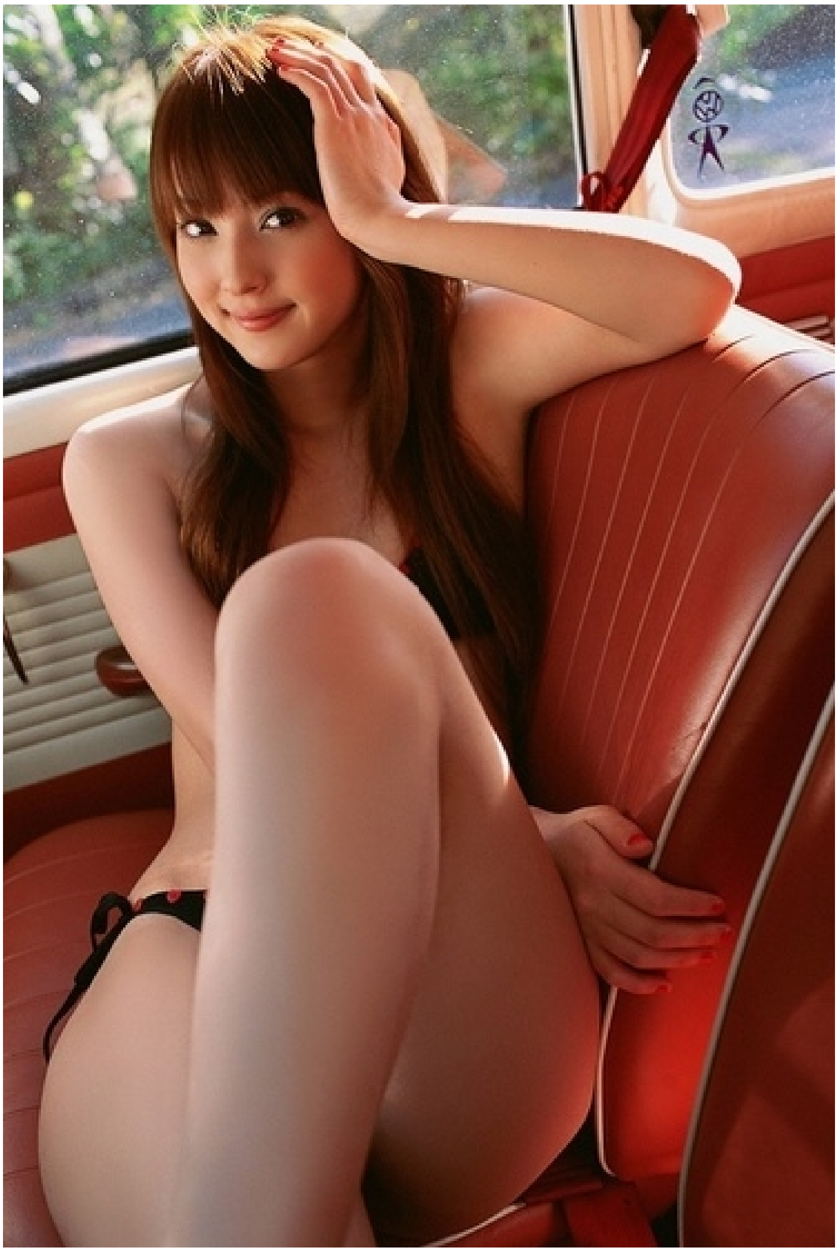
高等代數葵花寶典



博士家园 YINZHE

慾練神功 必先自宫
無慾無窮 方能行功

目录

1 考点整理

1.1 秩不等式总结

最简单的秩不等式是 $r(AB) \leqslant \min\{r(A), r(B)\}$, $r(A+B) \leqslant r(A)+r(B)$, 这里就不再证明了. 证明秩不等式的方法有矩阵的初等变换和几何的方法, 矩阵的方法在大多数时候是最方便的. 应用矩阵方法的根据就是下面这个不等式:
如果设

$$M=\begin{pmatrix} A & 0 \\ C & B \end{pmatrix}$$

这里 A, B 都是方阵, 那么 $r(M) \geqslant r(A)+r(B)$.
证明思路就是先把 A, B 都化成标准形: 设 $P_1AQ_1=E_r$, $P_2BQ_2=E_s$, 那么

$$\begin{pmatrix} P_1 & 0 \\ 0 & P_2 \end{pmatrix}\begin{pmatrix} A & 0 \\ C & B \end{pmatrix}\begin{pmatrix} Q_1 & 0 \\ 0 & Q_2 \end{pmatrix}=\begin{pmatrix} E_r & 0 \\ P_2CQ_1 & E_s \end{pmatrix}=\begin{pmatrix} E_r & 0 \\ D & E_s \end{pmatrix}$$

然后可以用 E_r 干掉 D 的前 r 列, 用 E_s 干掉 D 的后 s 行, M 就变成这个样子

$$M=\begin{pmatrix} E_r & 0 & 0 \\ 0 & * & 0 \\ 0 & 0 & E_s \end{pmatrix}$$

这样就得到了结论.
在证明有关秩的不等式的时候, 基本思路就是从准对角矩阵出发, 化成上三角或者下三角的分块矩阵, 来应用上面这个结论.

Frobenius 秩不等式的证明

设 A 是数域 $\mathbb{F}$ 上的 $m \times n$ 矩阵, B 是 $n \times k$ 矩阵, C 是 $k \times s$ 矩阵, 则

$$r(AB)+r(BC) \leqslant r(ABC)+r(B)$$

证明:

$$\begin{pmatrix} AB & 0 \\ B & BC \end{pmatrix} \xrightarrow{\text{第二行左乘以 } -A \text{ 加到第一行}} \begin{pmatrix} 0 & -ABC \\ B & BC \end{pmatrix}$$

$$\xrightarrow{\text{第一列右乘以 } -C \text{ 加到第二列}} \begin{pmatrix} 0 & -ABC \\ B & 0 \end{pmatrix}$$

比较两端的大的矩阵秩就得到结论.
注意这里其实是把真实的思路“倒着”写出来的. 在草纸上应该是从

$$\begin{pmatrix} ABC & 0 \\ 0 & B \end{pmatrix}$$

出发来做变换的. 为了形式上和前面的引理一致才改成了这个样子.

北大高代三题

问题 1: (2005北京大学) 设 A 是数域 F 上的 n 维向量空间上的线性变换, 求证

$$A^3 = I \Leftrightarrow r(I - A) + r(I + A + A^2) = n$$

证明: 注意到存在 $u(x), v(x)$ 使得 $u(x)(1 - x) + v(x)(1 + x^2 + x) = 1$, 考虑大的 $2n$ 阶的方阵

$$\begin{aligned}
&\begin{pmatrix} I - A & 0 \\ 0 & I + A + A^2 \end{pmatrix} \xrightarrow{\text{第一列乘以 } u(A) \text{ 加到第二列}} \begin{pmatrix} I - A & u(A)(I - A) \\ 0 & I + A + A^2 \end{pmatrix} \\
&\xrightarrow{\text{第二行乘以 } v(A) \text{ 加到第一列}} \begin{pmatrix} I - A & I \\ 0 & I + A + A^2 \end{pmatrix} \\
&\xrightarrow{\text{第二列乘以 } A - I \text{ 加到第一列}} \begin{pmatrix} 0 & I \\ A^3 - I & I + A + A^2 \end{pmatrix} \\
&\xrightarrow{\text{消去右下角的 } I + A + A^2} \begin{pmatrix} 0 & I \\ A^3 - I & 0 \end{pmatrix}
\end{aligned}$$

所以这个矩阵的秩是 n 当且仅当 $A^3 - I = 0$, 这就得到了证明.

问题 2: (2006北京大学) 设 A, B 是 n 阶矩阵, 证明

$$r(A - ABA) = r(A) + r(I - BA) - n$$

证明:

$$\begin{aligned}
&\begin{pmatrix} A & 0 \\ 0 & I - BA \end{pmatrix} \xrightarrow{\text{第一行左乘以 } B \text{ 加到第二行}} = \begin{pmatrix} A & 0 \\ BA & I - BA \end{pmatrix} \\
&\xrightarrow{\text{第一列加到第二列}} \begin{pmatrix} A & A \\ BA & I \end{pmatrix} \\
&\xrightarrow{\text{第二行左乘以 } -A \text{ 加到第一行}} \begin{pmatrix} A - ABA & 0 \\ BA & I \end{pmatrix} \\
&\xrightarrow{\text{第二列右乘以 } -BA \text{ 加到第一列}} \begin{pmatrix} A - ABA & 0 \\ 0 & I \end{pmatrix}
\end{aligned}$$

以上变换都是初等变换, 均保持秩不变. 从而等式成立.

问题 3: (2007北京大学) A,B 是两个 n 阶矩阵满足 $AB=BA$. 求证

$$r(A)+r(B)\geqslant r(A+B)+r(AB)$$

证明: 设 X 是齐次线性方程组 $AX=0$ 的解空间, Y 是齐次线性方程组 $BX=0$ 的解空间, Z 是齐次线性方程组 $ABX=BAX=0$ 的解空间, W 是齐次线性方程组 $(A+B)X=0$ 的解空间, 那么我们有 $X\subset Z, Y\subset Z$. 从而 $X+Y\subset Z$. 而且 $X\cap Y\subset W$. 由维数公式,

$$\dim X+\dim Y=\dim X\cap Y+\dim(X+Y)\leqslant \dim W+\dim Z$$

从而

$$n-r(A)+n-r(B)\leqslant n-r(A+B)+n-r(AB)$$

即 $r(A)+r(B)\geqslant r(A+B)+r(AB)$.
注: 这个题目也可以用矩阵的初等变换来做, 但是显然几何的方法更简洁.

科大教材课后习题

问题 1: 设 A,B 是数域 F 上的 n 阶方阵且满足 $AB=BA=0, r(A)=r(A^2)$, 求证

$$r(A+B)=r(A)+r(B)$$

下面分别用变换的方法和几何的方法给出两个证明, 它们的关键都在于对 $r(A)=r(A^2)$ 这个条件的利用.

变换的证明: 思路: 只要证明 $r(A+B)\geqslant r(A)+r(B)$ 即可. 用老办法, 从

$$\begin{pmatrix} A+B & 0\\ 0 & 0\end{pmatrix}$$

出发, 通过初等变换化为形如

$$\begin{pmatrix} B & *\\ 0 & A\end{pmatrix}$$

的矩阵, 就得到了结论. 在变换前先注意两点:
(1)由于 A^2 的列向量都是 A 的列向量的线性组合, 所以 $r(A)=r(A^2)$ 说明 A 的列向量组和 A^2 的列向量组是等价的, 也就是 A 的列向量可以被 A^2 的列向量线性表示出来. 设 α_i 是 A 的第 i 个列向量, 那么线性方程组 $A^2X=\alpha_i$ 有解 X_i, 令 $n\times n$ 矩阵 $P=(X_1,X_2,\cdots,X_n)$, 那么 $A^2P=A$.
(2)其次根据熟知的结论当 $r(A)=r(A^2)$ 时有 $r(A)=r(A^2)=r(A^3)=\cdots$
证明: 在 Frobenius 不等式

$$r(B)+r(ABC)\geqslant r(AB)+r(BC)$$

中令 $B=C=A$ 就不难得出结论.
下面进行变换:

$$\begin{aligned}\begin{pmatrix} A+B & 0 \\ 0 & 0 \end{pmatrix} &\xrightarrow{\text{第一行左乘以 } A \text{ 加到第二行}} = \begin{pmatrix} A+B & 0 \\ A^2 & 0 \end{pmatrix} \\ &\xrightarrow{\text{第一列右乘以 } A \text{ 加到第二列}} \begin{pmatrix} A+B & A^2 \\ A^2 & A^3 \end{pmatrix} \\ &\xrightarrow{\text{第二列右乘以 } -P \text{ 加到第一列}} \begin{pmatrix} B & A^2 \\ 0 & A^3 \end{pmatrix}\end{aligned}$$

从而 $r(A+B) \geqslant r(B)+r(A^3)=r(B)+r(A)$, 得证.

几何的方法: 这个方法的思路来自对条件 $r(A^2)=r(A)$ 的另一种解读. 首先证明下面的引理:

$$r(A^2)=r(A) \Rightarrow V=\mathrm{Ker}A \oplus \mathrm{Im}A.$$

证明: 设 $\eta_1, \eta_2, \cdots, \eta_{n-r}$ 是 $\mathrm{Ker}A$ 的一组基, $A\varepsilon_1, A\varepsilon_2, \cdots, A\varepsilon_r$ 是 $\mathrm{Im}A$ 的一组基. 我们只需要说明

$$\eta_1, \cdots, \eta_{n-r}, A\varepsilon_1, \cdots, A\varepsilon_r$$

构成 V 的一组基即可. 这样 $\mathrm{Ker}A+\mathrm{Im}A$ 就是直和, 且维数是 n, 就得到了引理的证明.
设不全为零的 $c_1, \cdots, c_n$ 使得

$$\sum_{i=1}^{n-r} c_i\eta_i + \sum_{j=1}^{r} c_iA\varepsilon_j = 0 \qquad (*)$$

那么在两边用 A 作用:

$$A^2(\sum_{j=1}^{r} c_i\varepsilon_j) = 0$$

但是 $A^2X=0$ 和 $AX=0$ 的解空间是相同的, 所以

$$A(\sum_{j=1}^{r} c_i\varepsilon_j) = 0$$

即

$$\sum_{j=1}^{r} c_iA\varepsilon_j = 0$$

代回$(*)$式可得

$$\sum_{i=1}^{n-r} c_i\eta_i = 0$$

所以 c_i 全部都是0, 引理得证.
最后完成原题的证明: 不难验证 Ker$A$ 和 ImA 都是 B 的不变子空间. 而且 B 在 ImA 上的限制是0. 所以在基 $\eta_1, \cdots, \eta_{n-r}, A\varepsilon_1, \cdots, A\varepsilon_r$ 下 A, B 的矩阵形如

$$A = \begin{pmatrix} 0 & 0 \\ 0 & A_1 \end{pmatrix} \quad B = \begin{pmatrix} B_1 & 0 \\ 0 & 0 \end{pmatrix}$$

现在 $r(A+B) = r(A) + r(B)$ 是很显然的了.

注: 如果对矩阵运算技巧足够熟悉的话, 这个题目可以直接看出来. 可以假设 A 是 Jordan 标准形, 那么 $r(A) = r(A^2)$ 说明 A 的0特征值对应的 Jordan 块都是一阶的. 从而 A 形如

$$\begin{pmatrix} A_1 & & & \\ & 0 & & \\ & & \ddots & \\ & & & 0 \end{pmatrix} = \begin{pmatrix} A_1 & 0 \\ 0 & \mathbf{0} \end{pmatrix}$$

这里 A_1 是可逆矩阵. $AB = BA = 0$ 说明 B 形如

$$\begin{pmatrix} B_1 & 0 \\ 0 & B_2 \end{pmatrix}$$

(这个不用算. 因为 A_1 没有特征值0, 所以与 A 交换的矩阵必然也是准对角矩阵. 详细情况见1.4) 这里 $A_1B_1 = B_1A_1 = 0$. 由于 A_1 可逆, 所以 $B_1 = 0$. 这样就得到了证明.

问题 2: 设 A, B 为 n 阶方阵, 求证 $r(A+B) = r(A) + r(B)$ 的充要条件是存在可逆方阵 P, Q 使得

$$PAQ = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$$
$$PBQ = \begin{pmatrix} 0 & 0 \\ 0 & E_s \end{pmatrix}$$

其中 r, s 分别是 A, B 的秩, $r + s \leqslant n$.

证明: 只证必要性. 可以先选 P, Q 使得

$$PAQ = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$$

所以不妨假设 A 就是标准形, 下面证明可以用初等变换不改变 A 的形状, 把 B 变成想要的形状.
设

$$B = \begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix}$$

$$A+B=\begin{pmatrix} E_r+B_1 & B_2 \\ B_3 & B_4 \end{pmatrix}$$

$$s=r(B)\geqslant r\begin{pmatrix} B_2 \\ B_4 \end{pmatrix}\geqslant r(A+B)-r=s$$

第二个不等号是因为从 $A+B$ 中删去前 r 列秩最多减少 r.
所以

$$r\begin{pmatrix} B_2 \\ B_4 \end{pmatrix}=r(B)$$

从而有列变换

$$B=\begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix}\xrightarrow{\text{列的变换}}\begin{pmatrix} 0 & B_2 \\ 0 & B_4 \end{pmatrix}$$

同理

$$r\begin{pmatrix} 0 & 0 \\ B_3 & B_4 \end{pmatrix}=r(B)$$

所以有行变换

$$B=\begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix}\xrightarrow{\text{行的变换}}\begin{pmatrix} 0 & 0 \\ B_3 & B_4 \end{pmatrix}$$

注意这两个变换不改变 A 的形状. 由于行列变换可以互换, 所以

$$B=\begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix}\xrightarrow{\text{行, 列的变换}}\begin{pmatrix} 0 & 0 \\ 0 & B_4 \end{pmatrix}$$

B 可以继续化为标准形, 而且不改变 A 的形状.

$$B=\begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix}\to\begin{pmatrix} 0 & 0 \\ 0 & B_4 \end{pmatrix}\to\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & E_s \end{pmatrix}$$

这就完成了证明.

1.2 不能被有限个真子空间覆盖

设基域 F 有无穷多个元素(比如是数域), 那么 F 上的向量空间 V 不能被它的有限多个真子空间覆盖. 也就是不存在真子空间 $V_1, V_2, \cdots, V_m$ 使得 $V \subseteq V_1 \cup V_2 \cup \cdots \cup V_m$.
证明: 反证法: 设 $V \subseteq V_1 \cup V_2 \cup \cdots \cup V_m$ 是 V 的所有真子空间覆盖中长度最小的(任何 $m-1$ 个真子空间都不能覆盖 V). 那么对任何$1 \leqslant i \leqslant m$,

$$V_i \subsetneqq V_1 \cup V_2 \cup \cdots \cup V_{i-1} \cup V_{i+1} \cup \cdots \cup V_m.$$

否则与 m 是极小长度矛盾. 从而每个 V_i 都有一个自己的"独有向量". 取 V_1 的独有向量 α 和 V_2 的独有向量 β, 下面论证线性组合 $\lambda\alpha + \beta$ 中有无穷多个不属于 $V_1, V_2, \cdots, V_m$ 中的任何一个. 这是因为 $\lambda\alpha + \beta$ 总不属于 V_1, 其次当 $i \geqslant 2$ 的时候如果两个不同的 λ_1 和 λ_2 使得$\lambda_1\alpha + \beta, \lambda_2\alpha + \beta$落在同一个 V_i 中, 那么 $(\lambda_1\alpha + \beta) - (\lambda_2\alpha + \beta) = (\lambda_1 - \lambda_2)\alpha$ 也落在 V_i 中, 与 α 的定义矛盾! 所以每个 V_i 中至多有一个形如 $\lambda\alpha + \beta$的向量. 而基域有无穷多个元素, 从而集合$\{\lambda\alpha + \beta \mid \lambda \in F\}$中有无穷多个不同的向量, 所以必有无穷多个不能被 $V_1, \cdots, V_m$ 覆盖.

当 V 是有限维的情形, 有一个简洁的证明: 设 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 是 V 的一组基, 考察

$$\alpha_i = \varepsilon_1 + i\varepsilon_2 + \cdots + i^{n-1}\varepsilon_n$$

由范德蒙行列式的知识知道无穷向量序列 $\alpha_1, \alpha_2, \cdots, \alpha_n, \cdots$ 中任何 n 个都构成 V 的一组基, 而有限多个真子空间只能包含其中有限多个向量, 从而有无穷多个 α_i 不能被覆盖.

下面这个结论极其重要.
如果向量空间 V 上的线性变换 A 的特征多项式与极小多项式相等, 那么存在一个向量 $v \in V$ 使得 $v, Av, \cdots, A^{n-1}v$ 是 V 的一组基.
证明: 设 $f(x)$ 是 A 的特征多项式, $m(x)$ 是 A 的极小多项式. 根据熟知的结论, V 中任一元素 v 的极小化零多项式 $m_v(x)$ 整除 A 的极小多项式 $m(x)$. 但是 $m(x)$ 只有有限多个因子, 所以当 v 跑遍 V 时, $m_v(x)$ 只有有限多个不同的可能 $m_1(x), m_2(x), \cdots, m_k(x)$. 考虑$V_i = \{v \in V \bigm| m_i(A)v = 0\}$, 那么 V 属于 $V_1, \cdots, V_k$ 的并, 从而 $V_1, \cdots, V_k$ 不能都是真子空间, 所以必有某个 V_i 使得 $V = V_i$, 那么就有 $m(x) = m_i(x)$, 即 $f(x) = m_i(x)$. 也就是说存在 V 中的一个向量 v 使得 v 的极小化零多项式就是 A 的特征多项式. 那么 n 个向量 $v, Av, \cdots, A^{n-1}v$ 必然是线性无关的（想想为什么）, 从而它们构成 V 的一组基.
在本文档中这个 v 被称为"循环向量".

1.3 A, B 同时上三角化

设 A 和 B 是复数域上的两个 $n\times n$ 矩阵, 而且 $AB=BA$. 求证：
(a) A, B 可以同时相似于上三角形.
(b)如果 A, B 都可以对角化, 那么它们可以同时对角化.
证明: 当 $AB=BA$ 时, 有两个性质十分重要:
引理1: A 的特征子空间是 B 的不变子空间. 这是因为 $A(B\alpha)=BA\alpha=\lambda B\alpha$.
引理2: A, B 有共同的特征向量. 这是因为 B 限制在 A 的特征子空间上仍然是一个线性变换, 从而有特征向量.
(a)的证明: 对空间的维数 n 进行归纳: $n=1$ 时显然. 设命题对 $n-1$ 的情形成立, 考察 n 的情形: 由引理2, 这时 A, B有共同的特征向量 α: $A\alpha=\lambda\alpha, B\alpha=\mu\alpha$. 将 α 扩充为 V 的一组基. 在这组基下, A 和 B 分别形如

$$A=\begin{pmatrix}\lambda & c\\ 0 & A_1\end{pmatrix}\qquad B=\begin{pmatrix}\mu & d\\ 0 & B_1\end{pmatrix}$$

由 $AB=BA$ 可得 $A_1B_1=B_1A_1$. 由归纳假设, A_1, B_1 同时相似于上三角阵. 设 P 使得 $P^{-1}A_1P$, $P^{-1}B_1P$ 都是上三角阵, 不难验证

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

可以将 A, B 同时上三角化, 从而(a)得证.
(b)的证明:
先证明一个引理: 如果 A 有特征子空间直和分解

$$V=\bigoplus_{i=1}^{s}M_i\,,$$

这里 M_i 对应的特征值 λ_i 互不相同. 那么对任何 $A-$ 不变子空间 N 都有

$$N=\bigoplus_{i=1}^{r}N\cap M_i\,.$$

直和是显然的, $N\supseteq\oplus_{i=1}^{r}N\cap M_i$ 也是显然的, 只要证明“$\subseteq$”. 设$\alpha\in N$, 则存在 $v_i\in M_i$ 有

$$\begin{aligned}\alpha &= v_1+v_2+\cdots+v_r\\ A\alpha &= \lambda_1v_1+\lambda_2v_2+\cdots+\lambda_rv_r\\ &\cdots\\ A^{r-1}\alpha &= \lambda_1^{r-1}v_1+\lambda_2^{r-1}v_2+\cdots+\lambda_r^{r-1}v_r\end{aligned}$$

根据范德蒙行列式的性质, 我们可以用 α, $A\alpha$, $\cdots$, $A^{r-1}\alpha$ 反解出 v_1, v_2, $\cdots$, v_r 来. 然而由于 N 是 A 的不变子空间, 所以 α, $A\alpha$, $\cdots$, $A^{r-1}\alpha$ 都在 N 内, 从而 v_1, v_2, $\cdots$, v_r 也都

在 N 内. 这就证明了反向的包含关系, 从而证明了引理.
现在完成(b)的证明. 设 A 有特征子空间直和分解

$$V = \bigoplus_{i=1}^{r} M_i,$$

B 有特征子空间直和分解

$$V = \bigoplus_{j=1}^{s} N_j.$$

其中 M_i 是 A 的对应特征值 λ_i 的特征子空间, N_j 是 B 的对应特征值 μ_j 的特征子空间. 那么根据刚证明的引理,

$$V = \bigoplus_{i=1}^{r} \bigoplus_{j=1}^{s} M_i \cap N_j$$

由引理1, 每一个 $M_i \cap N_j$ 都是 A 和 B 共同的不变子空间, 从而也是共同的特征子空间, 所以上式给出了 V 同时关于 A 和 B 的特征子空间的直和分解. 在每个 $M_i \cap N_j$ 中取一组基, 合并起来得到 V 的一组基, 在这组基下 A 和 B 同时为对角形.

下面举个常见的例子:
A, B 是两个 n 阶方阵, $AB = BA$, 且 A 是幂零的, 求证 $|A + B| = |B|$.
证明: A, B 可以同时上三角化, 特别 A 的对角线上都是0, 从而结论是显然的.

$\mathrm{rank}(AB - BA) \leqslant 1$的情形

求证当 $\mathrm{rank}(AB - BA) \leqslant 1$时 A, B 仍然可以同时上三角化.
从前面的证明不难想到, 关键是证明 A, B 有共同的特征向量. 首先我们可以假设 A 不是可逆的, 否则的话就以 $A - \lambda I$ 代替 A. 这个假定的目的就是保证 $\mathrm{Ker}A$ 和 $\mathrm{Im}A$ 都是非平凡的 $A-$ 不变子空间. 我们要论证 $\mathrm{Ker}A$ 和 $\mathrm{Im}A$ 中至少有一个是 B 的不变子空间. 这样就可以降低空间的维数, 便于应用归纳法. 记齐次线性方程组 $(AB - BA)X = 0$ 的解空间为 X. 分两种情况讨论:
(1) $\mathrm{Ker}A \subset X$. 设$x \in \mathrm{Ker}A$, 立刻有 $ABx = BAx = 0$. 从而 $\mathrm{Ker}A$ 也是 B 的不变子空间.
(2) $\mathrm{Ker}A \subsetneqq X$. 从而存在向量 $x \in \mathrm{Ker}A$ 但是 $y = (AB - BA)x \neq 0$. 注意到 $\mathrm{rank}(AB - BA) = 1$, 所以 $\mathrm{Im}(AB - BA) = \{y\}$. 但是

$$y = (AB - BA)x = ABx - BAx \in \mathrm{Im}A$$

这就说明 $\mathrm{Im}(AB - BA) \subset \mathrm{Im}A$. 那么对任何 $u = Av \in \mathrm{Im}A$, 我们有

$$BAv = ABv - (AB - BA)v \in \mathrm{Im}A$$

从而 $\mathrm{Im}A$ 确实是 B 的不变子空间.
剩下的就是复制(a)的证明: 首先用归纳法不难证明 $\mathrm{rank}(AB - BA) \leqslant 1$时 A, B 总是有共

同的特征向量. 这是因为 A, B 在共同的不变子空间（ $\mathrm{Ker}A$ 或者 $\mathrm{Im}A$ 之一）上的限制仍然满足 $\mathrm{rank}(AB-BA) \leqslant 1$, 从而使用归纳假设即可. 其次对空间的维数 n 归纳, 假设维数小于 n 的时候结论成立, 论证维数等于 n 的情形: 将 A, B 共同的特征向量开拓为空间的一组基, 在这组基下有

$$A = \begin{pmatrix} \lambda & c \\ 0 & A_1 \end{pmatrix} \qquad B = \begin{pmatrix} \mu & d \\ 0 & B_1 \end{pmatrix}$$

只要证 A_1, B_1 同时相似于上三角阵. 利用 $\mathrm{rank}(AB-BA) \leqslant 1$有

$$\mathrm{rank}[\begin{pmatrix} \lambda & c \\ 0 & A_1 \end{pmatrix}\begin{pmatrix} \mu & d \\ 0 & B_1 \end{pmatrix} - \begin{pmatrix} \mu & d \\ 0 & B_1 \end{pmatrix}\begin{pmatrix} \lambda & c \\ 0 & A_1 \end{pmatrix}] \leqslant 1$$

从而 $\mathrm{rank}(A_1B_1 - B_1A_1) \leqslant 1$. 由归纳假设, A_1, B_1 可以同时被矩阵 P 同时上三角化, 那么

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

可以将 A, B 同时上三角化, 从而(c)得证.

这一节我几乎进行了重写. 上一版中犯了两个基本的错误: 设 N 是 V 的子空间, 那么由

$$V = V_1 \bigoplus V_2$$

是推不出

$$N = (N \cap V_1) \bigoplus (N \cap V_2)$$

来的, 所以那个引理是很必要的. 很多书上对 (b) 采用的是矩阵的证法, 但是都漏掉了说明"当准对角矩阵 A 可以对角化的时候, A 的对角线上的矩阵也都可以对角化"这一点. 其次 $\mathrm{rank}(AB-BA) \leqslant 1$ 的情形的证明我是采用了教科书的证明, 以前我自己想的那一个被发现是错误的, 因为三个子空间 V_1, V_2, V_3 两两是直和的话不能推出 $V_1 + V_2 + V_3$ 是直和, 所以只好忍痛割爱. 仔细检查了上一版整个文档以后发现习题中还有一道也是犯了这样的错误, 在这里特别指出来, 希望大家对上面两个错误特别注意.

1.4 矩阵方程 $AX = XB$ 的讨论

设 A, B 分别是数域 F 上的 n 阶和 m 阶方阵, X 是 $n \times m$ 矩阵. 那么所有满足矩阵方程 $AX = XB$ 的 X 构成一个 F 上的向量空间. 我们有如下的结论:
(a)这个矩阵方程的有解与否和解空间的维数在数域的扩张下保持不变.
(b)矩阵方程有非零解当且仅当 A, B 有共同的特征值.
(c) 矩阵方程解空间的维数是 $\sum_{i,j} \deg(f_i(x), g_j(x))$. 这里 $f_i(x)$ 跑遍 A 的不变因子, $g_j(x)$ 跑遍 B 的不变因子, $(f_i(x), g_j(x))$ 表示 $f_i(x)$ 和 $g_j(x)$ 的最大公因式.

(a) 的证明: 解矩阵方程最直接的办法就是全部展开然后解线性方程组. 即把 $X = (x_{ij})$ 等同于一个 mn 维的向量

$$(x_{11}, x_{12}, \cdots, x_{1m}, \cdots, x_{n1}, \cdots, x_{nm}),$$

把 $AX = XB$ 看成一个关于 mn 个变量 x_{ij} 的齐次线性方程组. 不难验证

$$X \to (x_{11}, x_{12}, \cdots, x_{1m}, \cdots, x_{n1}, \cdots, x_{nm})$$

是矩阵方程的解空间 X 到齐次线性方程组的解空间 X' 的一个一一的线性的对应, 也就是向量空间 X 和 X' 同构. 然而解线性方程组只用到加减乘除运算, 在数域 F 内就能进行, 所以当数域 $K \supseteq F$ 并把 A, B 看成 K 上的矩阵的话, 矩阵方程 $AX = XB$ 的解空间的维数并不会改变.
(b) 的证明: $\Rightarrow$: 设 A, B 无共同的特征值, 且有 $n \times m$ 阶矩阵 X 满足 $AX = XB$, 那么 $A^2X = AXB = XB^2$, 用归纳法可得对任何正整数 k 都有 $A^kX = XB^k$. 从而对任何多项式 $f(x)$ 都有 $f(A)X = Xf(B)$. 特别令 $f(x)$ 是 B 的特征多项式, 由 Hamilton-Cayley 定理, $f(B) = 0$. 但是 $f(A)$ 是可逆矩阵, 必有 $X = 0$! （想一想, 为什么?）
$\Leftarrow$: 见(c)的具体计算.
(c) 的证明: 下面来计算解空间 X 的维数. 根据(a)的结论, 完全可以假设 $F = \mathbb{C}$, 这样并不改变解空间的维数. 设 P 是 n 阶的可逆矩阵, Q 是 m 阶的可逆矩阵, 在等式 $AX = XB$ 两边作变换

$$(P^{-1}AP)(P^{-1}XQ) = (P^{-1}XQ)(Q^{-1}BQ)$$

我们看到 $X \to P^{-1}XQ$ 给出了 X 到自身的一个可逆线性变换, 它也不改变 X 的维数. 所以我们可以在 A, B 都是 Jordan 标准型的前提下考虑问题. 设 A, B 的标准型中的全部 Jordan 块分别是(k_i, m_j 代表阶数)

$$J_{k_1}(\lambda_1), J_{k_2}(\lambda_2), \cdots, J_{k_s}(\lambda_s)$$

$$L_{m_1}(\mu_1), L_{m_2}(\mu_2), \cdots, L_{m_t}(\mu_t)$$

相应地把将 X 分成 st 个子块:

$$X = \begin{pmatrix} X_{11} & X_{12} & \cdots & X_{1t} \\ X_{21} & X_{22} & \cdots & X_{2t} \\ \vdots & \vdots & \ddots & \vdots \\ X_{s1} & X_{s2} & \cdots & X_{st} \end{pmatrix}$$

其中 X_{ij} 是 $k_i \times m_j$ 阶矩阵. 于是从 $AX = XB$ 得到 st 个方程

$$J_iX_{ij} = X_{ij}L_j \quad 1 \leqslant i \leqslant s,\ 1 \leqslant j \leqslant t$$

当 J_i 对应的特征值 λ_i 与 L_j 对应的特征值 μ_j 不相等的时候, 根据前面的证明, $X_{ij} = 0$. 而当 $\lambda_i = \mu_j$ 时,

$$N_iX_{ij} = X_{ij}N_j$$

这里N_i形如

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

那么直接展开就可以得到 X_{ij} 形如上三角分层矩阵:

$$\begin{pmatrix} 0 & \cdots & 0 & b_p & b_{p-1} & \cdots & b_2 & b_1 \\ & & & & b_p & b_{p-1} & \cdots & b_2 \\ \vdots & \vdots & \vdots & & \ddots & \ddots & & \vdots \\ & & & & & & b_p & b_{p-1} \\ 0 & \cdots & 0 & \cdots & \cdots & \cdots & 0 & b_P \end{pmatrix}$$

上面给出的是行大于列时的样子, 列等于行的时候的样子是上三角阵

$$\begin{pmatrix} b_p & b_{p-1} & \cdots & b_2 & b_1 \\ & b_p & b_{p-1} & \cdots & b_2 \\ & & \ddots & \ddots & \vdots \\ & & & b_p & b_{p-1} \\ & & & & b_p \end{pmatrix}$$

列小于行的时候类似. 可以看到, 当 $\lambda_i = \mu_j$ 时X_{ij}含有

$$\min\{X_{ij}\text{的行数},X_{ij}\text{的列数}\} = \min\{k_i, m_j\}$$

个参数. 显然它还等于 J_i, L_j 对应的初等因子的最大公因式的次数. 用初等因子来看的好处是把两种情况统一起来了, 如果 $\lambda_i \neq \mu_j$, 那么 J_i 和 L_j 对应的初等因子互素, 从而最大

公因式的次数是零, 也和 $X_{ij} = 0$ 吻合. 由于整个 X 含有的参数的个数是全部 X_{ij} 的参数个数的和, 所以我们只要两两求出 A 和 B 的初等因子的最大公因式, 然后把它们的次数加起来, 就得到了 X 的维数. 进一步, 两两求初等因子的最大公因式的次数再相加, 和两两求出不变因子的最大公因式的次数再相加, 结果是一样的. 所以我们得到 X 的维数就是

$$\sum_{i=1}^{p}\sum_{j=1}^{q}\deg(f_i(x), g_j(x))$$

其中 $f_1(x), \cdots, f_p(x)$是 A 的全部不变因子, $g_1(x), \cdots, g_q(x)$是 B 的全部不变因子.

在这里我们用不变因子得出了解空间的确切维数. 由于不变因子是不依赖于域的概念, 所以我们就验证了(1)中的结论.

可以看到, A 和 B “相似”程度越高, 矩阵方程解空间的维数就越大, 可以说解空间的维数反映了 A 和 B 之间的相似程度. 但是解空间的维数是不随域变化而变化的, 那么这是不是暗示着 A 和 B 之间的相似程度也不随着域的变化而变化呢? 这就是下面要讨论的.

矩阵的相似性与域无关

设 $F \subset K$ 是两个数域, 如果 F 上的两个方阵 A 和 B 在 K 上相似, 则它们在 F 上相似.

我们已经知道矩阵方程 $AX = XB$ 解空间的维数是不随域的扩大而改变的, 而现在的问题是证明如果 $AX = XB$ 在 K 上有一个可逆解的话, 那么它在 F 上也有一个可逆解. 首先把 $AX = XB$ 看成数域 F 上的矩阵方程, 并且设 $X_1, X_2, \cdots, X_s$ 是解空间的一组基, 那么 $X_1, X_2, \cdots, X_s$ 也构成数域 K 上的矩阵方程 $AX = XB$ 的解空间的一组基. 所以 $AX = XB$ 在 K 上有一个可逆解说明存在 $X_1, X_2, \cdots, X_s$ 的 $K-$ 线性组合

$$a_1X_1 + a_2X_2 + \cdots + a_sX_s, \qquad a_i \in K$$

使得 $|a_1X_1 + a_2X_2 + \cdots + a_sX_s| \neq 0$. 我们要证明一定有 $X_1, X_2, \cdots, X_s$ 的 $F-$ 线性组合

$$b_1X_1 + b_2X_2 + \cdots + b_sX_s, \qquad b_i \in F$$

使得 $|b_1X_1 + b_2X_2 + \cdots + b_sX_s| \neq 0$. 自然我们要考察 s 个变元 $t_1, t_2, \cdots, t_s$ 的多元多项式

$$f(t_1, t_2, \cdots, t_s) = |t_1X_1 + t_2X_2 + \cdots + t_sX_s|$$

f 是一个数域 F 上的多元多项式（当然也是 K 上的）, $f(a_1, a_2, \cdots, a_s) \neq 0$ 说明 f 不是零多项式. 由于数域有无穷多个元素而 f 只能有有限多个零点, 所以必定存在 F 中的 $b_1, b_2, \cdots, b_s$ 使得 $f(b_1, b_2, \cdots, b_s) \neq 0$. 这就说明了 A, B 在 F 上也是相似的.

1.5 把与 A 交换的矩阵表示为 A 的多项式

何时任何与矩阵 A 交换的矩阵总能表示为A的多项式? 下面将从变换和矩阵两个角度入手, 对这个问题进行一下深入的探讨, 力求把想法写清楚. 总共分为三个部分, 第一部分是变换的角度, 第二部分是矩阵的角度, 第三部分把二者统一起来.

变换的角度

如果任何与 A 交换的变换 B 总能表示为 A 的多项式, 那么 A 应该有怎样的性质? 首先由于 A 的不变子空间也是 $f(A)$ 的不变子空间, 所以我们看到如果 V 有 $A-$ 不变子空间的直和分解 $V = V_1 \oplus V_2 \oplus \cdots \oplus V_r$ 的话, 那么它一定也是 V 的 $B-$ 不变子空间直和分解. 换句话说, 任何与 A 交换的变换 B 不能把 V_i 中的向量映到别的 V_j 中去. 从而我们得到第一个有用的信息：

设 $V = V_1 \oplus V_2 \oplus \cdots \oplus V_r$ 是 V 的 $A-$ 不变子空间直和分解, 那么当 $i \neq j$ 时不存在非零变换 $C : V_i \to V_j$ 使得 $A_jC = CA_i$. 这里用 A_i 表示 A 在 V_i 上的限制.

若不然, 有这样的 C 存在, 那么当 k 不等于 i 时定义 B 在所有的 V_k 上都是0, 在 V_i 上就等于 C, 则不难验证 $AB = BA$, 但是 V_i 不是 B 的不变子空间, B 是不能表示为 A 的多项式的. 这是一个重要的发现, 这让我们想起1.4讨论过的:

设 A_1, A_2 分别是复数域上有限维向量空间 V_1 和 V_2 上的线性变换, 那么存在非零的线性映射 $B:\ V_1 \to V_2$ 使得 $BA_1 = A_2B$ 的充要条件是 A_1 和 A_2 有共同的特征值.

所以我们得到如下重要结论:
当 A 有不变子空间直和分解时, A 在任何两个不变子空间上的限制不能有共同的特征值.
进一步我们的推理, 由于 V 总是可以分解为关于 A 的循环不变子空间的直和, 在每个子空间上 A 呈 Jordan 形, 我们立刻得到
A 的 Jordan 标准形中属于任何特征值 λ 的 Jordan 块只能有一块.
这个话的等价说法是
A 的极小多项式与特征多项式相同.
还有另一种说法是
存在 V 中的向量 v 上使得 $v,\ Av,\ \cdots,\ A^{n-1}v$ 是 V 的一组基.
现在我们已经找到了 A 应该满足的必要条件: A 有一个循环向量. 那么这个条件是不是充分的呢? 答案是肯定的:
证明: 设 V 在 A 的作用下成为一个循环空间, v 是一个生成元, 我们断言循环空间上一个与 A 交换的变换由它在生成元 v 处的值完全决定: 由于 V 中的任何向量 m 可以表示为$\{v,Av,\cdots,A^{n-1}v\}$的线性组合, 也就是存在多项式 $f(x)$ 使得 $m = f(A)v$. 特别有 $Bv = f(A)v$. 不难验证就有 $B = f(A)$ 成立. 从而我们已经从变换的角度完全解决了这一

问题. 把它写成
命题1: 任何与 A 交换的矩阵都能写成 A 的多项式的充要条件是 V 在 A 的作用下有一个循环向量.

矩阵的角度

我们已经在1.4中计算了矩阵方程 $AX = XB$ 的解空间的维数, 特别当 $A = B$ 时, 就是与 A 交换的矩阵构成的向量空间的维数. 如果把 A 的不变因子按照次数高低从大到小排成 $f_1(x), f_2(x), \cdots, f_p(x)$, 次数分别是 $n_1, n_2, \cdots, n_p$, 那么这个维数等于

$$\sum_{i=1}^{p}\sum_{j=1}^{p}\deg(f_i(x), f_j(x)) = n_1 + n_2 + \cdots + (2p-1)n_p$$

从而我们得到了
命题2: 与 A 交换的矩阵构成的向量空间的维数是

$$n_1 + 3n_2 + \cdots + (2p-1)n_p$$

其中 n_i 是 A 的不变因子的次数, 按照从大到小的顺序排列.
注意到总是有

$$n_1 + 3n_2 + \cdots + (2p-1)n_p \geqslant n_1 + n_2 + \cdots + n_p = n$$

等号成立当且仅当 $p = 1$, 即 A 仅有一个次数大于0的不变因子, 也就是 A 的极小多项式等于它的特征多项式. 从而从矩阵的角度再次证明了当 A 有循环向量时与 A 交换的矩阵总能表示为 A 的多项式.

一个类似的问题

问题: 方阵 C 和每一个与 A 可交换的方阵都可交换, 求证 C 可以表示为 A 的多项式.
证明: 注意这里的 A 没有什么限制, 可能是任意的矩阵. 如果问题要证的结论成立的话, A 的不变子空间必然也是 C 的不变子空间. 即
引理2: A 的不变子空间直和分解必定也是 C 的不变子空间直和分解.
证明: 设 V 在 A 的作用下分解为不变子空间的直和

$$V = V_1 \oplus V_2 \oplus \cdots \oplus V_r$$

设 P 是从 V 到 V_1 上的投影, 则不难验证有 $AP = PA$, 那么根据已知有 $CP = PC$. 所以 V_1 也是 C 的不变子空间. 类似地所有 V_i 都是 C 的不变子空间. 引理得证.
现在立刻想到用上前面的结论, 根据 Frobenius 有理标准型的理论, V 可以分解为循环不变子空间的直和

$$V = \{v_1\} \oplus \{v_2\} \oplus \cdots \oplus \{v_r\},$$

而 C 在每个 $\{v_i\}$ 上的限制 C_i 与 A 在 $\{v_i\}$ 上的限制 A_i 交换, 所以可以表示为 A_i 的多项式 $C_i = g_i(A_i)$. 问题是这些 $g_i(x)$ 相等吗? 如果是同一个多项式 $g(x)$ 那问题就OK了, 在每一个不变子空间上都有 $C = g(A)$, 当然全空间也是. 但是这个不那么显然. 所以得分析的再细一点. 约定每个 $\{v_i\}$ 对应的极小多项式 $f_i(x)$ 是 A 的不变因子, 次数从大到小排列. 我们有

引理3: 存在与 A 交换的变换 B 使得 $Bv_1 = v_2$.

证明: 这个比较直观. 因为 $\{v_1\}$ 和 $\{v_2\}$ 是两个循环子空间, $\{v_2\}$ 的不变因子整除 $\{v_1\}$ 的不变因子, 所以 $\{v_1\}$ 有一个子空间和 $\{v_2\}$ 同构. 这就类似于两个循环群, 如果其中一个的阶数整除另一个, 那么它就同构于另一个的一个子群.

设 $f_1(x) = f_2(x)h(x)$, 考察 $\{v_1\}$ 的关于 A 的不变子空间 $\{h(A)v_1\}$, 它对应的极小多项式是 $f_2(x)$, 所以它和 $\{v_2\}$ 对应的极小多项式相同, 定义

$$\varphi : \{h(A)v_1\} \to \{v_2\}$$

$$h(A)v_1 \to v_2$$

不难验证这是一个双射, 而且与 A 交换. 下面考虑映射

$$v_1 \xrightarrow{h(A)} h(A)v_1 \xrightarrow{\varphi} v_2$$

两个与 A 交换的变换的复合, 还与 A 交换. 这就把 v_1 变成了 v_2.

注: B 在 $V_2, \cdots, V_r$ 上可以定义为恒等变换, 零变换等等, 我们要的是局部的“不平凡的”与 A 交换的变换.

下面完成最后的证明: 根据前面的结论, 存在一个从 $\{v_1\}$ 到 $\{v_2\}$ 的线性变换 B 使得 $Bv_1 = v_2$ 而且 B 与 A 交换:

$$CB = BC \Rightarrow CBv_1 = BCv_1 \Rightarrow Cv_2 = Bg_1(A)v_1 \Rightarrow g_2(A)v_2 = g_1(A)v_2$$

再引用一遍那句话: 循环空间上一个与 A 交换的变换由它在生成元 v 处的值完全决定. 所以 $g_1(A)$ 和 $g_2(A)$ 相等, 那么 $g_1(x)$ 和 $g_2(x)$ 可以取成相同的多项式. 类似地, 所有的 $g_i(x)$ 全部可以取成同样的多项式.

命题3: 如果变换 C 与任何与 A 交换的变换交换, 那么 C 可以表示为 A 的多项式.

进一步的解释

设 A, B 分别是复数域上 n 维向量空间 U 和 m 维向量空间 V 上的线性变换, 称线性变换 $f : U \to V$ 是一个模同态, 如果有

$$f(Au) = Bf(u) \quad u \in U$$

成立. 即 f 保持 A 和 B 的作用. 或者说下图交换:

$$\begin{tikzcd} u \arrow[r] \arrow[d, "f"] & Au \arrow[d, "f"] \\ v \arrow[r] & Bv \end{tikzcd}$$

换成矩阵的说法就是有 $XA = BX$.
一个最熟悉的例子就是诱导变换

$$\begin{tikzcd} v \arrow[r] \arrow[d, "\pi"] & Av \arrow[d, "\pi"] \\ \bar{v} \arrow[r] & \bar{A}v \end{tikzcd}$$

模同态同样有同态基本定理

$$U/Kerf \cong Imf$$

注意这里的同构不再是简单的向量空间之间的同构, 它保持 A 在 $U/\mathrm{Ker}f$ 上的诱导变换与 B 在 $\mathrm{Im}f$ 上的限制之间的作用.
同态意味着局部的相似. 所以现在可以解释为什么当 A, B 没有共同的特征值时矩阵方程 $AX = XB$ 仅有零解了, 因为有非零解说明 B 在某个非零的商空间上的诱导变换与 A 在某个非零子空间上的限制是相似的, 而相似的变换有相同的特征值, 从而 B 和 A 有共同的特征值.

最后用模的语言把结果表述一下: 向量空间 V 可以看作是一个有限维的 $\mathbb{C}[A]-$ 模, 记 $R = \mathbb{C}[A]$, 那么 R 是 V 上的全体线性变换 $\mathrm{Hom}_k(V, V)$ 的一个子环, 所以可以考虑 R 在 $\mathrm{Hom}_k(V, V)$ 内的中心化子

$$D = \mathrm{Hom}_R(V, V) = \{B \in \mathrm{Hom}_k(V, V) \mid AB = BA\}$$

以及 D 在 $\mathrm{Hom}_k(V, V)$ 中的中心化子

$$E = \mathrm{Hom}_D(V, V) = \{C \in \mathrm{Hom}_k(V, V) \mid CB = BC, \forall B \in D\}$$

对于交换环 $R = \mathbb{C}[A]$, 不难看出有 $R \in E \in D$. 命题1和2说的就是 $R = D$ 当且仅当 A有一个循环向量, 命题3说的是 $R = E$ 无条件成立.(和具体的 A 没关系) 这个叫做双中心化子性质: 对 R 连求两次中心化子以后又得到了 R.

1.6 A 的多项式的 Jordan 标准形

已知 A 的标准形和多项式 $f(x)$, 怎样求 $f(A)$ 的标准形? 显然这个问题可以化为对 Jordan 块 J, 求 $f(J)$ 的标准形. 首先来介绍 $f(J)$ 的计算技巧. 设

$$J=\begin{pmatrix}\lambda_0 & 1 & & \\ & \lambda_0 & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_0\end{pmatrix}_{r\times r}$$

对于给定的多项式 $f(x)$, 作 Taylor 展开 $f(x)=a_0+a_1(x-\lambda_0)+\cdots+a_n(x-\lambda_0)^n$, 那么

$$f(J)=\begin{pmatrix}a_0 & a_1 & \cdots & a_{r-1} \\ 0 & a_0 & \ddots & \vdots \\ \vdots & & \ddots & a_1 \\ 0 & 0 & \cdots & a_0\end{pmatrix}$$

可见 Jordan 块的多项式是很好算的. 反过来, 有一个常见的习题是证明当矩阵 B 与 J 交换的时候 B 是 J 的多项式. 我们已经知道这样的 B 必定是上三角分层矩阵, 所以 $f(x)$ 直接就可以“还原”出来.
继续, $f(J)$ 的 $\lambda-$ 矩阵

$$\lambda I-f(J)=\begin{pmatrix}\lambda-a_0 & -a_1 & \cdots & -a_{r-1} \\ 0 & \lambda-a_0 & \ddots & \vdots \\ \vdots & & \ddots & -a_1 \\ 0* & 0 & \cdots & \lambda-a_0\end{pmatrix}$$

显然它的特征多项式是 $D_r\lambda=(\lambda-a_0)^r$, 所以它的初等因子都形如 $(\lambda-a_0)^k, 0\leqslant k\leqslant r$. 所以其 Jordan 块都是些小的以 a_0 为特征值的子块.
(1) $a_1\neq 0$. 考察加了 $*$ 标记的0的代数余子式, 展开以后的 $(r-1)!$ 项中除了对角线乘积 $(-a_1)^{r-1}$ 以外都含有因式 $\lambda-a_0$, 所以在 $a_1\neq 0$ 的假设下这个余子式与 $\lambda-a_0$ 互素, 这说明 $D_{r-1}=0$, 所以 $f(J)$ 有唯一的初等因子 $D_r\lambda=(\lambda-a_0)^r$. 这说明 $a_1\neq 0$ 的时候 $f(J)$ 与 J 有同样的标准形, 仅仅把对角线上的元素换成 $f(\lambda_0)$.
(2) $a_1=\cdots=a_{k-1}=0, a_k\neq 0$. 记

$$B=f(J)-a_0I,\quad r_i=\operatorname{rank}B^i$$

注意由于 B^i 都是严格的上三角分层矩阵, 它的秩由最靠近主对角线的非零准对角线决定, 与更高的准对角线上的元素无关. 作带余除法

$$r=qk+p,\quad q\geqslant 1, 0\leqslant p<k$$

回忆一下 N 每乘一次幂秩都减少1, 那么 N^k 每乘一次幂秩都减少 k. 而 B^i 与 $(N^k)^i$ 秩相同. 所以 $r_0 = r, r_1 = r - k, r_2 = r - 2k, \cdots, r_q = r - qk$, 再往后就都是0了. 直观上看, 随着乘幂, B 的元素“k 步 k 步”地向右上角跳跃, 每跳一次秩减少 k, 剩下不足 k 步的时候“一跃为0”. 从而根据计算公式, $f(J)$ 的阶为 i 的 Jordan 块的个数是 $r_{i+1} + r_{i-1} - 2r_i$, 知道 $f(J)$ 的阶为 i 的 Jordan 块的个数是

阶数	0	1	$\cdots$	$q-1$	q	$q+1$	$q+2$	$\cdots$
个数	0	0	$\cdots$	0	$k-p$	p	0	$\cdots$

这结果很有意思, $f(J)$ 分解为阶数相邻或者相同的一些 Jordan 块. 而且 λ_0 是 $f(x) - f(\lambda_0)$ 的几阶零点 $f(J)$ 就会分解为几块. 下面是两个值得记住的推论:
(1)可逆矩阵 A 的任意次幂 A^k 与 A 有完全相同的 Jordan 形, 仅仅把对角元换成 λ_i^k.
(2)对于不可逆矩阵, 如果特征值0对应的 Jordan 块都是一阶的, 那么它的平方与 A 有完全相同的 Jordan 形, 仅仅把对角元换成 λ_i^2; 如果特征值0有阶数大于等于2的 Jordan 块, 那么如果这个块是 $2q$ 阶的, 就分解为2个 q 阶的子块, 如果它是 $2q+1$ 阶的, 就分解为一个 q 阶和一个 $q+1$ 阶的子块.

1.7 几种矩阵的讨论

1.7.1 秩1方阵

命题1: n 阶矩阵 A 的秩是1的充要条件是存在 $n\times 1$ 的非零向量 X,Y 使得 $A=XY'$.
证明: 若 $\mathrm{rank}(A)=1$, 任取 A 的一个非零的列向量 X, 则其他列向量都是 X 的数乘, 即 A 形如

$$A=(c_1X,c_2X,\cdots,X,\cdots,c_nX)$$

记 $Y'=(c_1,c_2,\cdots,1,\cdots,c_n)$ 则 Y 非零且 $A=XY'$. 充分性是显然的.
命题2: 秩1方阵的特征多项式是 $\lambda^{n-1}(\lambda-<X,Y>)$.
这是因为 XY' 与 $Y'X=<X,Y>$ 的特征多项式只差因子 λ^{n-1}.
注: 这里利用了一个应当熟知的结论: 设 A 是 $n\times m$ 矩阵, B 是 $m\times n$ 矩阵, 那么 AB 和 BA 的特征多项式仅相差因子 λ^{m-n}（假设 m 较大的话）.
命题3: A 相似于对角形当且仅当 $<X,Y>\neq 0$.
证明: 如果 $<X,Y>=0$, 那么 A 的特征值全是0, 但是由于 $\mathrm{rank}A=1$, 所以特征值0的几何重数是 $n-1$, 小于代数重数, 不可对角化. 如果 $<X,Y>\neq 0$, 那么特征值0的几何重数是 $n-1$, 特征值 $<X,Y>$ 的几何重数至少是1, 所以可以对角化.
命题4: n 阶矩阵 A 的伴随矩阵 A^* 可以表示为 A 的多项式.
讨论: (1) $\mathrm{rank}(A)<n-1$, 这时 $A^*=0$, 结论显然成立.
(2) $\mathrm{rank}(A)=n$, 这时由 Hamilton-Cayley 定理,

$$A_n+a_{n-1}A^{n-1}+\cdots+a_1A+a_0=0$$

由于 A 可逆, 所以 $a_0=(-1)^n|A|\neq 0$. 那么

$$A(A^{n-1}+a_{n-1}A^{n-2}+\cdots+a_1)=-a_0,$$

$$A^{-1}=-\frac{1}{a_0}(A^{n-1}+a_{n-1}A^{n-2}+\cdots+a_1)$$

即 A^{-1} 可表为 A 的多项式, 从而 A^* 也能.
(3) $\mathrm{rank}(A)=n-1$. 这时 $r(A^*)=1$. 分三步走:
第一步: 论证向量空间 $M=\{B\in M_n(F)\,|\,AB=BA=0\}$ 是一维的. 首先不难看出 M 中的矩阵都是秩1方阵. 设 α 是方程组 $AX=0$ 解空间的基, β' 是方程组 $YA=0$ 的解空间的基, 那么 M 中的任何矩阵 B 都可以写为 $B=c\alpha\beta'$, 从而 M 是一维的.
第二步: 找一个非零矩阵, 它是 A 的多项式且在 M 中. 设 A 的极小多项式为 $m(x)=xg(x)$, 由 $Ag(A)=g(A)A=0$ 知道 $g(A)\in M$. 但 $g(A)$ 不能是0, 否则与 $m(x)$ 是极小多项式矛盾.
第三步: 现在 $g(A)$ 和 A^* 都是矩阵方程 $AX=XA=0$ 的非零解, 而且这个解空间还是一维的, 所以存在一个常数 λ 使得 $A^*=\lambda g(A)$, 这就完成了证明.

1.7.2 幂等矩阵

设 $A_1,\cdots,A_m$ 为 n 阶方阵, 适合条件$A_1+A_2+\cdots+A_m=I_n$. 试证明下面三个条件等价:
(1) $A_1,\cdots,A_m$ 都是幂等矩阵;
(2) $\mathrm{rank}(A_1)$+$\mathrm{rank}(A_2)$+$\cdots$+$\mathrm{rank}(A_m)=n$;
(3)当 $1\leqslant i,j\leqslant n$ 且 $i\neq j$ 时有 $A_iA_j=0$.
证明: (1) $\Rightarrow$ (2): 用熟知的结论, 幂等矩阵的秩等于它的迹, 可得

$$\sum_{i=1}^{m}\mathrm{rank}(A_i)=\sum_{i=1}^{m}\mathrm{tr}(A_i)=\mathrm{tr}(I_n)=n.$$

(2) $\Rightarrow$ (3): 考虑大的 $n^2\times n^2$ 的准对角方阵

$$\begin{pmatrix} A_1 & & & & \\ & A_2 & & & \\ & & A_3 & & \\ & & & \ddots & \\ & & & & A_m \end{pmatrix} \xrightarrow{\text{都加到第一行}} \begin{pmatrix} A_1 & A_2 & A_3 & \cdots & A_m \\ & A_2 & & & \\ & & A_2 & & \\ & & & \ddots & \\ & & & & A_m \end{pmatrix}$$

$$\xrightarrow{\text{都加到第一列}} \begin{pmatrix} I_n & A_2 & A_3 & \cdots & A_m \\ A_2 & A_2 & & & \\ A_3 & & A_3 & & \\ \vdots & & & \ddots & \\ A_m & & & & A_m \end{pmatrix}$$

$$\xrightarrow{\text{消去第一行中的 } A_i} \begin{pmatrix} I_n & 0 & 0 & \cdots & 0 \\ A_2 & A_2-A_2^2 & -A_2A_3 & \cdots & -A_2A_m \\ A_3 & -A_3A_2 & A_3-A_3^2 & \cdots & -A_3A_m \\ \vdots & & & \ddots & \vdots \\ A_m & -A_mA_2 & -A_mA_3 & \cdots & A_m-A_m^2 \end{pmatrix}$$

$$\xrightarrow{\text{消去第一列中的 } A_i} \begin{pmatrix} I_n & 0 & 0 & \cdots & 0 \\ 0 & A_2-A_2^2 & -A_2A_3 & \cdots & -A_2A_m \\ 0 & -A_3A_2 & A_3-A_3^2 & \cdots & -A_3A_m \\ \vdots & & & \ddots & \vdots \\ 0 & -A_mA_2 & -A_mA_3 & \cdots & A_m-A_m^2 \end{pmatrix}$$

所以 $\mathrm{rank}(A_1)+\mathrm{rank}(A_2)+\cdots+\mathrm{rank}(A_m)=n$ 当且仅当右下角的大的矩阵为0, 即 A_i 都是幂等矩阵而且互相正交.

(3) $\Rightarrow$ (1): 在已知等式两边乘以 A_1 :

$$A_1^2+A_1(A_2+\cdots+A_m)=A_1$$

即 $A_1^2=A_1, A_1$ 是幂等的. 对其它的 A_i 也完全类似. 这就完成了全部的证明.

1.7.3 幂零矩阵

矩阵幂零等价于所有的特征根都是0, 这个比较显然. 要验证这一点, 只要证明对所有的 $1 \leqslant k \leqslant n$ 都有 $\mathrm{tr}A^k = 0$. 为什么呢? 原因是迹函数 $\mathrm{tr}A, \mathrm{tr}A^2, \cdots, \mathrm{tr}A^n$ 完全决定了 A 的特征多项式. 设 σ_k 是 $\lambda_1, \cdots, \lambda_n$ 的 k 次初等对称多项式, 那么

$$\mathrm{tr}A^k = \lambda_1^k + \lambda_2^k + \cdots + \lambda_n^k = s_k$$

根据 Newton 恒等式有

$$\begin{aligned} \sigma_1 &= s_1 \\ s_1\sigma_1 - 2\sigma_2 &= s_2 \\ s_2\sigma_1 - s_1\sigma_2 + 3\sigma_3 &= s_3 \\ \cdots\cdots \\ s_k\sigma_1 - s_{k-1}\sigma_2 + \cdots + (-1)^{k+1}k\sigma_k &= s_k \end{aligned}$$

用 Cramer 法则依次解出 $\sigma_1, \cdots, \sigma_n$ 来:

$$\sigma_k = \frac{1}{k!}\begin{vmatrix} s_1 & 1 & 0 & 0 & \cdots & 0 \\ s_2 & s_1 & 1 & 0 & \cdots & 0 \\ s_3 & s_2 & s_1 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ s_{k-1} & s_{k-2} & \cdots & \cdots & \ddots & 1 \\ s_k & s_{k-1} & \cdots & \cdots & \cdots & s_1 \end{vmatrix}$$

从而 $s_1, s_2, \cdots, s_n$ 完全决定了 $\sigma_1, \cdots, \sigma_n$, 也就完全决定了 A 的特征多项式.
一个常见的例子是
设 $[A, B] = AB - BA$, 如果有方阵 $A_1, A_2, \cdots, A_m$ 和 $B_1, B_2, \cdots, B_m$ 使得

$$C = [A_1, B_1] + [A_2, B_2] + \cdots + [A_m, B_m]$$

而且 C 与 $A_1, A_2, \cdots, A_m$ 都交换, 那么 C 是幂零的.
证明: 用上面的方法, 归纳证明对任何正整数 k 都有 $\mathrm{tr}C^k = 0$.

1.7.4 Gram 矩阵

设 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 是 n 维欧式空间中的 s 个线性无关的向量, 称 $s \times s$ 矩阵

$$G_s = \begin{pmatrix} (\alpha_1, \alpha_1) & (\alpha_1, \alpha_2) & \cdots & (\alpha_1, \alpha_s) \\ (\alpha_2, \alpha_1) & (\alpha_2, \alpha_2) & \cdots & (\alpha_2, \alpha_s) \\ \vdots & \vdots & \ddots & \vdots \\ (\alpha_s, \alpha_1) & (\alpha_s, \alpha_2) & \cdots & (\alpha_s, \alpha_s) \end{pmatrix} = \begin{pmatrix} G_{s-1} & x \\ x' & (\alpha_s, \alpha_s) \end{pmatrix}$$

为向量组 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 的 Gram 矩阵.
下面来求 G 的行列式: 以 $\alpha_1, \cdots, \alpha_s$ 为列向量排成 $n \times s$ 矩阵 A, 那么 $G_s = A'A$. 设 $\alpha_1, \cdots, \alpha_s$ 施密特正交化后得到的向量为 $\beta_1, \cdots, \beta_s$. 以 $\beta_1, \cdots, \beta_s$ 为列向量排成 $n \times s$ 矩阵 B, 那么 $B = AQ$, 这里 Q 是一个对角线上都是1的上三角矩阵. 从而 $B'B = Q'G_sQ$. 注意这里

$$B'B = \begin{pmatrix} (\beta_1,\beta_1) & (\beta_1,\beta_2) & \cdots & (\beta_1,\beta_s) \\ (\beta_2,\beta_1) & (\beta_2,\beta_2) & \cdots & (\beta_2,\beta_s) \\ \vdots & \vdots & \ddots & \vdots \\ (\beta_s,\beta_1) & (\beta_s,\beta_2) & \cdots & (\beta_s,\beta_s) \end{pmatrix} = \begin{pmatrix} (\beta_1,\beta_1) & & & \\ & (\beta_2,\beta_2) & & \\ & & \ddots & \\ & & & (\beta_s,\beta_s) \end{pmatrix}$$

而且 $|Q'| = |Q| = 1$. 所以在 $B'B = Q'G_sQ$ 两边求行列式即得 $|G_s| = \|\beta_1\|^2\|\beta_2\|^2 \cdots \|\beta_s\|^2$. 由于 A 的行列式就是 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 张成的平行多面体的体积, 所以 G 的行列式就是这个体积的平方.

1.7.5 循环矩阵

对于给定的 $\mathbb{C}^n$ 中的向量 $\alpha = (a_0, a_1, \cdots, a_{n-1})$, 定义由 α 生成的循环矩阵

$$circ[\alpha] = circ[a_0, a_1, \cdots, a_{n-1}] = \begin{pmatrix} a_0 & a_1 & \cdots & a_{n-1} \\ a_{n-1} & a_0 & \cdots & a_{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ a_1 & a_2 & \cdots & a_0 \end{pmatrix}$$

设所有循环矩阵组成的集合为 S.
命题1: S 是 $M_n(\mathbb{C})$ 的一个子代数: S 既是一个向量空间, 又是一个环(对乘法封闭).
证明: S 是一个向量空间:

$$c \cdot circ[a_0, a_1, \cdots, a_{n-1}] = circ[ca_0, ca_1, \cdots, ca_{n-1}]$$

$$circ[a_0, a_1, \cdots, a_{n-1}] + circ[b_0, b_1, \cdots, b_{n-1}] = circ[a_0 + b_0, a_1 + b_1, \cdots, a_{n-1} + b_{n-1}]$$

S 对乘法封闭: 设 $\pi = circ[0, 1, \cdots, 0]$, 换言之 π 就是轮换 $(123 \cdots n)$ 对应的置换矩阵, 那么 $\pi^n = I$, 不难验证 $circ[\alpha] = a_0 I + a_1\pi + a_2\pi^2 + \cdots + a_{n-1}\pi^{n-1}$, 从而所有循环矩阵都可以表示为 π 的多项式, 那么它们任何两个的乘积也能.
命题2: 对于循环矩阵 A, 其伴随矩阵也是循环的. 特别如果 A 可逆, 那么 A^{-1} 也是循环的.
证明: 根据1.7.1的结论, 伴随矩阵总是原矩阵的多项式, 从而根据命题1的结论也是循环的.
命题3: 所有循环矩阵可以同时对角化.
证明: 设 ε 是一个本原 n 次单位根, 矩阵

$$\begin{pmatrix} 1 & \varepsilon & \cdots & \varepsilon^{n-1} \\ \varepsilon & \varepsilon^2 & \cdots & \varepsilon^n \\ \vdots & \vdots & \ddots & \vdots \\ \varepsilon^{n-1} & \varepsilon^n & \cdots & \varepsilon^{2n-2} \end{pmatrix}$$

是可逆的(Vandermonde 行列式), 所以其列向量线性无关. 但每一个列向量都是任意一个循环矩阵的特征向量, 这是因为如果设 $f(x)=a_0+a_1x+\cdots+a_{n-1}x^{n-1}$, 那么

$$\begin{pmatrix} a_0 & a_1 & \cdots & a_{n-1} \\ a_{n-1} & a_0 & \cdots & a_{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ a_1 & a_2 & \cdots & a_0 \end{pmatrix}\begin{pmatrix} \varepsilon^i \\ \varepsilon^{i+1} \\ \vdots \\ \varepsilon^{i+n-1} \end{pmatrix}=f(\varepsilon)\begin{pmatrix} \varepsilon^i \\ \varepsilon^{i+1} \\ \vdots \\ \varepsilon^{i+n-1} \end{pmatrix}$$

从而所有循环矩阵可以同时对角化.
特别注意 S 作为一个代数仅由一个元素 π 生成, 但是看作一个向量空间却是 n 维的. 这是因为 π 的极小多项式为 x^n-1, 所以 $I,\pi,\cdots,\pi^{n-1}$ 线性无关.

1.7.6 Hilbert 矩阵

以下记 F 为数域, $F[x]$ 为 F 上的一元多项式环. $F[x]$ 上的两个重要的线性变换是

$$Af(x)=f'(x),\quad Bf(x)=xf(x).$$

容易验证 A,B 满足关系 $AB-BA=I$. 我们知道在有限维空间 V 上是找不出一对变换 A,B 使得 $AB-BA=I$ 来的, 因为两边求迹就得出矛盾. 但是 $F[x]$ 是无限维的, 所以有这样的 A,B 存在.
下面只考虑 $\mathbb{R}[x]$ 的由$1,x,\cdots,x^{n-1}$ 生成的 n 维子空间 V. 在内积

$$<f,g>=\int_0^1 f(x)g(x)\,\mathrm{d}x$$

下 V 成为一个欧式空间. 这个内积在基$1,x,\cdots,x^{n-1}$ 下的度量矩阵为

$$G=(\frac{1}{i+j+1}),\quad 0\leqslant i,j\leqslant n-1.$$

G 又叫做 Hilbert 矩阵, 它是正定矩阵, 因为内积的度量矩阵总是正定的.
下面来算 G 的行列式: 更一般地, 考虑 n 阶矩阵

$$G=(\frac{1}{\alpha_i+\beta_j}),\quad 1\leqslant i,j\leqslant n$$

下面证明

$$\det(G)=\frac{\prod\limits_{1\leqslant i<j\leqslant n}(\alpha_i-\alpha_j)(\beta_i-\beta_j)}{\prod\limits_{1\leqslant i,j\leqslant n}(\alpha_i+\beta_j)}$$

这是因为设

$$\det(G)\prod_{1\leqslant i,j\leqslant n}(\alpha_i+\beta_j)=F(\alpha_1,\cdots,\alpha_n,\beta_1,\cdots,\beta_n)$$

那么 F 是一个 n^2-n 次的多项式. 如果某个 $\alpha_i=\alpha_j$, 那么 G 的第 i,j 两行相同, $\det(G)=0$, 所以 F 有因子

$$\prod_{1\leqslant i<j\leqslant n}(\alpha_i-\alpha_j)$$

类似地, F 还有因子

$$\prod_{1\leqslant i<j\leqslant n}(\beta_i-\beta_j)$$

而

$$\prod_{1\leqslant i<j\leqslant n}(\alpha_i-\alpha_j)(\beta_i-\beta_j)$$

已经是一个 n^2-n 次的多项式, 再比较一下首项系数就可得证.
特别令 $\alpha_i=i+\frac{1}{2},\beta_j=j+\frac{1}{2}$, 就可得到 G 的行列式为

$$\det(G)=\frac{\prod\limits_{0\leqslant i<j\leqslant n-1}(i-j)^2}{\prod\limits_{0\leqslant i,j\leqslant n-1}(i+j+1)}=\frac{[1!2!\cdots(n-1)!]^3}{n!(n+1)!\cdots(2n-1)!}$$

1.7.7 正规矩阵

正规变换有通常有两种定义:
(1)形式定义: $AA^*=A^*A$.
(2)内蕴定义: 如果 M 是 A 的不变子空间, 那么 $M^\perp$ 也是 A 的不变子空间.
这两个定义同样重要. 形式定义可以让我们快速判断一个变换是否是正规变换, 内蕴定义让我们从结构上把握正规变换的性质. 在进入结构定理前先列举一些特殊的正规变换:

欧式空间	酉空间
对称变换 $A'=A$	Hermite变换 $A^*=A$
正交变换 $A'=A^{-1}$	酉变换 $A^*=A^{-1}$
反对称变换 $A'=-A$	反对称变换 $A^*=-A$

结构定理

酉空间上的正规变换结构: 存在一组标准正交基使得在这组基下成为对角形.
证明: 对维数归纳, 取一个模长为1的特征向量 $A\alpha=\lambda\alpha$, 记 α 生成的一维子空间为 M, 考察 $M^\perp$: 由于是正规变换, $M^\perp$ 也是 A 的不变子空间, 所以由归纳假设有 $M^\perp$ 的一组标准正交基 $\beta_2,\cdots,\beta_n$ 使得在这组基下 A 在 $M^\perp$ 上的限制是对角形, 把 $\alpha,\beta_2,\cdots,\beta_n$ 合起来作为空间的一组标准正交基, 在这组基下 A 就是对角形.

欧式空间上的正规变换结构: 存在一组标准正交基使得在这组基下为准对角形(矩阵中的都是实数)

$$\begin{pmatrix} \begin{pmatrix} a_1 & b_1 \\ -b_1 & a_1 \end{pmatrix} & & & & & \\ & \ddots & & & & \\ & & \begin{pmatrix} a_i & b_i \\ -b_i & a_i \end{pmatrix} & & & \\ & & & \lambda_{2i+1} & & \\ & & & & \ddots & \\ & & & & & \lambda_n \end{pmatrix}$$

证明: 仍是对 n 归纳: 如果有实特征值, 就取长度为1的特征向量, 记该向量生成的一维子空间为 M, 然后对 $M^{\perp}$ 用归纳假设即可. 如果没有实特征值, 取一个复特征值 λ 和对应的复特征向量 $X_0: AX_0 = \lambda X_0$. 从两边分别来考察 $X_0'AX_0$:

(1)由于 $AX_0 = \lambda X_0$, 所以 $X_0'AX_0 = \lambda X_0'X_0$.

(2)由于 X_0 是 A 对应特征值 λ 的特征向量, 那么 X_0 也是 $A^* = A'$ 对应特征值 $\bar{\lambda}$ 的特征向量: $A'X_0 = \bar{\lambda}X_0$. 两边求转置: $X_0'A = \bar{\lambda}X_0'$. 从而 $X_0'AX_0 = \bar{\lambda}X_0'X_0$. 所以 $\lambda X_0'X_0 = \bar{\lambda}X_0'X_0$. 但是 λ 不是实数, 所以只能有 $X_0'X_0 = 0$.

然后设 $\lambda = a + ib, X_0 = \alpha + i\beta$, 这里 α, β 是实向量且不妨设 α 的长度是1, 那么 $AX_0 = \lambda X_0$ 说明

$$(A\alpha, A\beta) = (\alpha, \beta)\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$$

即 α, β 生成一个2维不变子空间 M. $X_0'X_0 = 0$ 说明 $\alpha'\alpha = \beta'\beta = 1, \alpha'\beta = \beta'\alpha = 0$. 即 α, β 构成 M 的一组标准正交基. 剩下的对 $M^{\perp}$ 用归纳假设即可.

注: 若 X_0 是 A 的对应特征值 λ 的特征向量, 则 X_0 是 A^* 的对应特征值 $\bar{\lambda}$ 的特征向量. 这个是因为正规变换有这样的性质: $|Ax| = |A^*x|$. 所以 $(A - \lambda)\alpha = 0 \Rightarrow (A^* - \bar{\lambda})\alpha = 0$.

正规变换与特征值的关系:

Hermite变换	$\Leftrightarrow$	特征值都是实数的正规变换
酉变换	$\Leftrightarrow$	特征值的模长都是1的正规变换
反对称变换	$\Leftrightarrow$	特征值的实部是0的正规变换

两个引理

引理1: 设 A 是正规变换, B 是线性变换, 如果 $AB = BA$, 那么有 $A^*B = BA^*$.

证明: 由于 V 可以分解为 A 的特征子空间的直和, 所以 A, B 交换等价于 A 的特征子空间都是 B 的不变子空间. 但是 A 的特征子空间也恰好是 A^* 的特征子空间, 所以 A^* 的特征子空间也都是 B 的不变子空间. 所以 A^* 和 B 交换.

引理2: 设 A, B 是正规矩阵而且 $AC = CB$, 那么 $A^*C = CB^*$.
证明: 考察

$$X = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \quad S = \begin{pmatrix} 0 & C \\ 0 & 0 \end{pmatrix}$$

那么 X 仍是正规矩阵而且 $XS = SX$. 所以由引理1, $X^*S = SX^*$. 也就是 $A^*C = CB^*$.

正规矩阵 A^* 都是 A 的多项式

证明 A 是正规变换当且仅当 A^* 可以表示为 A 的多项式.
证明: 设 A 是 n 维酉空间 V 上的正规变换, $\lambda_1, \lambda_2, \cdots, \lambda_s$ 是 A 的全部不同的特征值, $V_1, V_2, \cdots, V_s$ 是对应的特征子空间, $P_1, P_2, \cdots, P_s$ 是 V 到 $V_1, V_2, \cdots, V_n$ 的射影, 则

$$A = \lambda_1 P_1 + \lambda_2 E_2 + \cdots + \lambda_s E_s$$

其中 $P_i^2 = P_i, P_iP_j = 0, i \neq j$. 那么

$$A^2 = \lambda_1^2 P_1 + \cdots + \lambda_s^2 P_s,$$

可见对任何正整数 k,

$$A^k = \lambda_1^k P_1 + \cdots + \lambda_s^k P_s,$$

从而对任何多项式 $f(x)$,

$$f(A) = f(\lambda_1)P_1 + \cdots + f(\lambda_s)P_s$$

设 $p(x)$ 是 Lagrange 插值多项式

$$p(\lambda) = \prod_{i \neq 1} \frac{\lambda - \lambda_i}{\lambda_1 - \lambda_i}, \quad p(\lambda_1) = 1, \quad p(\lambda_i = 0), \quad i = 2, \cdots, s$$

则 $p(A) = \sum\limits_{i=1}^{s} p(\lambda_i)P_i = P_1$, 即 P_1 可以表示为 A 的多项式, 同理每个 P_i 都可以表示为 A 的多项式, 从而

$$A^* = \bar{\lambda}_1 P_1 + \cdots + \bar{\lambda}_s P_s$$

可以表示为 A 的多项式.

复正规矩阵相似等价于酉相似

设 A, B 是正规矩阵且存在可逆矩阵 T 使得 $T^{-1}AT = B$, 则存在酉矩阵 U 使得 $U^{-1}AU = B$.
思路: 设可逆矩阵 T 使得 $T^{-1}AT = B$, 要找一个酉矩阵 U 使得 $U^{-1}AU = B$. 怎么从相似过渡到酉相似呢? 用极分解, 取正定 Hermite 矩阵 S 和酉矩阵 U 使得 $T = SU$, 那么$U^{-1}(S^{-1}AS)U = B$. 如果能证明 S 与 A 交换, 那结论就成立了. 而 S 又是 TT^* 的平方

根, 所以只要证明 A 与 TT^* 交换. 下面两个证明都是要证这件事.
证法1: 首先由引理2, $A^*T = TB^*$, 求共轭就得到 $T^*A = BT^*$. 那么

$$AT = TB \Rightarrow ATT^* = TBT^* = TT^*A$$

即 A 与 TT^* 交换.
证法2: 由于 A, B 相似, 所以它们有同样的特征值, 所以存在同样的多项式 $f(x)$ 使得 $A^* = f(A), B^* = f(B)$ (插值多项式是只用特征值作出来的). 那么显然有 $T^{-1}A^*T = B^*$. 求共轭: $T^*A(T^*)^{-1} = B$. 所以

$$T^{-1}AT = T^*A(T^*)^{-1} = B$$

这就得到了 A 与 TT^* 交换.

实矩阵酉相似等价于正交相似

设 A, B 是实矩阵, U 是酉矩阵, $U^{-1}AU = B$, 则存在正交矩阵 O 使得 $O^{-1}AO = B$.
思路: 设 $U = P + iQ$, P, Q 是实矩阵, 那么对任何实数 λ 都有 $A(P + \lambda Q) = (P + \lambda Q)B$, 特别取一个实数 λ 使得 $T = P + \lambda Q$ 可逆, 那么就得到了一个可逆的实矩阵 T 使得 $T^{-1}AT = B$. 所以只要证明 A 与 TT^* 交换, 就可以复制上面的极分解的手段了.
首先由 $U^{-1}AU = B$ 可得 $U^*A = BU^*$, 从而 $(P^* + \lambda Q^*)A = B(P^* + \lambda Q^*)$, $T^*A = BT^*$. 那么 $ATT^* = TBT^* = TT^*A$, 得证.

2

问题集

2.1 问题

问题 1: 设 $A \in M_n(\mathbb{K})$ 且 $\mathrm{tr}A = 0$, 求证 A 相似于 $M_n(\mathbb{K})$ 中一个对角线上都是0的方阵.

问题 2: 设 A 是 $m \times n$ 矩阵, B 是 $n \times m$ 矩阵. 求证存在 $m \times n$ 矩阵 C 使得 $A = ABC$ 当且仅当 $r(A) = r(AB)$.

问题 3: A, B 是 n 阶正交矩阵且 $|A| = -|B|$, 求证 $|A+B| = 0$.

问题 4: A, B 是两个奇数阶的实方阵且 $AB = BA$, 求证 A, B 有共同的特征向量.

问题 5: A, B 是数域 F 上的 n 阶方阵, 若 $I - BA$ 可逆, 则 $I - AB$ 也可逆, 并求其逆.

问题 6: 求证在 n 维欧式空间中两两夹钝角的向量的个数的最大值是 $n+1$.

问题 7: 求证 n 阶实矩阵 A 是对称矩阵的充要条件是 $A^2 = A'A$.

问题 8: 设 $A, B \in M_n(\mathbb{R})$ 且 A, B对称. 求证 $\mathrm{tr}[(AB)^2] \leqslant \mathrm{tr}(A^2B^2)$.

问题 9: 设 A, B 是 n 阶实方阵满足 $A'A = B'B$, 求证存在正交矩阵 O 使得 $B = OA$.

问题 10: 设 A, B 为 n 阶实正交方阵, 证明: $|A| = |B|$ 当且仅当 $n - r(A+B)$ 为偶数.

问题 11: 设 A 是欧式空间 V 上的正交变换, 且$\det A = 1$. 求证存在 V 上的正交变换 B 使得 $A = B^2$.

问题 12: 设 A, C 是 n 阶正定矩阵, 而且矩阵方程 $AX + XA = C$ 有唯一的解 B, 求证 B 也是正定的.

问题 13: A, B 都是 n 阶半正定矩阵, 求证 AB 的特征值都是实数.

问题 14: 已知 n 阶复矩阵 A 的特征值互不相同, 求证与 A 交换的方阵可以表示为 A 的多项式.

问题 15: A, B 都是数域 F 上的 n 阶方阵.
(1)若 AB 是幂零矩阵(存在一个正整数 k 使得 $(AB)^k = 0$). 那么 BA 是否也是幂零矩阵?
(2)若 AB 是幂等矩阵 ($(AB)^2 = AB$), 那么 BA 是否也是幂等矩阵?
证明你的结论.

问题 16: 实的 n 维向量空间 V 上是否存在线性变换 A 使得 $A^2 = -I$?

问题 17: 设 n 阶方阵 A 满足 $A + A' = I$, 求证 A 可逆.

问题 18: (2005华中科技大学)求证不存在正交矩阵 A, B 满足 $A^2 = AB + B^2$.

问题 19: 设 A, B 都是 n 阶幂等矩阵: $A^2 = A, B^2 = B$, 而且 $I - A - B$ 可逆. 求证 $\mathrm{tr}(A) = \mathrm{tr}(B)$.

问题 20: (2006中科院)设 a 为实数,

$$A = \begin{pmatrix} a & 1 & & \\ & a & \ddots & \\ & & \ddots & 1 \\ & & & a \end{pmatrix} \in \mathbb{R}^{100\times 100}$$

求 A^{50} 的第一行元素之和.

问题 21: 设 A, B 是实方阵且 A, B 的特征值都是正数, $A^2 = B^2$, 求证 $A = B$.

问题 22: 判断 n 阶实矩阵

$$A = \begin{pmatrix} & & & & a_1 \\ & & & a_2 & \\ & & \iddots & & \\ & a_{n-1} & & & \\ a_n & & & & \end{pmatrix}$$

何时在实数域内相似于对角形.

问题 23: 设 A 是数域 F 上的 n 阶方阵, 而且 A 的秩是 r, 求证 A 的极小多项式的次数至多是 $r+1$.

问题 24: 已知 A,B 都是 n 阶复方阵, 且 $AB-BA$ 是 A 和 B 的线性组合, 求证 A,B 可以同时上三角化.

问题 25: 设 A,B 是 n 阶方阵且 B 可逆, $\mathrm{rank}(I-AB)+\mathrm{rank}(I+BA)=n$, 求证 A 也可逆.

问题 26: 设 A,B 是两个 n 阶矩阵满足 $\mathrm{rank}(A+B)=\mathrm{rank}(A)+\mathrm{rank}(B)$, 且 $A+B$ 是幂等的. 求证 A 和 B 也都是幂等矩阵而且 $AB=BA=0$.

问题 27: (中科院2007)设二次型 $f=x_1^2+x_2^2+x_3^2+2ax_1x_2+2x_1x_3+4bx_2x_3$ 通过正交变换化成标准形 $f=y_2^2+2y_3^2$, 求参数 a,b 及所用的正交变换.

问题 28: 已知 $f(x)$ 是一个没有实根的实系数多项式, 求证存在实系数多项式 $g(x),h(x)$ 使得 $f(x)=[g(x)]^2+[h(x)]^2$.

问题 29: 设 A 是4维欧式空间上的正交变换, A 无特征值, 但是 A^2,A^3 均有特征值. 求 A 的极小多项式.

问题 30: (2009南开大学)设 F 是一个数域, T 是 $M_n(F)$ 上的线性变换满足对任何给定的 $A,B\in M_n(F)$, $T(AB)=T(A)T(B)$ 和 $T(AB)=T(B)T(A)$ 至少有一个成立. 求证: 要么对所有的 $A,B\in M_n(F)$ 都有 $T(AB)=T(A)T(B)$ 成立, 要么对所有的 $A,B\in M_n(F)$ 都有 $T(AB)=T(B)T(A)$ 成立.

问题 31: 设 V 是一个有理数域 $\mathbb{Q}$ 上的有限维向量空间, A 是 V 上一个线性变换满足 $A^3+A-I=0$. 求证 V 的维数 $\dim V$ 能被3整除.

问题 32: (科大教材523页习题15) 求证两个半正定矩阵可以同时合同于对角矩阵.

问题 33: 设 J 是一个 Jordan 块:

$$A=\begin{pmatrix} a & 1 & & \\ & a & \ddots & \\ & & \ddots & 1 \\ & & & a \end{pmatrix}_{n\times n}$$

其中 a 是实数, $n>1$. 如果存在 n 阶实矩阵 B 使得 $B^2=J$, 就称 J 可以在实数域内开平方. 证明当 $a>0$ 时 J 可以在实数域内开平方, $a\leqslant 0$ 时则不能.

2.2 解答

解答 1: 由于 A 不是数量矩阵, 所以存在向量 α 使得 α 与 $A\alpha$ 线性无关. 扩充为一组基 $\alpha, A\alpha, \beta_1, \cdots, \beta_{n-2}$, 则在这组基下 A 的矩阵形如

$$A = \begin{pmatrix} 0 & a_{12}\cdots & a_{1n} \\ 1 & a_{22}\cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ 0 & a_{n2}\cdots & a_{nn} \end{pmatrix} = \begin{pmatrix} 0 & M \\ N & B \end{pmatrix}$$

其中 B 的迹仍是0. 对 B 用归纳假设即可.

解答 2: $\Rightarrow$:

$$r(A) = r(ABC) \leqslant r(AB) \leqslant r(A)$$

等号必须全部成立, 从而 $r(A) = r(AB)$.

$\Leftarrow$: 如果 $r(A) = r(AB)$, 注意到 AB 的列向量是 A 的列向量的线性组合, 从而 $r(A) = r(AB)$ 说明 A 和 AB 的列向量是等价的, 它们可以互相线性表出. 那么设 $\alpha_1, \cdots, \alpha_n$ 是 A 的列向量, 每个非齐次线性方程组 $ABX = \alpha_i$ 都是有解 X_i 的. 把 $X_1, \cdots, X_n$ 作为列向量排成矩阵 C, 就有 $ABC = A$. 得证.

解答 3: 考虑正交矩阵 AB^{-1}, 问题是说第二类正交变换必有-1为其特征值.

解答 4: 证明: 关键是看到 A 有一个奇数维的实的根子空间 W, 易见 W 也是 B 的不变子空间, B 限制在 W 上还是一个奇数次的线性变换, 所以只要在 W 中去看问题, 问题转化为

设 A, B 是两个奇数阶的实方阵且 $AB = BA$, A 的特征值都是实数, 求证 A, B 有共同的特征向量.

这个时候调过来看 B: B 必定有实的特征子空间 N, 由于交换性所以 N 还是 A 的不变子空间. A 限制在其上必有复特征值和复特征向量, 但是由于 A 的特征值都是实数, 所以对应的特征向量也是实的, 这就找到了共同的特征向量.

解答 5: 首先用纯形式的推导找出这个逆来. 把矩阵看成数, 作幂级数展开:

$$\begin{aligned}(I - AB)^{-1} &= I + AB + (AB)^2 + \cdots \\ &= I + A(I + BA + (BA)^2 + \cdots)B \\ &= I + A(I - BA)^{-1}B\end{aligned}$$

所以 $(I - AB)^{-1} = I + A(I - BA)^{-1}B$, 这就从形式上找出来了结果, 剩下的只是验证而已.

解答 6: 用归纳法分两步, 首先证明至多有 $n+1$ 个向量两两夹钝角:

$n = 1$ 是显然的, 设命题对 $n-1$ 维欧式空间成立, 考察 n 维的情形: 如果有

$n+2$ 个向量 $\alpha_1, \alpha_2, \cdots, \alpha_{n+2}$ 两两之间夹钝角, 考虑 α_1 及其生成的一维子空间 $\{\alpha_1\}$ 的正交补 $M^\perp$, 任何 $\alpha_2, \cdots, \alpha_{n+2}$ 有唯一的表示

$$\alpha_i = c_i\alpha_1 + \beta_i \quad \beta_i \in M^\perp, i = 2, 3, \cdots, n+2$$

显然 $c_i < 0$. 考察 $n-1$ 维欧式空间 $M^\perp$ 中的 $n+1$ 个向量 $\beta_2, \cdots, \beta_{n+2}$, 我们有

$$(\alpha_i\,, \alpha_j) = (c_i\alpha_1 + \beta_i\,, c_j\alpha_1 + \beta_j) = c_ic_j|\alpha_1|^2 + (\beta_i\,, \beta_j)$$

即

$$(\beta_i\,, \beta_j) = (\alpha_i\,, \alpha_j) - c_ic_j|\alpha_1|^2 < 0$$

即 $\beta_2, \cdots, \beta_{n+2}$ 两两夹钝角, 但 $M^\perp$ 的维数是 $n-1$, 这与归纳假设矛盾! 所以 n 维欧式空间中至多有 $n+1$ 个向量两两夹钝角.
其次证明至少有 $n+1$ 个向量两两夹钝角: 仍是对 n 归纳, $n=1$ 显然, 设 $n-1$ 维的时候结论成立, 考察 n 维的情形:
首先取一个 $n-1$ 维的子空间 M, 根据归纳假设, 在 M 内有 n 个两两夹钝角的向量 $\alpha_1, \alpha_2, \cdots, \alpha_n$, 似乎只要再找一个向量与它们都夹钝角即可. 而从 $n=1$ 到 $n=2$ 的情形告诉我们这走不通. (验证一下!) 所以要对 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 进行一下微调, 给它们同时加上一个向量. 设 β 是 $M^\perp$ 中的非零单位向量, 取正数 λ 使得其满足

$$-2\lambda^2 > \max\{(\alpha_i\,, \alpha_j), \quad 1 \leqslant i < j \leqslant n\}$$

那么

$$(\alpha_i - \lambda\beta\,, \beta) = -\lambda < 0$$
$$(\alpha_i - \lambda\beta\,, \alpha_j - \lambda\beta) = (\alpha_i\,, \alpha_j) + \lambda^2 < -\lambda^2 < 0$$

从而

$$\alpha_1 - \lambda\beta, \alpha_2 - \lambda\beta, \cdots, \alpha_n - \lambda\beta, \beta$$

就是满足要求的 $n+1$ 个向量.

解答 7: 两边求迹:

$$\sum_{i=1}^n\sum_{j=1}^n a_{ij}a_{ji} = \sum_{i=1}^n\sum_{j=1}^n a_{ij}^2$$

两边乘以2:

$$\sum_{i=1}^n\sum_{j=1}^n 2a_{ij}a_{ji} = \sum_{i=1}^n\sum_{j=1}^n 2a_{ij}^2 = \sum_{i=1}^n\sum_{j=1}^n a_{ij}^2 + a_{ji}^2$$

变形:

$$\sum_{i=1}^n\sum_{j=1}^n (a_{ij} - a_{ji})^2 = 0$$

从而 $a_{ij} = a_{ji}$, 即 A 对称.
或者也可以利用 $\mathrm{tr}(A-A')(A-A')' \geqslant 0$ 且等号成立当且仅当 $A = A'$.

解答 8: 令 $C = AB - BA$, 则 $\mathrm{tr}CC' \geqslant 0$. 展开整理:

$$\mathrm{tr}(ABBA + BAAB) \geqslant \mathrm{tr}(ABAB + BABA)$$

由于 $\mathrm{tr}(ABB \cdot A) = \mathrm{tr}(A \cdot ABB) = \mathrm{tr}(A^2B^2)$, $\mathrm{tr}(B \cdot AAB) = \mathrm{tr}(AAB \cdot B) = \mathrm{tr}(A^2B^2)$, 所以

$$\mathrm{tr}(ABBA + BAAB) = 2\mathrm{tr}(A^2B^2)$$

同理可得 $\mathrm{tr}(ABAB + BABA) = 2\mathrm{tr}[(AB)^2]$. 从而问题得证.

解答 9: 用极分解: 设 $A = O_1P, B = O_2Q$, 这里 O_1, O_2 是正交矩阵, P, Q 是半正定矩阵, 那么 $A'A = P'P = P^2 = Q^2 = B'B$, 从而根据平方根的唯一性知道 $P = Q$. 所以 $B = (O_2O_1^{-1})A$.

解答 10: 考察正交矩阵 $C = AB^{-1}$, 那么 $\mathrm{rank}(A+B) = \mathrm{rank}(C+I)$, 而且 $n-\mathrm{rank}(C+I)$ 就是 C 的特征值中-1的个数. 所以 C 是第一类的当且仅当C的特征值中-1的个数为偶数.

解答 11: 首先存在一组标准正交基使得 A 在这组基下的矩阵形如

$$A = \begin{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} & & & & & \\ & \ddots & & & & \\ & & \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} & & & \\ & & & 1 & & \\ & & & & \ddots & \\ & & & & & 1 \end{pmatrix}$$

注意由于 $\det A = 1$, 所以 A 的特征值中-1的个数是偶数, 所以可以把特征值中的-1两两组合使得 A 成为上面的形状. 现在根据旋转的几何直观不难有

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} \cos\frac{\theta}{2} & -\sin\frac{\theta}{2} \\ \sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{pmatrix}^2$$

$$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} \cos\frac{\pi}{2} & -\sin\frac{\pi}{2} \\ \sin\frac{\pi}{2} & \cos\frac{\pi}{2} \end{pmatrix}^2$$

可见只要令

$$B=\begin{pmatrix} \begin{pmatrix} \cos\frac{\theta}{2} & -\sin\frac{\theta}{2} \\ \sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{pmatrix} & & & & & \\ & \ddots & & & & \\ & & \begin{pmatrix} \cos\frac{\pi}{2} & -\sin\frac{\pi}{2} \\ \sin\frac{\pi}{2} & \cos\frac{\pi}{2} \end{pmatrix} & & & \\ & & & 1 & & \\ & & & & \ddots & \\ & & & & & 1 \end{pmatrix}$$

即可.

解答 12: 显然 B' 也是解, 所以由解的唯一性知道 $B=B'$, 即 B 对称. 所以只要再证明 B 的特征值都大于0即可. 设 λ 是 B 的特征值: $B\alpha=\lambda\alpha$, 那么由 C 的正定性知道

$$\alpha'(AB+BA)\alpha=\alpha'C\alpha>0$$

即

$$2\lambda\alpha'A\alpha>0$$

再根据 A 的正定性就得到 $\lambda>0$, 从而问题得证.

解答 13: 用平方根分解, 设 $A=P^2$, 这里 P 是半正定矩阵, 那么 $AB=P^2B$, 且 P^2B 与 PBP 有同样的特征根, 而 PBP 是对称矩阵, 其特征根都是实数, 所以 AB 的特征根都是实数.

解答 14: 设 λ_i 对应的特征向量为 α_i, 去证明 $v=\alpha_1+\alpha_2+\cdots+\alpha_n$ 是循环向量, 即 $v,Av,\cdots,A^{n-1}v$ 是空间的一组基. 从而由1.5的命题1的结论即得所证.

解答 15: (1) BA 是幂零的. 你可以用 AB 和 BA 的特征值是相同的这个结论, 也可以从

$$AB\cdots AB=A(BA\cdots BA)B$$

来做.

(2) BA 不一定是幂等的. 比如

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

这个反例是怎么找呢? 当然不是凑出来的. 如果 AB 幂等的话那么

$$BABABA=B(ABAB)A=BABA$$

也就是 $(BA)^3=(BA)^2$. 那么有什么矩阵三次方和二次方相等但是不是幂等的呢? 0特征值的2阶 Jordan 块就可以. 这样就好办了.

解答 16: n 是偶数的时候有, n 是奇数的时候就没有. 显然 A 的特征根只有 $\pm i$, 而奇数阶矩阵必有实特征根, 所以 n 不能是奇数. 至于偶数的情形,

$$A = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}$$

就满足要求.

注: 实际上任何满足 $A^2 + I = 0$ 的实方阵都相似于上面的矩阵. 这是因为 A 在复数域内可以对角化为

$$A = \begin{pmatrix} iI_n & 0 \\ 0 & -iI_n \end{pmatrix}$$

那么用酉矩阵

$$U = \frac{1}{\sqrt{2}} \begin{pmatrix} I_n & -iI_n \\ I_n & iI_n \end{pmatrix}$$

作相似变换

$$U^{-1}AU = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}$$

从而 A 复相似于

$$\begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}$$

当然也就实相似于这个实矩阵.

解答 17: 问题本质就是 A 可以表示为一个正定矩阵 $(I/2)$ 与一个反对称矩阵的和, 所以行列式大于0. 但是有更简洁的证法: 对于非零的 n 维向量 X 总有 $X'X = X'(A + A')X = 2X'AX > 0$, 所以必然 A 可逆.

解答 18: 在等式两边左乘以 A^{-1}, 右乘以 B^{-1} 可得 $AB^{-1} = I + A^{-1}B$. 即 $AB' = I + A'B$. 两边求迹:

$$\mathrm{tr}(AB') = n + \mathrm{tr}(A'B)$$

但是

$$\mathrm{tr}AB' = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}b_{ij} = \mathrm{tr}A'B$$

(或者 $\mathrm{tr}AB' = \mathrm{tr}B'A = \mathrm{tr}(B'A)' = \mathrm{tr}A'B$)

从而 $0 = n$, 矛盾!

解答 19: 只要注意到 $(I - A - B)A = B(I - A - B)$ 就可得到 A, B 相似.

解答 20: 设 $x^{50}=a_0+a_1(x-a)+\cdots+a_{50}(x-a)^{50}$, 那么

$$A^{50}=\begin{pmatrix} a_0 & a_1 & \cdots & a_{50} & \cdots & 0\\ & a_0 & a_1 & \ddots & \ddots & \vdots\\ & & \ddots & \ddots & \ddots & a_{50}\\ & & & \ddots & \ddots & \vdots\\ & & & & \ddots & a_1\\ & & & & & a_0 \end{pmatrix}$$

第一行元素之和就是 $a_0+a_1+\cdots+a_{50}=(a+1)^{50}$.

解答 21: 我们有 $A(A-B)=-(A-B)B$, 如果 $A-B$ 不是0的话根据1.4的结论 A 和 $-B$ 应当有共同的特征值, 这不可能, 从而必有 $A-B=0$, 即 $A=B$, 得证.
注: 从这个证明方法可以看出来当 A,B 的特征值的实部大于0时结论也成立. 或者当"A,B 的特征值的凸包不包含原点" 时结论也成立.

解答 22: 设 n 维向量空间 V 上的线性变换 $\mathcal{A}$ 在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的矩阵为 A. 那么

$$\mathcal{A}\varepsilon_i=a_{n-i+1}\varepsilon_{n-i+1},\quad \mathcal{A}\varepsilon_{n-i+1}=a_i\varepsilon_i,$$

令 $M_i=L(\varepsilon_i,\varepsilon_{n-i+1})$, 则 M_i 为 $\mathcal{A}$ 的二维不变子空间.
(1)如果 $n=2m$ 为偶数, 那么

$$V=M_1\oplus M_2\oplus\cdots\oplus M_m$$

所以 $\mathcal{A}$ 可以对角化当且仅当 $\mathcal{A}$ 在每一个 M_i 上的限制 $\mathcal{A}|_{M_i}$ 可以对角化. $\mathcal{A}|_{M_i}$ 在基 $\varepsilon_i,\varepsilon_{n-i+1}$ 下的矩阵为

$$A_i=\begin{pmatrix} 0 & a_i\\ a_{n-i+1} & 0\end{pmatrix}$$

如果 $a_i=a_{n-i+1}=0$, 那么 A_i 自然是对角矩阵, 如果 a_i,a_{n-i+1} 中恰好有一个是0, 那么 A_i 是 Jordan 形, 不能对角化. 如果 a_i,a_{n-i+1} 都不是0, 那么 $a_ia_{n-i+1}<0$ 时 A_i 不能对角化, 因为这时 A_i 在实数域内没有特征根. 而 $a_ia_{n-i+1}>0$ 时 A_i 可以对角化, 因为这时 A_i 有两个互异的特征根.
(2) $n=2m+1$ 是奇数, 那么

$$V=M_1\oplus M_2\oplus\cdots\oplus M_m\oplus\{\varepsilon_{m+1}\}$$

可见重复上面的讨论即可.

解答 23: 设 $\eta_1,\eta_2,\cdots,\eta_{n-r}$ 是 $AX=0$ 的一组基础解系, 扩充为 $\mathbb{F}^n$ 的一组基 $\eta_1,\cdots,\eta_{n-r},\ \xi_1,\cdots,\xi_r$. 那么在这组基下 A 的矩阵形如

$$A=\begin{pmatrix} 0 & A_1\\ 0 & B_1\end{pmatrix}$$

这里 $A_1 \in \mathbb{M}_{(n-r)\times r}(\mathbb{F}), B_1 \in \mathbb{M}_{r\times r}(\mathbb{F})$. 容易验证

$$A^k = \begin{pmatrix} 0 & A_1 B_1^{k-1} \\ 0 & B_1^k \end{pmatrix}$$

现在 B_1 是一个 r 阶方阵, 其极小多项式 $p(x)$ 的次数至多为 r, 那么

$$Ap(A) = \begin{pmatrix} 0 & A_1 p(B_1) \\ 0 & B_1 p(B_1) \end{pmatrix} = 0$$

所以 A 的极小多项式的次数至多是 $r+1$. 问题得证.

解答 24: 关键还是证明 A, B 有共同的特征向量. 记 $C = AB - BA = aA + bB$, 那么 $AC - CA = bC$, 从而 $CA - AC$ 与 C 交换, 所以 C 是幂零的, 从而方程组 $CX = 0$ 的解空间 X 是非平凡的子空间. 其次在 A, B 在这个解空间 X 上的限制满足 $AB = BA$, 所以 A, B 在 X 上有共同的特征向量. 剩下的用归纳法即可.

解答 25: 注意到 $r(I + BA) = r[B(B^{-1} + A)B] = r(I + AB)$, 用正交幂等元的的结论有 $(I + AB)(I - AB) = 0, (AB)^2 = I$, 所以 A 可逆.

解答 26: 由已知不难得到 $r(A) + r(B) + r(I - A - B) = n$. 所以根据正交幂等元的结论, A, B 都是幂等的而且 $A(I - A - B) = (I - A - B)A = 0$, 即 $AB = BA = 0$. .

解答 27: 由题意,

$$\begin{pmatrix} 1 & a & 1 \\ a & 1 & 2b \\ 1 & 2b & 1 \end{pmatrix} \quad 与 \quad \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

是正交相似的. 当然有同样的特征多项式. 计算二者的特征多项式可得 $a = 2b$ 和 $a = 0$, 从而 $a = b = 0$. $f = x_1^2 + x_2^2 + x_3^2 + 2x_1x_3 = x_2^2 + (x_1 + x_3)^2$, 这个正交变换很明显了. 令 $y_2 = x_2, y_1 = \dfrac{(x_3 - x_1)}{\sqrt{2}}, y_3 = \dfrac{(x_3 + x_1)}{\sqrt{2}}$, 那么 $f = y_2^2 + 2y_3^2$. 反解出 y_1, y_2, y_3 来可得正交矩阵为

$$T = \begin{pmatrix} -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{pmatrix}$$

解答 28: 如果你对这个问题没有思路的话, 看看这个问题: 若正整数 m, n 都能写成两个整数的平方和, 那么 mn 也能. 设 $m = a^2 + b^2, n = c^2 + d^2$, 那么

$$\begin{aligned} mn &= (a^2 + b^2)(c^2 + d^2) = (a + b\mathrm{i})(a - b\mathrm{i})(c + d\mathrm{i})(c - d\mathrm{i}) \\ &= [(a + b\mathrm{i})(c + d\mathrm{i})][(a - b\mathrm{i})(c - d\mathrm{i})] \\ &= (ac - bd)^2 + (ad + bc)^2 \end{aligned}$$

现在你应该会做这道题了吧.

解答 29: A 没有特征值说明 A 的标准形为

$$B=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}\quad 和\quad C=\begin{pmatrix}\cos\varphi & -\sin\varphi\\ \sin\varphi & \cos\varphi\end{pmatrix}$$

的直和. 这里 θ 和 φ 都不能是0, π 或 2π. 其次 A^2 有特征值说明 2θ 或者 2φ 之一等于 π 或者 2π, 当然只能等于 π, 不妨设 $2\theta=\pi$, $\theta=\pi/2$. A^3 有特征值说明 3θ 或者 3φ 之一等于 π 或者 2π, 当然只能是 3φ 等于 π 或者 2π. 从而 φ 等于 $\pi/3$ 或者 $2\pi/3$. 所以极小多项式是 $(\lambda^2+1)(\lambda^2+\lambda+1)$ 或者 $(\lambda^2+1)(\lambda^2-\lambda+1)$. 前一种情况对应于 $\theta=\pi/2$, $\varphi=2\pi/3$, 后一种情况对应于 $\theta=\pi/2$, $\varphi=\pi/3$.

解答 30: 对于固定的矩阵 A, 设 V_1 是所有满足 $T(AB)=T(A)T(B)$ 的方阵 B 构成的集合, V_2 是所有满足 $T(AC)=T(C)T(A)$ 的方阵 C 构成的集合, 你可以直接验证 V_1,V_2 都是 $M_n(\mathbb{F})$ 的子空间, 而且根据已知, $M_n(\mathbb{F})=V_1\cup V_2$. 那么根据1.2的结论, V_1,V_2 不可能都是 $M_n(\mathbb{F})$ 的真子空间, 所以必有 $V_1=M_n(\mathbb{F})$ 或者 $V_2=M_n(\mathbb{F})$ 之一成立. 所以要么对所有的 B 都有 $T(AB)=T(A)T(B)$, 要么对所有的 B 都有 $T(AB)=T(B)T(A)$ 成立.

证明完了吗? 还没有. 因为有可能对某个 A_1 总是有 $T(A_1B)=T(A_1)T(B)$ 成立, 而对某个 A_2 总是有 $T(A_2B)=T(B)T(A_2)$ 成立. 我们还需要排除这种可能. 这回我们再来用一遍上面的招数: 设 U_1 是总有 $T(A_1B)=T(A_1)T(B)$ 的那些个 A_1 构成的集合, U_2 是总有 $T(A_2B)=T(B)T(A_2)$ 的那些个 A_2 构成的集合, 显然 U_1,U_2 也都是子空间. 注意由我们刚才证得的结论, $M_n(\mathbb{F})=U_1\cup U_2$. 从而 U_1,U_2 不能都是真子空间. 这样就得到了证明.

这个题目源自华罗庚的一个定理(代数学引论131页). 命题对一般的环都是成立的. 证明完全可以仿照上面的做法, 只是需要说明"一个 Abel 群不能表示为它的两个真子群的并".

解答 31: 首先不难证明题中的多项式(记作 $p(x)$)是一个有理数域上的不可约多项式, 其次要证明特征多项式 $f(x)$ 是它的若干次幂. 否则设 $f(x)=[p(x)]^l\cdot h(x)$, 这里 $p(x)$ 不整除 $h(x)$, 从而 $(p(x),h(x))=1$, 从而存在有理数域上的多项式 $u(x),v(x)$ 使得 $u(x)p(x)+v(x)h(x)=1$. 但是特征多项式和极小多项式有相同的根, 所以 $h(x)$ 的根都是 $p(x)$ 的根. 任意代 $h(x)$ 的一个根进去就得到$0=1$, 矛盾.

解答 32: 以下约定 A,B 为半正定实矩阵, X,Y 为 $n\times 1$ 向量.

引理(1): 如果 A 半正定, 那么 $X'AX=0$ 等价于 $AX=0$.

引理(1)的证明: 由于 A 半正定, 所以存在 A 的平方根分解 $A=P'P$, 那么

$$X'AX=0\Rightarrow (PX)'(PX)=0\Rightarrow PX=0\Rightarrow AX=0.$$

反之 $AX=0\Rightarrow X'AX=0$ 是显然的.

引理(2): 如果半正定矩阵的某个对角元是0, 那么该对角元所在的行和列的

所有元素都是0.

引理(2)的证明: 不妨设第 i 个对角元 $a_{ii}=0$, 那么 $e_i'Ae_i=a_{ii}=0$. 由引理1, $e_i'A=Ae_i=0$, 从而 $e_i'Ae_j=e_j'Ae_i=0$, 即 $a_{ij}=a_{ji}=0$. 这说明 A 的第 i 行和第 j 列全是0.

下面进入原问题的证明: 首先半正定性是在合同变换下保持不变的, 而且合同变换的复合仍然是合同变换. 所以首先作合同变换把 A 化成标准形

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

这时 B 仍然是半正定的, 所以我们不妨从一开始就假设 A 就是如上的标准形, 我们的思路是在保持 A 的形状不变的前提下不断地对 B 作合同变换, 把 B 化成想要的形式. 先把 B 分块:

$$B=\begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} \quad B_{12}=B_{21}$$

这里面的子块 B_{11} 和 B_{22} 都是对称的, 从而可以选取 $r\times r$ 正交矩阵 T_1 和 $(n-r)\times(n-r)$ 正交矩阵 T_2 使得 $T_1'B_{11}T_1$ 和 $T_2'B_{22}T_2$ 都是对角形:

$$T_1'B_{11}T_1=\begin{pmatrix} D_1 & 0 \\ 0 & 0 \end{pmatrix}, \quad T_2'B_{22}T_2=\begin{pmatrix} D_2 & 0 \\ 0 & 0 \end{pmatrix}$$

这里 D_1,D_2 是对角线上都不是0的对角形. 注意正交矩阵保持 A 的形状不变, 如果你用一般的矩阵给 B 作合同变换的话就不一定能保持 A 的形状不变了. 那么在矩阵

$$\begin{pmatrix} T_1 & 0 \\ 0 & T_2 \end{pmatrix}$$

的合同变换下, A 保持不变, 仍是原来的标准形, B 变为

$$\begin{pmatrix} D_1 & 0 & C & 0 \\ 0 & 0 & 0 & 0 \\ C & 0 & D_2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

其中

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

是 $B_{12}=B_{21}$ 在合同变换后的结果, 注意到根据引理2, 它们必须是这种形式.

然后用合同变换把 C 干掉:

$$\begin{pmatrix} E & 0 & -CD_2^{-1} & 0 \\ 0 & E & 0 & 0 \\ 0 & 0 & E & 0 \\ 0 & 0 & 0 & E \end{pmatrix} \begin{pmatrix} D_1 & 0 & C & 0 \\ 0 & 0 & 0 & 0 \\ C & 0 & D_2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} E & 0 & 0 & 0 \\ 0 & E & 0 & 0 \\ -CD_2^{-1} & 0 & E & 0 \\ 0 & 0 & 0 & E \end{pmatrix}$$
$$= \begin{pmatrix} D_1 - CD_2^{-1} & 0 & C & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & D_2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

注意由于 C, D_2 都是对角矩阵, 所以它们既对称又可以交换, 从而 $(CD_2^{-1})' = CD_2^{-1}$, 而且在

$$\begin{pmatrix} E & 0 & 0 & 0 \\ 0 & E & 0 & 0 \\ -CD_2^{-1} & 0 & E & 0 \\ 0 & 0 & 0 & E \end{pmatrix}$$

的合同变换下 A 仍然保持标准形不变. 这个时候由于 $D_1 - CD_2^{-1}C$, D_2 都是对角矩阵, 所以

$$\begin{pmatrix} D_1 - CD_2^{-1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & D_2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

是对角矩阵, 这就把 A, B 同时化为了对角形.

解答 33: 利用1.6的结论都不难证明. 当 $a > 0$ 的时候有一个巧妙的办法给出了 B 的算法: 记$J = aI + N$, 则 $N^n = 0$. 考虑 $\sqrt{x+a}$ 的 n 次 Taylor 多项式 $P(x)$, 则 $P(x)^2 - (x+a) = x^n Q(x)$, 而且 $Q(x)$ 也是一个多项式（不要以为是个幂级数）. 我们还有 $P(x), Q(x)$ 都是实系数的. 那么

$$P(N)^2 = N + aI + N^n Q(N) = N + aI = J$$

所以 $B = P(N)$ 满足要求. 当 $a = 0$ 的时候可以这样做, 设 $B^2 = N$, 那么 $B^{2n-2} = N^{n-1} \neq 0, B^{2n} = N^n = 0$. 但是 $B^{2n} = 0$ 意味着 $B^n = 0$ （极小多项式次数不会大于 n ）, 从而 $n \geqslant 2$ 时有 $B^{2n-2} = B^n B^{n-2} = 0$, 矛盾!