots \\ &\quad - Q^{(2n-1)}(x)R(x)] \Big|_0^1 + \int_0^1 Q^{(2n)}(x)R(x)dx. \end{aligned}$$

由 $R(0) = R'(0) = \dots = R^{(2n)}(0) = R(1) = R'(1) = \dots = R^{(2n)}(1) = 0$ 得

$$\int_0^1 X_{2n}(x)Q(x)dx = \int_0^1 Q^{(2n)}(x)R(x)dx = 0.$$

特别地, 由 $\int_0^1 X_{2n}(x) dx = 0$ 及 $\deg(X_{2n}) = 2n$ 得存在 $x \in [0, 1]$ 使 $X_{2n}(x) < 0$. 通过选取适当的常数 $c \in \mathbb{R}^+$, 可以使得 $w(x) = cX_{2n}(x) + 1$ 在 $[0, 1]$ 上有取值为负的点. \square

注 2.0.22 本题的构造来自 **Legendre 多项式**, 这是一类正交多项式.

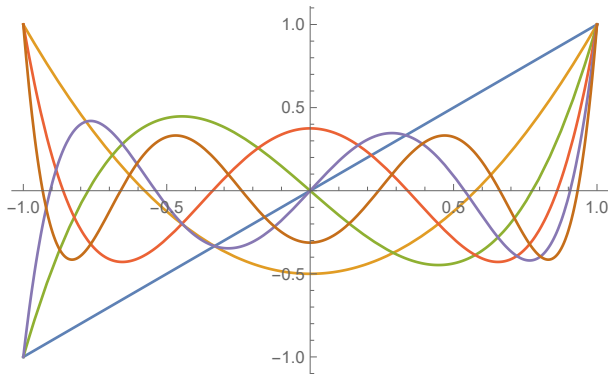


图 2.4: Legendre 多项式 $P_1(x) \sim P_6(x)$

Legendre 多项式的通式可由 **Rodrigues 公式** 给出: $P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} (x^2 - 1)^n$.

$$P_0(x) = 1$$

$$P_1(x) = x$$

$$P_2(x) = \frac{1}{2} (3x^2 - 1)$$

$$P_3(x) = \frac{1}{2} (5x^3 - 3x)$$

$$P_4(x) = \frac{1}{8} (35x^4 - 30x^2 + 3)$$

$$P_5(x) = \frac{1}{8} (63x^5 - 70x^3 + 15x)$$

$$P_6(x) = \frac{1}{16} (231x^6 - 315x^4 + 105x^2 - 5)$$

习题 10.1.8 设 $\alpha_1, \alpha_2, \dots, \alpha_m$ 是 n 维实内积空间 (V, ρ) 中的一组非零向量. 证明:

(1) 若对于任意 $i < j$ 都有 $\rho(\alpha_i, \alpha_j) < 0$, 则 $m \leq n + 1$.

(2) 若对于任意 $i < j$ 都有 $\rho(\alpha_i, \alpha_j) \leq 0$, 则 $m \leq 2n$.

证明 (1) 用反证法. 假设 $\alpha_1, \dots, \alpha_{n+2} \in V$ 是满足对任意 $i < j$ 都有 $\rho(\alpha_i, \alpha_j) < 0$ 的一组非零向量, 则存在不全为 0 的 x_1, \dots, x_{n+1} 使得 $\sum_{i=1}^{n+1} x_i \alpha_i = \mathbf{0}$, 记 $S^+ = \{1 \leq i \leq n+1 \mid x_i >$

$0\}$, $S^- = \{1 \leq i \leq n+1 \mid x_i < 0\}$. 由 $\sum_{i=1}^{n+1} x_i \rho(\alpha_i, \alpha_{n+2}) = 0$ 及 $\rho(\alpha_i, \alpha_{n+2}) < 0$ ($\forall i$) 知 S^+, S^-

均非空. 设 $\beta^+ = \sum_{i \in S^+} x_i \alpha_i, \beta^- = \sum_{i \in S^-} x_i \alpha_i$, 由 $\rho(\beta^+, \alpha_{n+2}) < 0$ 及 $\rho(\beta^-, \alpha_{n+2}) > 0$ 知 β^+, β^- 均非零向量. 由 $\beta^+ + \beta^- = \mathbf{0}$ 得 $0 > \rho(\beta^+, -\beta^-) = - \sum_{i \in S^+} \sum_{j \in S^-} x_i x_j \rho(\alpha_i, \alpha_j) < 0$, 矛盾.

(2) 对 n 归纳. 当 $n = 1$ 时任意 α_i 均共线, 只能有 $m \leq 2$, 结论成立.

假设结论对 $n - 1$ 维实内积空间成立. 若存在非零向量 $\alpha_1, \dots, \alpha_{2n+1} \in V$ 满足对任意 $i < j$ 都有 $\rho(\alpha_i, \alpha_j) \leq 0$, 将 α_{2n+1} 扩充为 V 的标准正交基 S , α_{2n+1} 在 S 下的坐标为 \mathbf{e}_n . 由 $\rho(\alpha_i, \alpha_{2n+1}) \leq 0$ ($1 \leq i \leq 2n$) 知 α_i ($1 \leq i \leq 2n$) 在 S 下第 n 个坐标分量均不大于 0. 又至多还存在一个向量与 α_{2n+1} 共线, 故可不妨设 $\alpha_1, \dots, \alpha_{2n-1}$ 在 S 下前 $n - 1$ 个分量均不全为 0. 由归纳假设, 存在 $1 \leq i < j \leq 2n - 1$, α_i 与 α_j 在 S 下前 $n - 1$ 个分量内积大于 0, 而它们的第 n 个分量均不大于 0, 从而 $\rho(\alpha_i, \alpha_j) > 0$, 矛盾. 故结论对 n 维实内积空间也成立. \square

习题 10.2.2 (2) 求 $V = \mathbb{R}[x]$ 上的一个内积 ρ , 使得 S 构成 (V, ρ) 的正交基.

$S = \{T_n(x) \mid n \in \mathbb{N}\}$, 其中 $T_n(\cos \theta) = \cos(n\theta)$ 是第一类 Chebyshev 多项式.

$$\text{解 由 } \int_{-1}^1 T_n(x) T_m(x) \frac{dx}{\sqrt{1-x^2}} = \begin{cases} 0, & n \neq m \\ \pi, & n = m = 0 \\ \frac{\pi}{2}, & n = m \neq 0 \end{cases} \text{ 可知内积 } \rho(f, g) = \int_{-1}^1 f(x) g(x) \frac{dx}{\sqrt{1-x^2}}$$

满足要求. \square

习题 10.2.2 (3) 求 $V = \mathbb{R}[x]$ 上的一个内积 ρ , 使得 S 构成 (V, ρ) 的正交基.

$S = \{U_n(x) \mid n \in \mathbb{N}\}$, 其中 $U_n(\cos \theta) = \frac{\sin((n+1)\theta)}{\sin \theta}$ 是第二类 Chebyshev 多项式.

$$\text{解 由 } \int_{-1}^1 U_n(x) U_m(x) \sqrt{1-x^2} dx = \begin{cases} 0, & n \neq m \\ \frac{\pi}{2}, & n = m \end{cases} \text{ 可知内积 } \rho(f, g) = \int_{-1}^1 f(x) g(x) \sqrt{1-x^2} dx$$

满足要求. \square

习题 10.2.5 设 $A \in \mathbb{R}^{m \times n}, B \in \mathbb{R}^{k \times n}, \alpha \in \mathbb{R}^{m \times 1}, \mathbf{x} \in \mathbb{R}^{n \times 1}$. 求 $\|A\mathbf{x} - \alpha\|$ 在条件 $B\mathbf{x} = \mathbf{0}$ 下的最小值.

解 设 $B = P \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} Q$, 其中 $P \in \mathbb{R}^{k \times k}, Q \in \mathbb{R}^{n \times n}$ 可逆, 则 $B\mathbf{x} = \mathbf{0} \iff \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} Q\mathbf{x} = \mathbf{0} \iff Q\mathbf{x} = \begin{pmatrix} O & O \\ O & I_{n-r} \end{pmatrix} \mathbf{y}, \mathbf{y} \in \mathbb{R}^{n \times 1}$. 记 $\tilde{A} = A Q^{-1} \begin{pmatrix} O & O \\ O & I_{n-r} \end{pmatrix}$, 则问题转化为求 $\|\tilde{A}\mathbf{y} - \alpha\|$ 在条件 $\mathbf{y} \in \mathbb{R}^{n \times 1}$ 下的最小值. 此后按例 (6.9) 求最小二乘解方法求解即可. \square

习题 10.2.6 证明: 在实内积空间中, 不含零向量的正交向量组一定是线性无关的.

证明 任取正交向量组中非零向量 $\alpha_1, \dots, \alpha_n$. 设 $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ 使得 $\lambda_1\alpha_1 + \dots + \lambda_n\alpha_n = \mathbf{0}$, 则对任意 $1 \leq i \leq n$, 由 $0 = \rho\left(\alpha_i, \sum_{j=1}^n \lambda_j\alpha_j\right) = \lambda_i\rho(\alpha_i, \alpha_i)$ 可得 $\lambda_i = 0$, 故 $\alpha_1, \dots, \alpha_n$ 线性无关, 从而不含零向量的正交向量组一定是线性无关的. \square

习题 10.2.7 举例: U 是实内积空间 V 的子空间, $U = (U^\perp)^\perp$, $V \neq U \oplus U^\perp$.

解 设 $V = \text{Span}(\mathbf{1}, \mathbf{e}_1, \mathbf{e}_2, \dots)$, 对 $\mathbf{x} = (x_n)_{n \in \mathbb{N}^*}$ 与 $\mathbf{y} = (y_n)_{n \in \mathbb{N}^*}$ 定义内积 $\rho(\mathbf{x}, \mathbf{y}) = \sum_{n=1}^{\infty} \frac{x_n y_n}{n^2}$. 设 $U = \text{Span}(\mathbf{e}_1, \mathbf{e}_3, \dots)$, 则 $U^\perp = \text{Span}(\mathbf{e}_2, \mathbf{e}_4, \dots)$. \square

习题 10.2.7 旧 举例: 实内积空间 V 有标准正交基, 并且 $\dim V$ 是不可数的.

解 设 V 是实线性空间, $\dim V$ 不可数, $\{e_i \mid i \in I\}$ 是 V 的 Hamel 基. 对任意 $\alpha, \beta \in V$, 存在有限子集 $I_1, I_2 \subset I$, 使得 $\alpha = \sum_{i_1 \in I_1} \lambda_{i_1} e_{i_1}, \beta = \sum_{i_2 \in I_2} \mu_{i_2} e_{i_2}$, 其中 $\lambda_{i_1}, \mu_{i_2} \in \mathbb{R}$. 对 α, β 定义内积 $\rho(\alpha, \beta) := \sum_{i_1 \in I_1} \sum_{i_2 \in I_2} \lambda_{i_1} \mu_{i_2} \delta_{i_1 i_2}$, 则基 $\{e_i \mid i \in I\}$ 是标准正交向量组. \square

习题 10.2.8 设 U_1, U_2 是实内积空间 V 的子空间.

- (1) 证明: $(U_1 + U_2)^\perp = U_1^\perp \cap U_2^\perp$, $(U_1 \cap U_2)^\perp \supset U_1^\perp + U_2^\perp$.
- (2) 证明: 若 $U_1 + U_2$ 是有限维的, 则 $(U_1 \cap U_2)^\perp = U_1^\perp + U_2^\perp$.
- (3) 举例: $(U_1 \cap U_2)^\perp \neq U_1^\perp + U_2^\perp$.

解 (1) ① 对任意 $\alpha \in (U_1 + U_2)^\perp$, $\alpha \in U_i^\perp$ ($i = 1, 2$), 因此 $\alpha \in U_1^\perp \cap U_2^\perp$, $(U_1 + U_2)^\perp \subset U_1^\perp \cap U_2^\perp$. 对任意 $\beta \in U_1^\perp \cap U_2^\perp$ 与 $\gamma = \gamma_1 + \gamma_2 \in U_1 + U_2$, 其中 $\gamma_i \in U_i$ ($i = 1, 2$), 有 $\rho(\beta, \gamma) = \rho(\beta, \gamma_1) + \rho(\beta, \gamma_2) = 0$, 因此 $\beta \in (U_1 + U_2)^\perp$, $U_1^\perp \cap U_2^\perp \subset (U_1 + U_2)^\perp$. 故 $(U_1 + U_2)^\perp = U_1^\perp \cap U_2^\perp$.

② 由 $U_i^\perp \subset (U_1 \cap U_2)^\perp$ ($i = 1, 2$) 得 $(U_1 \cap U_2)^\perp \supset U_1^\perp + U_2^\perp$.

(2) 将 (1) 中第一式的 U_i 用 U_i^\perp ($i = 1, 2$) 替换得到

$$(U_1^\perp + U_2^\perp)^\perp = (U_1^\perp)^\perp \cap (U_2^\perp)^\perp.$$

由定理 10.6 知, 若 U 是有限维的, 则 $(U^\perp)^\perp = U$. 由 $U_1 + U_2$ 是有限维的可知 U_1, U_2 都是有限维的. 又从 $U_1^\perp \subset U_1^\perp + U_2^\perp$ 可得 $(U_1^\perp + U_2^\perp)^\perp \subset (U_1^\perp)^\perp = U_1$. 因此上式两端均是有限维子空间, 两边同时取正交补就得证.

(3) 设 $V = \ell^2$, $U_1 = \text{Span}\left(\left(1, \frac{1}{2}, \frac{1}{2^2}, \frac{1}{2^3}, \dots\right)\right)$, $U_2 = \text{Span}(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \dots)$. 则 $U_1 \cap U_2 = \{\mathbf{0}\}$, $U_2^\perp = \{\mathbf{0}\}$, 但 $U_1^\perp \neq V$. \square

习题 10.2.9 设 V 是实内积空间, I 是指标集合, $\{e_i\}_{i \in I}$ 是 V 中的一个极大标准正交向量组. 对于任意 $\alpha \in V$, 是否都有 $\|\alpha\|^2 = \sum_{i \in I} (e_i, \alpha)^2$?

解 不一定. 设 $V = \text{Span}(\mathbf{1}, e_2, e_3, \dots)$, 对 $\mathbf{x} = (x_n)_{n \in \mathbb{N}^*}, \mathbf{y} = (y_n)_{n \in \mathbb{N}^*} \in V$ 定义内积 $\rho(\mathbf{x}, \mathbf{y}) = \sum_{n=1}^{\infty} \frac{x_n y_n}{n^2}$. 易证 $\{e_2, e_3, \dots\}$ 是极大标准正交向量组 (只需注意到 $e_1 \notin V$), 但 $\|\mathbf{1}\|^2 = \sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$, $\sum_{i \in I} (e_i, \mathbf{1})^2 = \sum_{n=2}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6} - 1$. \square

习题 10.3.2 (3) 设 V 是实内积空间, $\mathcal{A}: V \rightarrow V$ 是任意映射.

举例: 线性映射 \mathcal{A} 满足 $(\mathcal{A}\alpha, \mathcal{A}\beta) = (\alpha, \beta), \forall \alpha, \beta \in V$, 但 \mathcal{A} 不是可逆映射.

解 设 $V = \mathbb{R}[x]$, 内积 $\left(\sum_i a_i x^i, \sum_i b_i x^i\right) = \sum_i a_i b_i$. 线性变换 $\mathcal{A}: f(x) \mapsto xf(x)$ 保内积但不是满射. \square

习题 10.3.4 设 V 是实内积空间, $\mathcal{A} \in \mathcal{L}(V)$ 可逆. 证明: 若 \mathcal{A} 保持任意两个向量的正交关系, 即

$$\alpha \perp \beta \implies \mathcal{A}\alpha \perp \mathcal{A}\beta, \quad \forall \alpha, \beta \in V,$$

则存在 $\lambda > 0$, 使得 $\lambda\mathcal{A}$ 是正交变换.

证明 ① 若 $V = \{\mathbf{0}\}$, 则结论平凡.

② 若 $V \neq \{\mathbf{0}\}$, 任取 $\mathbf{0} \neq \alpha \in V$, 由 \mathcal{A} 是单射知 $\mathcal{A}\alpha \neq \mathbf{0}$, 记 $\lambda = \frac{\|\alpha\|}{\|\mathcal{A}\alpha\|} > 0$. 对任意 $\mathbf{0} \neq \beta \in V$, 记 $\mu = \frac{\|\beta\|}{\|\alpha\|} > 0$, 由 $(\beta - \mu\alpha, \beta + \mu\alpha) = (\beta, \beta) - (\mu\alpha, \mu\alpha) = 0$ 得 $\mathcal{A}(\beta - \mu\alpha) \perp \mathcal{A}(\beta + \mu\alpha)$, 即 $\|\mathcal{A}\beta\| = \mu\|\mathcal{A}\alpha\| = \frac{\|\beta\|}{\lambda}$. 结合 $\mathcal{A}(\mathbf{0}) = \mathbf{0}$ 可知 $\|\lambda\mathcal{A}\gamma\| = \|\gamma\|, \forall \gamma \in V$. 由定义知 $\lambda\mathcal{A}$ 是正交变换. \square

注 2.0.23 若 \mathcal{A}^* 存在, 可参考 <https://math.stackexchange.com/a/1790693> 中的做法.

习题 10.3.5 (3)(4) 设 V 是实内积空间, $\mathcal{A} \in \mathcal{L}(V)$ 是正交变换. 证明:

$$(3) \quad \text{Ker}(\mathcal{I} - \mathcal{A}) = \text{Ker}(\mathcal{I} - \mathcal{A})^2, \quad \text{Ker}(\mathcal{I} + \mathcal{A}) = \text{Ker}(\mathcal{I} + \mathcal{A})^2.$$

$$(4) \quad \text{若 } d_{\mathcal{A}}(x) = \prod_{i=1}^k (x - \lambda_i), \text{ 其中 } \lambda_i \in \mathbb{C}, \text{ 则 } |\lambda_i| = 1 \text{ 并且 } \lambda_1, \lambda_2, \dots, \lambda_k \text{ 两两不同.}$$

证明 (3) 只需证第一式, 将 \mathcal{A} 用 $-\mathcal{A}$ 替换就可得第二式.

引理 设 V 是实内积空间, $\mathcal{A} \in \mathcal{L}(V)$ 是正交变换, 则 $\text{Ker}(\mathcal{I} - \mathcal{A}) \perp \text{Im}(\mathcal{I} - \mathcal{A})$.

【引理的证明: 任取 $\alpha \in \text{Ker}(\mathcal{I} - \mathcal{A})$ 与 $\mathcal{A}\beta - \beta \in \text{Im}(\mathcal{I} - \mathcal{A})$, 由 $\mathcal{A}\alpha = \alpha$ 得

$$(\alpha, \mathcal{A}\beta - \beta) = (\alpha, \mathcal{A}\beta) - (\alpha, \beta) = (\mathcal{A}\alpha, \mathcal{A}\beta) - (\alpha, \beta) = 0,$$

故 $\text{Ker}(\mathcal{I} - \mathcal{A}) \perp \text{Im}(\mathcal{I} - \mathcal{A})$.】

任取 $x \in \text{Ker}(\mathcal{I} - \mathcal{A})^2$, 则 $(\mathcal{I} - \mathcal{A})x \in \text{Ker}(\mathcal{I} - \mathcal{A})$. 由引理知 $\text{Ker}(\mathcal{I} - \mathcal{A}) \cap \text{Im}(\mathcal{I} - \mathcal{A}) = \{\mathbf{0}\}$, 因此 $(\mathcal{I} - \mathcal{A})x = \mathbf{0}$, 即 $x \in \text{Ker}(\mathcal{I} - \mathcal{A})$, $\text{Ker}(\mathcal{I} - \mathcal{A})^2 \subset \text{Ker}(\mathcal{I} - \mathcal{A})$. 再结合 $\text{Ker}(\mathcal{I} - \mathcal{A}) \subset \text{Ker}(\mathcal{I} - \mathcal{A})^2$ 即得证.

(4) 将 V 复化为 $V^{\mathbb{C}} := \{u + iv \mid u, v \in V\}$, \mathcal{A} 诱导 $V^{\mathbb{C}}$ 上的线性变换 $\mathcal{A}^{\mathbb{C}} : u + iv \mapsto \mathcal{A}u + i\mathcal{A}v$. 容易验证 $\mathcal{A}^{\mathbb{C}}$ 的最小多项式即为 $d_{\mathcal{A}}$ 且 $\mathcal{A}^{\mathbb{C}}$ 为酉变换. 由定理 9.15, $\lambda_1, \dots, \lambda_k$ 恰为 $\mathcal{A}^{\mathbb{C}}$ 的所有特征值, 于是 $|\lambda_i| = 1$ ($i = 1, \dots, k$).

对任意 λ_i , 任取 $\alpha \in \text{Ker}(\lambda\mathcal{I} - \mathcal{A}^{\mathbb{C}})$ 与 $\mathcal{A}^{\mathbb{C}}\beta - \lambda\beta \in \text{Im}(\lambda\mathcal{I} - \mathcal{A}^{\mathbb{C}})$, 由 $\mathcal{A}^{\mathbb{C}}\alpha = \lambda\alpha$ 得

$$(\alpha, \mathcal{A}^{\mathbb{C}}\beta - \lambda\beta) = (\alpha, \mathcal{A}^{\mathbb{C}}\beta) - (\alpha, \lambda\beta) = \frac{1}{\lambda} (\mathcal{A}^{\mathbb{C}}\alpha, \mathcal{A}^{\mathbb{C}}\beta) - \lambda(\alpha, \beta) = 0,$$

故 $\text{Ker}(\lambda\mathcal{I} - \mathcal{A}^{\mathbb{C}}) \perp \text{Im}(\lambda\mathcal{I} - \mathcal{A}^{\mathbb{C}})$. 从而 $\text{Ker}(\lambda\mathcal{I} - \mathcal{A}^{\mathbb{C}}) \cap \text{Im}(\lambda\mathcal{I} - \mathcal{A}^{\mathbb{C}}) = \{\mathbf{0}\}$. 类似 (3) 中引理可知 $\text{Ker}(\lambda\mathcal{I} - \mathcal{A}^{\mathbb{C}}) = \text{Ker}(\lambda\mathcal{I} - \mathcal{A}^{\mathbb{C}})^2$, 若 $d_{\mathcal{A}}(x)$ 有重根 λ_i , 则 $\frac{d_{\mathcal{A}}(x)}{x - \lambda_i}$ 也是 $\mathcal{A}^{\mathbb{C}}$ 的化零多项式, 与 $d_{\mathcal{A}}(x)$ 是最小多项式矛盾. 故 $\lambda_1, \dots, \lambda_k$ 两两不同. \square

习题 10.4.3 设 $V = \mathbb{R}^n$, 内积 $(x, y) = x^T G y$, $\mathcal{A}(x) = x + (\alpha, x)\beta$, 其中 $\alpha, \beta \in V$. 求 \mathcal{A}^* .

解 由 \mathcal{A} 的矩阵表示为 $I + \beta\alpha^T G$ 知 \mathcal{A}^* 的矩阵表示为 $G^{-1}(I + G\alpha\beta^T)G = I + \alpha\beta^T G$. 故 $\mathcal{A}^* : y \mapsto y + \alpha\beta^T G y = y + (\beta, y)\alpha$. \square

习题 10.4.5 设 $V = \mathbb{R}[x]$, 内积 $(f, g) = \int_{-1}^1 f(x)g(x)dx$, $\mathcal{A} \in \mathcal{L}(V) : f(x) \mapsto xf(-x)$. 求 \mathcal{A}^* .

解 由 $\int_{-1}^1 xf(-x)g(x)dx = \int_{-1}^1 f(x)\mathcal{A}^*g(x)dx$ 及 $\int_{-1}^1 xf(-x)g(x)dx = -\int_{-1}^1 xf(x)g(-x)dx$ 知 $\mathcal{A}^* : g(x) \mapsto -xg(-x)$. \square

习题 10.4.8 设 V 是实内积空间, $\mathcal{A} \in \mathcal{L}(V)$ 可逆, \mathcal{A}^* 存在. 证明:

- (1) \mathcal{A}^* 是单射, 并且 $(\text{Im } \mathcal{A}^*)^{\perp} = \{\mathbf{0}\}$.
- (2) 若 $(\mathcal{A}^{-1})^*$ 存在, 则 \mathcal{A}^* 可逆, 并且 $(\mathcal{A}^*)^{-1} = (\mathcal{A}^{-1})^*$.
- (3) 若 \mathcal{A}^* 可逆, 则 $(\mathcal{A}^{-1})^*$ 存在, 并且 $(\mathcal{A}^{-1})^* = (\mathcal{A}^*)^{-1}$.

证明 (1) 由

$$\begin{aligned} u \in \text{Ker } \mathcal{A}^* &\iff \mathcal{A}^*u = \mathbf{0} \iff (\mathcal{A}^*u, v) = 0, \forall v \in V \\ &\iff (u, \mathcal{A}v) = 0, \forall v \in V \iff u \in \text{Im } \mathcal{A}^{\perp} \end{aligned}$$

可知 $\text{Ker } \mathcal{A}^* = \text{Im } \mathcal{A}^\perp$. 用 \mathcal{A}^* 替换 \mathcal{A} 又可得 $\text{Ker } \mathcal{A} = (\text{Im } \mathcal{A}^*)^\perp$. 于是

$$\begin{aligned}\mathcal{A} \text{ 满} &\implies \text{Im } \mathcal{A}^\perp = \{\mathbf{0}\} \implies \text{Ker } \mathcal{A}^* = \{\mathbf{0}\} \implies \mathcal{A}^* \text{ 单}, \\ \mathcal{A} \text{ 单} &\implies \text{Ker } \mathcal{A} = \{\mathbf{0}\} \implies (\text{Im } \mathcal{A}^*)^\perp = \{\mathbf{0}\}.\end{aligned}$$

(2) 由

$$\begin{aligned}((\mathcal{A}^{-1})^* \mathcal{A}^* x, y) &= (\mathcal{A}^* x, \mathcal{A}^{-1} y) = (x, \mathcal{A} \mathcal{A}^{-1} y) = (x, y), \quad \forall x, y \in V, \\ \mathcal{A}^* (\mathcal{A}^{-1})^* x, y) &= ((\mathcal{A}^{-1})^* x, \mathcal{A} y) = (x, \mathcal{A}^{-1} \mathcal{A} y) = (x, y), \quad \forall x, y \in V\end{aligned}$$

即知 $\mathcal{A} \mathcal{A}^{-1} = \mathcal{A}^{-1} \mathcal{A} = \mathcal{I}$. 故 \mathcal{A}^* 可逆, 且 $(\mathcal{A}^*)^{-1} = (\mathcal{A}^{-1})^*$.

(3) 由

$$((\mathcal{A}^*)^{-1} x, y) = ((\mathcal{A}^*)^{-1} x, \mathcal{A} \mathcal{A}^{-1} y) = (\mathcal{A}^* (\mathcal{A}^*)^{-1} x, \mathcal{A}^{-1} y) = (x, \mathcal{A}^{-1} y), \quad \forall x, y \in V$$

可知 $(\mathcal{A}^{-1})^*$ 存在, 且 $(\mathcal{A}^{-1})^* = (\mathcal{A}^*)^{-1}$. □

习题 10.4.8 (4) 旧 设 V 是实内积空间, $\mathcal{A} \in \mathcal{L}(V)$ 可逆, \mathcal{A}^* 存在.

举例: \mathcal{A}^* 不是满射, 从而 \mathcal{A}^* 不可逆, $(\mathcal{A}^{-1})^*$ 不存在.

解 设 $V = \mathbb{R}[x]$, 内积 $\left(\sum_{i=0}^n f_i x^i, \sum_{i=0}^n g_i x^i\right) = \sum_{i=0}^n f_i g_i$. 设 $\mathcal{A} \in \mathcal{L}(V)$ 使得

$$\mathcal{A} x^i = \begin{cases} 1, & i = 0, \\ x^i - x^{i-1}, & i \geq 1. \end{cases}$$

易知 \mathcal{A} 是单射, 又 $\mathcal{A} \left(\sum_{i=0}^n x^i\right) = x^n$, 所以 \mathcal{A} 是满射, 从而 \mathcal{A} 可逆.

设 $\mathcal{B} \in \mathcal{L}(V)$ 使得 $\mathcal{B} x^i = x^i - x^{i+1}$, 则由

$$\begin{aligned}(\mathcal{A} x^i, x^j) &= (x^i - x^{i-1}, x^j) = \delta_{ij} - \delta_{i-1,j} = \delta_{ij} - \delta_{i,j+1} \\ &= (x^i, x^j - x^{j+1}) = (x^i, \mathcal{B} x^j), \quad \forall i \geq 1 \text{ 与 } j \geq 0\end{aligned}$$

及

$$(\mathcal{A}(1), x^j) = (1, x^j) = (1, x^j - x^{j+1}) = (1, \mathcal{B} x^j), \quad \forall j$$

可知 $\mathcal{B} = \mathcal{A}^*$. 显然 $1 \notin \text{Im } \mathcal{A}^*$, 即 \mathcal{A}^* 不是满射, 从而 \mathcal{A}^* 不可逆. 再由习题 10.4.8 (2) 的逆否命题知 $(\mathcal{A}^{-1})^*$ 不存在. □

注 2.0.24 虽然本题的反例只能是无限维内积空间, 但我们仍可以借助“无穷矩阵”进行构造. 这时的核心在于 \mathcal{A} 的“无穷矩阵表示”的每一列是基的像的坐标, 从而每列只有有限个元素非 0; 为使 $(\mathcal{A}^{-1})^*$ 不存在, 可以使 \mathcal{A} 的每列仍满足只有有限个元素非 0, 但存在某行有无穷个元素非 0. 以下是本题给出的构造中对应的几个“无穷矩阵”.

$$\mathcal{A} \text{ 对应于 } \begin{pmatrix} 1 & -1 & & & \\ & \ddots & \ddots & & \\ & & \ddots & -1 & \\ & & & 1 & \ddots \\ & & & & \ddots \end{pmatrix}, \quad \mathcal{A}^* \text{ 对应于 } \begin{pmatrix} 1 & & & & \\ -1 & \ddots & & & \\ & \ddots & \ddots & & \\ & & \ddots & -1 & 1 \\ & & & \ddots & \ddots \end{pmatrix},$$

$$\mathcal{A}^{-1} \text{ 对应于 } \begin{pmatrix} 1 & 1 & \cdots & 1 & \cdots \\ & \ddots & \ddots & \ddots & \ddots \\ & & \ddots & 1 & \ddots \\ & & & \ddots & \ddots \\ & & & & \ddots \end{pmatrix} \quad (\text{上三角全 1 方阵}).$$

习题 10.4.9 设 V 是实内积空间, $\mathcal{A} \in \mathcal{L}(V)$, \mathcal{A}^* 存在. 证明:

- (1) \mathcal{A} 是斜自伴变换当且仅当对于任意 $\alpha \in V$, 有 $(\alpha, \mathcal{A}\alpha) = 0$.
- (2) 若 V 是有限维的, \mathcal{A} 是斜自伴变换, 则 $(\mathcal{I} + \mathcal{A})^{-1}(\mathcal{I} - \mathcal{A})$ 是正交变换.
- (3) 若 \mathcal{A} 是斜自伴变换, $\mathcal{I} + \mathcal{A}$ 和 $\mathcal{I} - \mathcal{A}$ 是否一定是可逆变换?

证明 (1) \Rightarrow : 由 $(\alpha, \mathcal{A}\alpha) = (-\mathcal{A}\alpha, \alpha) = -(\alpha, \mathcal{A}\alpha)$ 即知 $(\alpha, \mathcal{A}\alpha) = 0, \forall \alpha \in V$.

\Leftarrow : 对任意 $x, y \in V$, 都有

$$0 = (x + y, \mathcal{A}(x + y)) = (x, \mathcal{A}x) + (x, \mathcal{A}y) + (y, \mathcal{A}x) + (y, \mathcal{A}y) = (x, \mathcal{A}y) + (y, \mathcal{A}x),$$

即

$$(x, \mathcal{A}y) = (-\mathcal{A}x, y), \quad \forall x, y \in V \implies \mathcal{A}^* = -\mathcal{A}.$$

(2) 先证明 $\mathcal{I} + \mathcal{A}$ 与 $\mathcal{I} - \mathcal{A}$ 均为可逆变换. 因为

$$(\mathcal{I} + \mathcal{A})x = \mathbf{0} \iff x = -\mathcal{A}x \iff \|x\|^2 = (x, -\mathcal{A}x) \stackrel{(1)}{=} 0 \iff x = \mathbf{0},$$

所以 $\mathcal{I} + \mathcal{A}$ 是单射, 又有限维空间线性变换单射与满射等价, 故 $\mathcal{I} + \mathcal{A}$ 是可逆变换. 同理可证 $\mathcal{I} - \mathcal{A}$ 是可逆变换. 于是 $(\mathcal{I} + \mathcal{A})^{-1}(\mathcal{I} - \mathcal{A})$ 可逆, 其逆为 $(\mathcal{I} - \mathcal{A})^{-1}(\mathcal{I} + \mathcal{A})$. 由例 (10.21),

只需证 $((\mathcal{I} + \mathcal{A})^{-1}(\mathcal{I} - \mathcal{A}))^* = (\mathcal{I} - \mathcal{A})^{-1}(\mathcal{I} + \mathcal{A})$. 而

$$\begin{aligned} ((\mathcal{I} + \mathcal{A})^{-1}(\mathcal{I} - \mathcal{A}))^* &= (\mathcal{I} - \mathcal{A})^* ((\mathcal{I} + \mathcal{A})^{-1})^* = (\mathcal{I} + \mathcal{A}) ((\mathcal{I} + \mathcal{A})^{-1})^* \\ &\stackrel{\text{习题 10.4.8 (2)}}{=} (\mathcal{I} + \mathcal{A}) ((\mathcal{I} + \mathcal{A})^*)^{-1} = (\mathcal{I} + \mathcal{A})(\mathcal{I} - \mathcal{A})^{-1}, \end{aligned}$$

则只需证 $(\mathcal{I} - \mathcal{A})^{-1}(\mathcal{I} + \mathcal{A}) = (\mathcal{I} + \mathcal{A})(\mathcal{I} - \mathcal{A})^{-1}$. 这可以参考习题 6.1.11 (1) 的证明 (这也是本题的矩阵版本).

(3) 不一定, 习题 10.4.5 中斜自伴变换 \mathcal{A} 即反例. 对任意 $0 \neq f \in V$, $\deg((\mathcal{I} \pm \mathcal{A})f) = \deg(f) + 1$, 故 $1 \notin \text{Im}(\mathcal{I} \pm \mathcal{A})$, $\mathcal{I} \pm \mathcal{A}$ 不是满射, 从而不可逆. \square

习题 10.4.10 设 V 是实内积空间, $\mathcal{A} \in \mathcal{L}(V)$, \mathcal{A}^* 和 $d_{\mathcal{A}}$ 都存在. 证明:

(1) $d_{\mathcal{A}^*} = d_{\mathcal{A}}$.

(2) \mathcal{A} 是规范变换 $\implies d_{\mathcal{A}}$ 无重根.

(3) \mathcal{A} 是规范变换 \iff 存在 $f(x) \in \mathbb{R}[x]$, 使得 $\mathcal{A}^* = f(\mathcal{A})$.

证明 (1) 注意到对 $f \in \mathbb{R}[x]$,

$$\begin{aligned} f(\mathcal{A}) = \mathcal{O} &\iff f(\mathcal{A})x = \mathbf{0}, \forall x \in V \\ &\iff (f(\mathcal{A})x, y) = 0, \forall x, y \in V \\ &\stackrel{\text{内积双线性}}{\iff} (x, f(\mathcal{A}^*)y) = 0, \forall x, y \in V \\ &\iff f(\mathcal{A}^*)y = \mathbf{0}, \forall y \in V \\ &\iff f(\mathcal{A}^*) = \mathcal{O}, \end{aligned}$$

因此 $d_{\mathcal{A}^*} = d_{\mathcal{A}}$.

(2) 同习题 10.3.5 (3)(4) 将 V 复化为 $V^{\mathbb{C}}$, 并由 \mathcal{A} 诱导出 $\mathcal{A}^{\mathbb{C}}$, 则 $d_{\mathcal{A}^{\mathbb{C}}} = d_{\mathcal{A}}$ 且 $\mathcal{A}^{\mathbb{C}}$ 为规范变换. 同习题 6.2.5 (2)(3) 证明中的法二可知 $\mathcal{A}^{\mathbb{C}}x = \lambda x \iff (\mathcal{A}^{\mathbb{C}})^*x = \bar{\lambda}x$.

• $\text{Ker}(\lambda\mathcal{I} - \mathcal{A}^{\mathbb{C}}) \perp \text{Im}(\lambda\mathcal{I} - \mathcal{A}^{\mathbb{C}})$. 任取 $x \in \text{Ker}(\lambda\mathcal{I} - \mathcal{A}^{\mathbb{C}})$ 与 $\mathcal{A}^{\mathbb{C}}y - \lambda y \in \text{Im}(\lambda\mathcal{I} - \mathcal{A}^{\mathbb{C}})$, 则

$$\begin{aligned} (x, \mathcal{A}^{\mathbb{C}}y - \lambda y) &= (x, \mathcal{A}^{\mathbb{C}}y) - (x, \lambda y) = ((\mathcal{A}^{\mathbb{C}})^*x, y) - (\bar{\lambda}x, y) \\ &= ((\mathcal{A}^{\mathbb{C}})^*x - \bar{\lambda}x, y) = (\mathbf{0}, y) = 0. \end{aligned}$$

因此 $\text{Ker}(\lambda\mathcal{I} - \mathcal{A}^{\mathbb{C}}) \perp \text{Im}(\lambda\mathcal{I} - \mathcal{A}^{\mathbb{C}})$.

• $\text{Ker}(\lambda\mathcal{I} - \mathcal{A}^{\mathbb{C}}) = \text{Ker}(\lambda\mathcal{I} - \mathcal{A}^{\mathbb{C}})^2$. 任取 $x \in \text{Ker}(\lambda\mathcal{I} - \mathcal{A}^{\mathbb{C}})^2$, 则 $(\lambda\mathcal{I} - \mathcal{A}^{\mathbb{C}})x \in \text{Ker}(\lambda\mathcal{I} - \mathcal{A}^{\mathbb{C}})$. 由上一点知 $\text{Ker}(\lambda\mathcal{I} - \mathcal{A}^{\mathbb{C}}) \cap \text{Im}(\lambda\mathcal{I} - \mathcal{A}^{\mathbb{C}}) = \mathbf{0}$, 因此 $(\lambda\mathcal{I} - \mathcal{A}^{\mathbb{C}})x = \mathbf{0}$, 即 $x \in \text{Ker}(\lambda\mathcal{I} - \mathcal{A}^{\mathbb{C}})$, $\text{Ker}(\lambda\mathcal{I} - \mathcal{A}^{\mathbb{C}})^2 \subset \text{Ker}(\lambda\mathcal{I} - \mathcal{A}^{\mathbb{C}})$. 结合 $\text{Ker}(\lambda\mathcal{I} - \mathcal{A}^{\mathbb{C}}) \subset \text{Ker}(\lambda\mathcal{I} - \mathcal{A}^{\mathbb{C}})^2$ 可知 $\text{Ker}(\lambda\mathcal{I} - \mathcal{A}^{\mathbb{C}}) = \text{Ker}(\lambda\mathcal{I} - \mathcal{A}^{\mathbb{C}})^2$.

由上面第二点, 若 $d_{\mathcal{A}^c}$ 有重根 λ_i , 则 $\frac{d_{\mathcal{A}^c}(x)}{x - \lambda_i}$ 也是 \mathcal{A}^c 的化零多项式, 与 $d_{\mathcal{A}^c}$ 是最小多项式矛盾. 故 $d_{\mathcal{A}^c}$ 即 $d_{\mathcal{A}}$ 无重根.

(3)

□

习题 10.5.6 (2) 设复线性空间 $V = \mathbb{C}^{m \times n}$, $\rho(X, Y) = \operatorname{tr}(X^H S Y)$, $\mathcal{A}(X) = P X Q$, 其中 S, P, Q 都是复方阵. 分别求矩阵对 (P, Q) 应满足的充分必要条件, 使得 \mathcal{A} 是复内积空间 (V, ρ) 上的酉变换/自伴变换/斜自伴变换/规范变换.

解 取定 $E_{11}, \dots, E_{m1}, \dots, E_{1n}, \dots, E_{mn}$ 为 V 的基, 则 \mathcal{A} 在这组基下的矩阵表示为 $Q^T \otimes P$ (推导参见习题 9.3.5 (2)). 由 $(E_{i_1 j_1}, E_{i_2 j_2}) = \operatorname{tr}(E_{j_1 i_1} S E_{i_2 j_2}) = \begin{cases} 0, & j_1 \neq j_2 \\ s_{i_1 i_2}, & j_1 = j_2 \end{cases}$ (其中 $S = (s_{ij})_{m \times m}$) 可知这组基的度量矩阵为 $I_n \otimes S$. 由定理 10.26(1) 知 \mathcal{A}^* 在这组基下的矩阵表示为 $(I_n \otimes S)^{-1} (Q^T \otimes P)^H (I_n \otimes S) = \overline{Q} \otimes (S^{-1} P^H S)$.

$$\textcircled{1} \mathcal{A} \text{ 是酉变换} \iff \mathcal{A}^* = \mathcal{A}^{-1} \iff (Q^T \overline{Q}) \otimes (P S^{-1} P^H S) = I \iff \begin{cases} Q^T \overline{Q} = \lambda I \\ P S^{-1} P^H S = \frac{1}{\lambda} I \end{cases} \quad (\text{其}$$

$$\text{中 } \lambda \in \mathbb{C} \setminus \{0\}) \iff \begin{cases} Q^H Q = \lambda I \\ P S^{-1} P^H S = \frac{1}{\lambda} I \end{cases} \quad (\text{其中 } \lambda \in \mathbb{C} \setminus \{0\}).$$

$$\textcircled{2} \mathcal{A} \text{ 是自伴变换} \iff \mathcal{A}^* = \mathcal{A} \iff \begin{cases} Q^T = \lambda \overline{Q} \\ P = \frac{1}{\lambda} S^{-1} P^H S \end{cases} \quad (\text{其中 } \lambda \in \mathbb{C} \setminus \{0\}) \iff$$

$$\begin{cases} Q = \lambda Q^H \\ P = \frac{1}{\lambda} S^{-1} P^H S \end{cases} \quad (\text{其中 } \lambda \in \mathbb{C} \setminus \{0\}).$$

$$\textcircled{3} \mathcal{A} \text{ 是斜自伴变换} \iff \mathcal{A}^* = -\mathcal{A} \iff \begin{cases} Q^T = -\lambda \overline{Q} \\ P = \frac{1}{\lambda} S^{-1} P^H S \end{cases} \quad (\text{其中 } \lambda \in \mathbb{C} \setminus \{0\}) \iff$$

$$\begin{cases} Q = -\lambda Q^H \\ P = \frac{1}{\lambda} S^{-1} P^H S \end{cases} \quad (\text{其中 } \lambda \in \mathbb{C} \setminus \{0\}).$$

$$\textcircled{4} \mathcal{A} \text{ 是规范变换} \iff \mathcal{A} \mathcal{A}^* = \mathcal{A}^* \mathcal{A} \iff \begin{cases} Q^H Q = Q Q^H \\ P S^{-1} P^H S = S^{-1} P^H S P \end{cases}. \quad \square$$

习题 10.6.2 设 V 是 \mathbb{F} 上的线性空间, $\operatorname{char} \mathbb{F} \neq 2$, ρ 是 V 上的双线性函数. 证明: 存在 V 上的对称双线性函数 f 和反对称双线性函数 g , 使得 $\rho = f + g$, 并且 f, g 是唯一的.

证明 设 $f(x, y) = \frac{\rho(x, y) + \rho(y, x)}{2}$, $g(x, y) = \frac{\rho(x, y) - \rho(y, x)}{2}$, 则 f, g 满足要求. 反过来, 若 $\rho = f + g$, 其中 f 是对称双线性函数, g 是反对称双线性函数, 则

$$\begin{cases} \rho(x, y) = f(x, y) + g(x, y), \\ \rho(y, x) = f(x, y) - g(x, y). \end{cases}$$

由此可知这种表达方式是唯一的. □

习题 10.6.3 (1) 设 V 是 \mathbb{F} 上的线性空间, ρ 是 V 上的非退化双线性函数. 证明: 对于任意非零向量 $\alpha \in V$, 存在 $\beta \in V$ 使得 $\rho(\alpha, \beta) \neq 0$ 且 $\rho(\beta, \alpha) \neq 0$.

证明 由 ρ 非退化, 对任意非零向量 $\alpha \in V$, 存在 $x, y \in V$, 使得 $\rho(\alpha, x) \neq 0$, $\rho(y, \alpha) \neq 0$.

① 若 $\rho(x, \alpha) \neq 0$ 或 $\rho(\alpha, y) \neq 0$, 则结论已成立.

② 若 $\rho(x, \alpha) = \rho(\alpha, y) = 0$. 由

$$\rho(\alpha, x + y) = \rho(\alpha, x) + \rho(\alpha, y) = \rho(\alpha, x) \neq 0,$$

$$\rho(x + y, \alpha) = \rho(x, \alpha) + \rho(y, \alpha) = \rho(y, \alpha) \neq 0$$

可知 $\beta = x + y$ 满足要求. □

习题 10.6.4 设 ρ 是线性空间 V 上的双线性函数. 证明下列两个叙述等价.

① 存在 $f, g \in V^*$, 使得 $\rho(\alpha, \beta) = f(\alpha)g(\beta)$.

② $\forall \alpha_i, \beta_i \in V$, $\begin{vmatrix} \rho(\alpha_1, \beta_1) & \rho(\alpha_1, \beta_2) \\ \rho(\alpha_2, \beta_1) & \rho(\alpha_2, \beta_2) \end{vmatrix} = 0$.

证明 ① \Rightarrow ②: 代入 $\rho(\alpha_i, \beta_i) = f(\alpha_i)g(\beta_i)$ ($i = 1, 2$) 直接计算可得该行列式为 0.

② \Rightarrow ①: 不妨设 ρ 取值不恒为 0, 取 $x, y \in V$ 使得 $\rho(x, y) \neq 0$. 则对任意 $\alpha, \beta \in V$, 由 $\begin{vmatrix} \rho(\alpha, \beta) & \rho(\alpha, y) \\ \rho(x, \beta) & \rho(x, y) \end{vmatrix} = 0$ 得 $\rho(\alpha, \beta) = \frac{\rho(\alpha, y)\rho(x, \beta)}{\rho(x, y)}$. 取 $f(\alpha) = \frac{\rho(\alpha, y)}{\rho(x, y)}$, $g(\beta) = \rho(x, \beta)$ 即满足要求. □

习题 10.6.5 设 \mathbb{F} 是二元域, $\alpha_1, \alpha_2, \dots, \alpha_m$ 是 $\mathbb{F}^{n \times 1}$ 中一组两两不同的向量. 证明:

(1) 若对于任意 i, j 都有 $\alpha_i^T \alpha_j = 0$, 则 $m \leq 2^{\lfloor \frac{n}{2} \rfloor}$.

(2) 若对于任意 i, j 都有 $\alpha_i^T \alpha_j = 1$, 则 $m \leq 2^{\lfloor \frac{n-1}{2} \rfloor}$.

证明 (1) 设 $A = (\alpha_1 \ \cdots \ \alpha_m) \in \mathbb{F}^{n \times m}$. 因为 $A^T A = O$, 由 Sylvester 秩不等式 (例 (4.9)) 得

$$\operatorname{rank}(A^T) + \operatorname{rank}(A) \leq \operatorname{rank}(A^T A) + n \implies r := \operatorname{rank}(A) \leq \frac{n}{2}.$$

再由 \mathbb{F} 是二元域, $m \leq |\operatorname{Span}(\alpha_1, \dots, \alpha_m)| = 2^r$, 结论得证.

(2) 当 n 是奇数时, 设 $\beta_i = \mathbf{1} - \alpha_i$, 其中 $\mathbf{1}$ 是全一向量, 则 β_1, \dots, β_m 两两不同, 且 $\beta_i^T \beta_j = 0, \forall i, j$. 由 (1) 得 $m \leq 2^{\frac{n-1}{2}}$. 当 n 是偶数时, 设 $\beta_i = \begin{cases} \mathbf{1} - \mathbf{u}_i, & \alpha_i = (0, \mathbf{u}_i) \\ \mathbf{u}_i, & \alpha_i = (1, \mathbf{u}_i) \end{cases}$, 则 $\beta_1, \dots, \beta_m \in \mathbb{F}^{(n-1) \times 1}$ 两两不同, 且 $\beta_i^T \beta_j = 0, \forall i, j$. 由 (1) 得 $m \leq 2^{\frac{n-2}{2}}$. \square

习题 10.6.6 设 $V = (\mathbb{F}^{2n}, \rho)$ 是辛空间, 其中辛内积 ρ 在 \mathbb{F}^{2n} 的标准基下的矩阵表示 $G = \begin{pmatrix} O & I_n \\ -I_n & O \end{pmatrix}$. 研究 V 上的辛变换、自伴变换、斜自伴变换、规范变换的性质.

解 参考 GTM 135 第 277 页 Symplectic Geometry. \square

第三章 考前自测清单

3.1 矩阵的相似

1. $f(A)$ 作用于 α 相当于 $f(\lambda)$ 与 α 数乘.
2. 怎样用秩表示几何重数.
3. A 与 B 相似 $\implies f(A)$ 与 $f(B)$ 相似 ($\forall f \in \mathbb{F}[x]$). 【特别地, 加上或减去一个纯量方阵也是一种多项式.】
4. 两两不同的特征值对应的特征向量线性无关. 【两种证法: Vandermonde 方阵/Lagrange 插值多项式.】
5. 用于相似的可逆方阵 P 的形状. 【这是后面多个重要证明的关键和灵感来源.】
6. “可相似对角化 \iff 每个特征值的几何重数等于代数重数”的证明. 【用计算秩的方法表示几何重数这一相似不变量.】
7. $f(A)$ 可逆当且仅当 $f(x)$ 与 $\varphi_A(x)$ 互素. 【如何表示 $f(A)$ 的特征值.】
8. A 的特征值 λ_i 的几何重数不超过 λ_i 的代数重数的证明. 【用 $\text{rank}(A - \lambda_i I)^k = n - n_i$ ($\forall k \geq n_i$).】
9. Cayley-Hamilton 定理的证明. 【算法流程描述/多项式技巧.】
10. 什么条件下 $\begin{pmatrix} A & C \\ O & B \end{pmatrix}$ 与 $\begin{pmatrix} A & O \\ O & B \end{pmatrix}$ 相似. 【如何把矩阵方程 $AX - XB = C$ 表示为线性方程组形式并构造相似? 按列向量进行分块乘法.】
11. 相似准对角/分块对角化.
12. d_A 整除 φ_A 且包含 φ_A 的所有根 $\implies \varphi_A \mid d_A^n$.
13. 准对角方阵的最小多项式. 【后面经常用.】
14. 为什么友方阵的最小多项式等于特征多项式? 【 A^k ($k = 0, 1, \dots, n-1$) 的第一列线性无关.】
15. 设 $A \in \mathbb{F}^{n \times n}$ 的所有特征值都属于 \mathbb{F} , 则 A 可以在 \mathbb{F} 上相似于对角方阵 $\iff d_A(x)$

无重根.

16. 化零多项式或特征多项式无重根 \implies 最小多项式无重根.
17. 置换方阵可在 \mathbb{C} 上相似对角化. 【最小多项式无重根.】
18. 特征多项式和最小多项式都相等 $\not\Rightarrow$ 两个矩阵相似. 【利用 Jordan 块以及准对角方阵的最小多项式分别构造满足特征多项式与最小多项式相等条件的矩阵.】
19. $(\alpha \quad A\alpha \quad \cdots \quad A^{k-1}\alpha)$ 列满秩, 其中 $k = \deg(d_{A,\alpha})$. 【注意这提供了一种求 $d_{A,\alpha}$ 的算法; 若 $d_{A,\alpha} = \varphi_A$, 则该列向量组是 \mathbb{F}^n 的一组基.】
20. 存在 α 使得 $d_{A,\alpha} = d_A$. 【类似中国剩余定理的证明.】
21. $d_A = \varphi_A \implies A$ 与 φ_A 的友方阵相似. 【如何构造相似的方阵?】
22. $AB = BA$ 且 $d_A = \varphi_A \implies \exists f \in \mathbb{F}[x]$ s.t. $B = f(A)$. 【先设出来再证明是想要的.】
23. $d_A = \varphi_A, g = \gcd(d_A, f) \implies \text{rank}(f(A)) = n - \deg(g)$. 【把 $f(A)$ 的零空间的元素用 $r(A)\alpha$ 的形式表示, 然后发现 $r(x)$ 必须是 $h(x)$ 的倍式, 而 $r(x)$ 本身次数又不超过 $n-1$.】
24. Jordan 块满足 $d_A = \varphi_A$. 【作为特例, $J_n(0)$ 与 x^n 的友方阵即 $J_n(0)^T$ 相似.】
25. 为什么有 Jordan 标准形? 【相似不变量.】
26. 如何从 Jordan 标准形读取代数重数、几何重数、 $d_A = \varphi_A$ 的等价条件、 $\text{rank}(A - \lambda_i I)^k$ 等信息? “ $\text{rank}(A - \lambda I) = n - 1 \implies A$ 的 Jordan 标准形即 $J_n(\lambda)$ ” 有何用处? 【参考习题 6.2.3 证明中用到的引理.】
27. 复数方阵与自身转置相似的证明. 【用“掉头 (中心对称)”方阵相似. 如何由此推出实方阵的情形?】
28. 可逆复数方阵可开方的证明. 【缺少“可逆”或“复数”时的反例.】
29. A 与 B 在 \mathbb{F} 上相似 $\iff xI - A$ 与 $xI - B$ 在 $\mathbb{F}[x]$ 上模相抵. 【多项式技巧体现在 x^j 前系数的比对; 先构造出 $AX = XB$, 再证明 X 可逆 (可记住 X^{-1} 的形式与 X “对偶”).】
30. 有理标准形. 【这个定理需要反复品读!】
31. 设 $A \in \mathbb{F}^{m \times m}, B \in \mathbb{F}^{n \times n}$. 若 $xI - B$ 与 $\text{diag}(f_1, \dots, f_n)$ 在 $\mathbb{F}[x]$ 上模相抵, 则 $I \otimes A - B \otimes I$ 与 $\text{diag}(f_1(A), \dots, f_n(A))$ 在 \mathbb{F} 上相抵. 【问题可能以求秩的形式出现.】
32. n 阶秩一方阵迹不等于 0, 则它相似于 E_{1n} .
33. 同阶方阵乘积交换不改变特征多项式.
34. 幂等方阵在扩域上一定可以相似对角化. 【化零多项式无重根 \implies 最小多项式无重根.】
35. 任意 $A \in \mathbb{F}^{n \times n}$ 可以在 \mathbb{F} 上相似于 Hessenberg 方阵. 【用有理标准形立即得到.】
36. 求所有与 A^2 相似的复数方阵 A . 【需要另一道习题作为引理.】

37. $\deg(d_A) \leq \text{rank}(A) + 1$. 【求 A 的最小多项式即求 A 的 Jordan 标准形的最小多项式. 而 A 的 Jordan 标准形中 0 的几何重数为 $n - \text{rank}(A)$, 因此在取这些 Jordan 块的最小多项式的最小公倍式时 x 的幂次至少减少 $n - \text{rank}(A) - 1$. A 可用 $\lambda I - A$ 代替.】

38. $\gcd(d_{A,\alpha}, d_{A,\beta}) = 1 \implies d_{A,\alpha+\beta} = d_{A,\alpha} d_{A,\beta}$. 【证明互相整除.】

3.2 正交方阵与酉方阵

1. 正交变换保向量范数和内积. 【证明逆命题用“一生二”.】

2. 二阶正交方阵有 2 种, 不要只记得其一! 【先固定第一个列向量.】

3. 正交方阵特征值模为 1, 实特征值只可能为 ± 1 . 虚特征值对应特征向量的“实部”与“虚部”向量正交且模相等. 【将 λ 设为 $\cos \theta + i \sin \theta$ 的形式.】

4. 若 $\lambda = \pm 1$ 是正交方阵 P 的特征值, 则存在单位实向量 α 为 λ 对应的特征向量. 【利用 $\lambda \in \mathbb{R}$ 进行实虚部对应并做标准化.】

5. Givens 方阵 $G_{ij}(\theta)$ 的性质: ① $\det(G_{ij}(\theta)) = 1$. 【旋转.】② $\text{rank}(G_{ij}(\theta) - I) \in \{0, 2\}$. ③ $G_{ij}(\theta)$ 可相似对角化. 【即说明 $\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$ (其特征值为 $\cos \theta \pm i \sin \theta$) 可相似对角化.】

6. Householder 方阵 H_v 的性质: ① $\det(H_v) = -1$. 【镜面反射.】② $\text{rank}(H_v - I) = 1$. 【Sylvester 行列式恒等式.】③ H_v 可相似对角化. 【即说明秩一方阵 vv^T 可相似对角化, 由习题 5.1.11 (2) 即等价于 $\text{tr}(vv^T) \neq 0$ 即 $\text{tr}(v^T v) \neq 0$ 即 $v \neq 0$.】

7. QR 分解: ① 对矩阵的要求? 在何种情况下有唯一性? ② 唯一性如何证明? 【正交的上三角方阵长什么样?】③ 左乘 Givens 方阵算法: 无法保证 R 对角元非负. ④ 左乘 Householder 方阵算法: 能保证 R 对角元都非负. v 如何构造? 【几何意义: 始末之差.】注意对子矩阵的处理方式. ⑤ 对列向量处理时, Givens 方阵是逐个清零, Householder 方阵则一步到位. 【故允许直接对子矩阵操作.】⑥ Gram-Schmidt 标准正交的待定系数法具体操作、 Q 与 R 的构造结果、怎样把 β_k 写成矩阵作用于 α_k 的形式.

8. Hadamard 不等式的证明 (为何需要分是否可逆讨论而不能直接 QR 分解?)、取等条件. 达到 Hadamard 界的例子: $H_0 = \begin{pmatrix} 1 \end{pmatrix}, H_{k+1} = \begin{pmatrix} H_k & H_k \\ -H_k & H_k \end{pmatrix} (k \in \mathbb{N})$.

9. 实方阵正交相似于准上三角方阵的形状、证明方法、推论 (正交方阵、对称实方阵、反对称实方阵、规范方阵).

10. 二阶规范实方阵的两种形状.

11. 实方阵 $\begin{pmatrix} A_1 & A_2 \\ O & A_3 \end{pmatrix}$ 规范 $\implies A_1, A_3$ 规范. 【先证 $A_2 = O$.】
12. 奇异值分解的证明.
13. 矩阵范数与最大奇异值的关系. 【联想单位圆到椭圆的线性变换.】
14. Moore–Penrose 广义逆的构造与唯一性. 【Penrose 等式.】
15. 复数对称方阵不能相似对角化的例子: $\begin{pmatrix} 1 & i \\ i & -1 \end{pmatrix}$.
16. 规范方阵可酉相似对角化. 借助这一点可以规范实方阵的许多证明在 \mathbb{C} 上进行简化.
17. 不与自身转置正交相似的实方阵: $\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}$. 【必须从 3 阶开始找.】
18. A 是 n 阶规范方阵 $\iff \|A\alpha\| = \|A^H\alpha\|$ ($\forall \alpha \in \mathbb{C}^{n \times 1}$).
19. 奇数阶正交方阵的行列式必为其特征值.

3.3 内积空间

1. 设 $V = \mathbb{R}^{n \times n}$, 内积 $\rho(X, Y) = \text{tr}(X^T Y)$. 设 $X = \begin{pmatrix} x_1 & \cdots & x_n \\ \vdots & \ddots & \vdots \\ x_{n^2-n+1} & \cdots & x_{n^2} \end{pmatrix}, Y = \begin{pmatrix} y_1 & \cdots & y_n \\ \vdots & \ddots & \vdots \\ y_{n^2-n+1} & \cdots & y_{n^2} \end{pmatrix}$, 则 $\rho(X, Y) = \sum_{i=1}^n x_i y_i$.
2. n 阶 Hilbert 矩阵是正定的. 【 $V = \mathbb{R}_n[x]$, $\rho(f, g) = \int_0^1 f(x)g(x) dx$ 在 V 的基 $1, x, x^2, \dots, x^{n-1}$ 下的度量矩阵 $G = \left(\frac{1}{i+j-1} \right)_{1 \leq i, j \leq n}$.】
3. 方阵 $G = ((i+j-2)!)_{1 \leq i, j \leq n}$ 正定. 【 $V = \mathbb{R}_n[x]$, $\rho(f, g) = \int_0^{+\infty} f(x)g(x) e^{-x} dx$ 在 V 的基 $1, x, x^2, \dots, x^{n-1}$ 下的度量矩阵.】
4. 设 G_1, G_2 分别是 ρ 在 V 的基 S_1, S_2 下的度量矩阵, P 是从 S_1 到 S_2 的过渡矩阵, 则 $G_2 = P^T G_1 P$. 【① 推论: 若 S_1, S_2 都是 V 的标准正交基, 则 P 为正交方阵. ② 用途: 给定度量矩阵找对应的基.】
5. \mathbb{R}^n 上 p 范数在 $p \geq 1$ 且 $p \neq 2$ 时不为任何内积导出. 【若某范数由内积诱导, 则该范数需满足平行四边形法则: $\|\alpha + \beta\|^2 + \|\alpha - \beta\|^2 = 2\|\alpha\|^2 + 2\|\beta\|^2$.】 $0 < p < 1$ 时不是范数.

$$\mathbf{【\|e_1 + e_2\|_p = 2^{\frac{1}{p}} > 2 = \|e_1\|_p + \|e_2\|_p. \mathbf{】}}$$