高等代数考研攻略

博士家园yinzhe整理

2009 年 1 月 19 日

目 录

1 高等代数十个问题

1.1 正规变换总结

1.1.1 基本性质

满足$AA^* = A^*A$的变换称为正规变换, 这是书上的通常的形式定义. 正规变换的内蕴定义是: **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. 特别注意这两个定义都不依赖于基域的选取. 下面列举特殊的正规变换:

实的版本	复的版本
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$A^* = -A$

正规变换的结构定理:

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(表中的都是实数)

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

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$X_0 : AX_0 = \lambda X_0$. 设$\lambda = a + ib, X_0 = \alpha + i\beta$, 这里$\alpha, \beta$是实向量且不妨设$\alpha$的长度是1, 则有

$$A\alpha = a\alpha - b\beta$$

$$A\beta = b\alpha + a\beta$$

这就说明α,β生成一个2维不变子空间. 对$AX_0=\lambda X_0$求转置: $X_0'A'=\lambda X_0'$, 再右乘X_0: $X_0'A'X_0=\lambda X_0'X_0$. 由于$A'X_0=\bar{\lambda}X_0$, 所以$\bar{\lambda}X_0'X_0=\lambda X_0'X_0$. 但是$\bar{\lambda}\neq\lambda$, 否则与$\lambda$不是实根矛盾, 所以$X_0'X_0=0$, 也就是$\alpha'\alpha=\beta'\beta=1,\alpha'\beta=\beta'\alpha=0$. 这就说明$\alpha,\beta$都是单位向量而且互相正交.

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. 这个可以用定义证明. 由于A是实矩阵, 所以$A'X_0=\bar{\lambda}X_0$.

从特征值判定正规变换的类型:

$$\begin{array}{ccc}
\text{Hermite变换} & \Leftrightarrow & \text{特征值都是实数的正规变换}\\
\text{酉变换} & \Leftrightarrow & \text{特征值的模长都是1的正规变换}\\
\text{反对称变换} & \Leftrightarrow & \text{特征值的实部是0的正规变换}
\end{array}$$

1.1.2 $A^*=f(A)$

证明A是正规变换当且仅当A^*可以表示为A的多项式.

解答: 设A是n维酉空间V上的正规变换, $\lambda_1,\lambda_2,\cdots,\lambda_s$是$A$的不同的特征值, $V_1,V_2,\cdots,V_s$是对应的特征子空间, $E_1,E_2,\cdots,E_s$ 是V到$V_1,V_2,\cdots,V_n$的射影, 则

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

$$A=\lambda_1E_1+\lambda_2E_2+\cdots+\lambda_sE_s$$

其中$E_i^2=E_i,E_iE_j=0,i\neq j$. 那么

$$A^2=\lambda_1^2E_1+\cdots+\lambda_s^2E_s,$$

显然可见对任何正整数k,

$$A^k=\lambda_1^kE_1+\cdots+\lambda_s^kE_s,$$

从而对任何多项式f,

$$f(A)=f(\lambda_1)E_1+\cdots+f(\lambda_s)E_s$$

设p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$p(A)=\sum\limits_{i=1}^{s}p(\lambda_i)E_i=E_1$, 即$E_1$可以表示为$A$的多项式, 同理每个$E_i$都可以表示为$A$的多项式, 从而

$$A^*=\bar{\lambda}_1E_1+\cdots+\bar{\lambda}_sE_s$$

可以表示为A的多项式.

1.1.3 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$A^*B = BA^*$.
这个性质不太引人注意, 但是很有用.
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X和S是交换的, 所以由引理1, X^*和S交换, 直接计算就得到了$A^*C = CB^*$.
下面是两个应用:

1.1.4 正规矩阵相似必酉相似

设A, B是正规矩阵而且存在可逆矩阵T使得$T^{-1}AT = B$, 则存在酉矩阵U使得$U^{-1}AU = B$.
给出两个证明:
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.

1.1.5 实矩阵酉相似必正交相似

设A,B是实矩阵, U是酉矩阵, $U^{-1}AU=B$, 求证存在正交矩阵O使得$O^{-1}AO=B$.
证明: 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$\lambda$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.

1.2 $T(X)=AX-XB$的讨论

(a)设A,B分别是复数域上的n,m阶矩阵, X是$n\times m$矩阵. 那么矩阵方程

$$AX=XB$$

有非零解当且仅当A,B有共同的特征值.
(b)设A,B是复数域上的n阶矩阵, 定义$M_n(\mathbb{C})$上的线性变换$T:T(X)=AX-XB$, 求T的全部特征值.
(c)T可对角化当且仅当A,B都是可对角化矩阵.
(a)的证明: $\Rightarrow$设A,B无共同的特征值, 但是有非零的$m\times n$矩阵X满足$AX=XB$, 那么$A^2X=AXB=XB^2$, 用归纳法可得对任何整数k, $A^kX=XB^k$, 从而对任何多项式$g(x)$, $g(A)X=Xg(B)$, 特别令g是A的特征多项式, 由Hamilton-Caley定理, $g(B)X=0$. 但是$g(B)$可逆, 必有$X=0$, 矛盾!(想一想,Why?)
充分性: 下面来计算解空间的维数(所有的$n\times m$解构成一个向量空间). 对于n阶的可逆矩阵P和m阶的可逆矩阵Q, 在方程$AX=XB$两边作变换

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

我们看到$X\to P^{-1}XQ$给出了解空间X到自身的一个可逆线性变换, 它不改变解空间的维数. 所以可以在A,B都是Jordan标准型的条件下考虑问题. 设A,B的标准型中的全部Jordan块分别是(k_i,m_j代表阶数)

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

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

按照A,B的方式将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. 如果记δ_{ij}当$\lambda_i=\mu_j$时值为1, 当$\lambda_i\neq\mu_j$时值为0, 那么整个X含有的参数的个数是全部X_{ij}的参数个数的和

$$\sum_{i=1}^{s}\sum_{j=1}^{t}\min\{k_i,m_j\}\delta_{ij}$$

这就是解空间的维数.

(b)如果$AX-XB=\lambda X, X\neq 0$, 那么$(A-\lambda I_n)X=XB$有非零解, 根据(a)的结论, $A-\lambda I_n$与B有共同的特征值, 所以存在A,B的特征值λ_A和μ_B使得$\lambda_A-\lambda=\mu_B$, 即$\lambda=\lambda_A-\mu_B$, 这说明T的特征值可以表示为A,B的特征值的差.

反之, 对于A,B的特征值λ_A和μ_B, 矩阵方程

$$AX-XB=(\lambda_A-\mu_B)X$$

一定有非零解, 这是因为$A-\lambda_A I_n$和$B-\mu_B I_n$有共同的特征值0. 所以$\lambda_A-\mu_B$是T的特征值. 这就说明了A,B的特征值两两之差都是T的特征值.

(c)只要计算T的特征值的代数重数和几何重数, 再加以比较即可. 仍然同上设A,B的Jordan块分别是

$$J_{k_1}(\lambda_1), J_{k_2}(\lambda_2), \cdots, J_{k_s}(\lambda_s)$$
$$J_{m_1}(\mu_1), J_{m_2}(\mu_2), \cdots, J_{m_t}(\mu_t)$$

T的对应于特征值$\lambda_i-\mu_j$的特征子空间的几何重数是方程

$$AX-XB=(\lambda_i-\mu_j)X$$

的解空间的维数, 是$\min\{k_i,m_j\}$. (注意如果$\lambda_i=\lambda_{i'}$和$\mu_j=\mu_{j'}$的话, 我们是把$\lambda_i-\mu_j$和$\lambda_{i'}-\mu_{j'}$对应的特征子空间分开算.) 则T的所有特征值的几何重数的和为

$$\sum_{i=1}^{s}\sum_{j=1}^{t}\min\{k_i,m_j\}$$

所有代数重数的和为

$$n^2=\sum_{i=1}^{s}k_i\sum_{j=1}^{t}m_j=\sum_{i=1}^{s}\sum_{j=1}^{t}k_i\cdot m_j$$

所以T可对角化$\Leftrightarrow$这两个和相等$\Leftrightarrow$对一切i,j有$k_im_j=\min\{k_i,m_j\}\Leftrightarrow k_i=m_j=1\Leftrightarrow$ A,BJordan块都是1阶的$\Leftrightarrow$ A,B都可以对角化.

1.3 秩为1的方阵的总结

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$rank(A)=1$, 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方阵的特征值: 其中$n-1$个是0, 还有一个是trA.

证明: 显然齐次线性方程组$AX=0$的解空间是$n-1$维的, 也就是特征值0有$n-1$个线性无关的特征向量(几何重数是$n-1$). 从而0至少是A的$n-1$重特征值(代数重数大于等于几何重数). 由于A的所有特征值的和是trA, 所以剩下一个特征值必然是trA.
特别注意$trA=x_1y_1+x_2y_2+\cdots+x_ny_n=<X,Y>=Y'X$.
命题3 A的极小多项式是$x^2-tr(A)x$.
证明:

$$A^2-trA\cdot A=XY'XY'-trA\cdot XY'=X(Y'X)Y'-trA\cdot XY'=trA(XY'-XY')=0$$

但极小多项式不能是x(这个显然), 也不能是$x-trA$, 否则要么$trA\neq 0$时$A=trA\cdot I$, A可逆, 或者$trA=0$时$A=0$, 都矛盾.
这个结论也可以用牛刀做: Jordan标准形只有两种可能

$$J=\begin{pmatrix} trA & & & \\ & 0 & & \\ & & \ddots & \\ & & & 0\end{pmatrix}\quad 或者\quad J=\begin{pmatrix} 0 & 1 & & \\ 0 & 0 & & \\ & & \ddots & \\ & & & 0\end{pmatrix}$$

命题4 A相似于对角形当且仅当trA非零, 也就是$<X,Y>\neq 0$.
证明: 这个时候极小多项式无重根.

1.3.1　例题

例题 1: 设X,Y是$n\times 1$列向量, 证明$\det|I_n+XY'|=1+Y'X$.
解: XY'的特征值有$n-1$个是0, 还有一个是$trXY'=Y'X$, 所以I_n+XY'的特征值有$n-1$个是1, 一个是$1+Y'X$, 所以$\det|I+XY'|=1+Y'X$

例题 2: 设$f(\alpha,\beta)$是一个双线性函数, 求证$f(\alpha,\beta)$可以分解为两个非退化线性函数的乘积

$$f(\alpha,\beta)=g(\alpha)h(\beta)$$

的充要条件是f的秩是1.
解: 如果f的秩是1, 那么f在一组基下的矩阵$A=XY'$, 那么

$$f(\alpha,\beta)=\beta'XY'\alpha=(X'\beta)(Y'\alpha)$$

由于X,Y都是非零向量, 所以非退化. 反过来用前面的命题1必要性就不难得证.

例题 3: (武汉大学2005)设α是一个$\mathbb{R}^n$中的向量, $A=I_n-\alpha\alpha'$. 求证α的模长为1的充要条件是$rank(A)<n$.
解: $rankA<n$等价于说1是方阵$\alpha\alpha'$的特征值. 而$\alpha\alpha'$是秩1方阵, 所以也就等价于$tr\alpha\alpha'=\alpha'\alpha=1$. 得证.

1.3.2 伴随矩阵是原矩阵的多项式

证明n阶矩阵A的伴随矩阵A^*是A的多项式.
讨论:
(1)$rank(A) < n-1$, 这时$A^* = 0$, 结论显然成立.
(2)$rank(A) = n$, 这时由Hamilton-Caley定理,

$$A_n + a_{n-1}A^{n-1} + \cdots + a_1 A + a_0 = 0$$

由于A可逆, 所以$a_0 = (-1)^n|A| \neq 0$. 那么

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

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

从而A^{-1}可表为A的多项式, 从而A^*也能.
(3)$rank(A) = n-1$. 分三步走:
第一步: 论证向量空间

$$M = \{B \in M_n(F) \mid AB = BA = 0\}$$

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中的任何元素B都可以写为$B = c\alpha\beta'$, 从而M是一维的.
第二步: 找一个非零的矩阵, 它是A的多项式, 且在M中. 设A的极小多项式

$$m(x) = xg(x)$$

那么根据

$$Ag(A) = g(A)A = 0$$

从而$g(A) \in M$, 但$g(A)$不能是0, 否则与极小多项式矛盾.
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.

1.4 A, B同时上三角化

1.4.1 A, B交换的情形

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, 从而有特征向量的缘故.
先来证(b):
如果A,B都相似于对角形, 那么有特征子空间分解

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

$$V=N_1\oplus N_2\oplus\cdots\oplus N_s$$

其中M_i是A的属于特征值λ_i的特征子空间, N_j是B的属于特征值μ_j的特征子空间. 那么

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

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.
再来证(a):
用归纳法对空间的维数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$\alpha$扩充为$V$的一组基$\alpha,\beta_1,\cdots,\beta_{n-1}$. 在这组基下, A和B分别形如

$$A_1=\begin{pmatrix}\lambda & c\\ 0 & A_2\end{pmatrix}\quad 和\quad B_1=\begin{pmatrix}\mu & d\\ 0 & B_2\end{pmatrix}$$

只要证A_1,B_1同时相似于上三角阵. 由$AB=BA$有$A_1B_1=B_1A_1$, 从而$A_2B_2=B_2A_2$. 由归纳假设, A_2,B_2同时相似于上三角阵. 设P使得$P^{-1}A_2P,P^{-1}B_2P$为上三角阵, 不难验证

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

可以将A_1,B_1同时上三角化, 从而(b)得证.
补充一句, $AB=0$的时候A,B也可以同时上三角化, 因为$KerA$是B的不变子空间, 再归纳即可.

1.4.2 rank$(AB-BA)=1$的情形

证明当$rank(AB-BA)=1$时, A,B仍然可以同时上三角化.
从前面的证明可以看到, 关键是证明A,B有共同的特征向量. 由于$rank(AB-BA)=1$, 所以齐次方程组

$$(AB-BA)X=0$$

的解空间X是$n-1$维的. 注意到如果A,B之一的某个特征子空间, 比如说A的关于特征值λ 的特征子空间A_λ, 完全落在X中, 那么对于$\alpha \in V_\lambda$, 有

$$AB\alpha = BA\alpha = \lambda B\alpha$$

从而$B\alpha$仍在V_λ之中, 当然也仍在X中. 继续

$$A(B^2\alpha) = AB(B\alpha) = BA(B\alpha) = \lambda B(B\alpha) = \lambda B^2\alpha$$

从而知道$B^2\alpha$仍然在V_λ中. 从而不难验证$\alpha, B\alpha, \cdots$都在A的特征子空间V_λ内, 那么$\{\alpha, B\alpha, \cdots\}$生成的循环子空间既是$A$的特征子空间, 又是$B$的不变子空间, B限制在其上必有特征向量, 即A,B有共同的特征向量.
但是如果X当中不能全部包含A或B的任何特征子空间, 我们断言A,B仍然有共同的特征向量. 否则A某个特征向量α不在X内, B某个特征向量β不在X内, 而且α与β线性无关, 那么

$$\{\alpha\}\bigoplus\{\beta\}\bigoplus X$$

是直和. 这个维数至少是$n+1$, 矛盾! 从而无论如何A,B都有共同的特征向量.
剩下的就是复制(a)的证明: 用归纳法, 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

$$A_1 = \begin{pmatrix} \lambda & c \\ 0 & A_2 \end{pmatrix} \quad \text{和} \quad B_1 = \begin{pmatrix} \mu & d \\ 0 & B_2 \end{pmatrix}$$

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

$$rank[\begin{pmatrix} \lambda & c \\ 0 & A_2 \end{pmatrix}\begin{pmatrix} \mu & d \\ 0 & B_2 \end{pmatrix} - \begin{pmatrix} \mu & d \\ 0 & B_2 \end{pmatrix}\begin{pmatrix} \lambda & c \\ 0 & A_2 \end{pmatrix}] = 1$$

从而$rank(A_2B_2 - B_2A_2) \leqslant 1$. 由归纳假设, A_2, B_2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_1, B_1同时上三角化, 从而(c)得证.

1.5 幂等与幂零

1.5.1 正交幂等元分解

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)rank(A_1)+rank(A_2)+$\cdots$+rank$(A_m)=n$;
(3)对于$i\neq j$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

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

(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rank(A_1)+rank(A_2)+$\cdots$+rank$(A_m)=n$当且仅当右下角的矩阵为0, 即A_i幂等而且互相正交.
(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.

1.5.2 幂零$\Leftrightarrow tr(A^k)=0$

幂零等价于所有的特征根都是0, 要验证这一点, 只要证明对所有$1\leqslant k\leqslant n$都有$\mathrm{tr}A^k=0$. 通常的证法是用Newton恒等式, 有一种更简单的:
如果A的特征值不全是0, 那么设A的全部**非零**的特征值是$\lambda_1,\cdots,\lambda_r(r\geqslant 1)$, 对应的重数是$n_1,\cdots,n_r$, 那么由$\mathrm{tr}A^k=0$得

$$n_1\lambda_1^k+n_2\lambda_2^k+\cdots+n_r\lambda_r^k=0\quad k=1,2,\cdots,r$$

把上面的式子看成以$n_1,\cdots,n_r$为变量的线性方程组, 行列式是一个Vandermonde行列式, 所以只有零解!
但是用Newton恒等式可以证明**迹函数**$trA,trA^2,\cdots,trA^n$**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,

$$s_k=\lambda_1^k+\lambda_2^k+\cdots+\lambda_n^k=trA^k$$

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:

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

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$trC^k=0$.

1.6 Gram矩阵, 循环矩阵和Hilbert矩阵

1.6.1 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.
Gram矩阵的几何意义是什么? 就是$\alpha_1,\alpha_2,\cdots,\alpha_s$张成的平行多面体的体积的平方.
这是因为一方面$|G_s|=|G_{s-1}|((\alpha_s,\alpha_s)-x'G_{s-1}x)$, 另一方面设$\beta=c_1\alpha_1+\cdots+c_{s-1}\alpha_{s-1}$为$\alpha_s$在$\mathrm{span}\{\alpha_1,\alpha_2,\cdots,\alpha_{s-1}\}$ 中的最佳逼近元, 那么有

$$(\alpha_s-\beta,\alpha_i)=0,\quad i=1,2,\cdots,s-1$$

即$c_i=(\alpha_s,\alpha_i)$. 从而α_s到$\mathrm{span}\{\alpha_1,\alpha_2,\cdots,\alpha_{s-1}\}$的距离平方为

$$\|\alpha_s-\beta\|^2=(\alpha_s-\beta,\alpha_s-\beta)=(\alpha_s,\alpha_s)-x'G_{s-1}x$$

那么乘以$\alpha_1,\alpha_2,\cdots,\alpha_{s-1}$的体积$|G_{s-1}|$就可以看出来结论.
显然

$$\begin{gathered}|G_1|=\|\alpha_1\|^2\\|G_2|=|G_1|\cdot\|\alpha_2-(\alpha_2,\alpha_1)\alpha_1\|^2\\|G_3|=|G_2|\cdot\|\alpha_3-(\alpha_3,\alpha_1)\alpha_1-(\alpha_3,\alpha_2)\alpha_2\|^2\\\cdots\\|G_s|=|G_{s-1}|\cdot\|\alpha_s-(\alpha_s,\alpha_1)\alpha_1-\cdots-(\alpha_s,\alpha_{s-1})\alpha_{s-1}\|^2\end{gathered}$$

由于$\alpha_1,\alpha_2-(\alpha_2,\alpha_1)\alpha_1,\cdots,\alpha_s-(\alpha_s,\alpha_1)\alpha_1-\cdots-(\alpha_s,\alpha_{s-1})\alpha_{s-1}$ 正是对$\alpha_1,\cdots,\alpha_s$作施密特正交化手续后得到的向量, 所以G_s的行列式就是把$\alpha_1,\cdots,\alpha_s$作施密特正交化手续后得到的向量的模长乘起来再平方. 这就是北大09年高代第8题要证明的结论.
G的行列式还可以这样求: 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$|G_s|=\|\beta_1\|^2\|\beta_2\|^2\cdots\|\beta_s\|^2$

1.6.2 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$\alpha$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$\pi$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$\pi$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1.3.2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$f(x) = a_0 + a_1 x + \cdots + a_{n-1}x^{n-1}$, 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.

1.6.3 Hilbert矩阵

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

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

容易验证A,B满足关系$AB-BA=E$.
在有限维空间V上是找不出一对变换A,B使得$AB-BA=E$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

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

下面来点分析的: 考虑内积空间$L^2[0,1]$以及其中的n个元素$f_1,f_2,\cdots,f_n$生成的子空间V, 对于$g\in L^2[0,1]$, 根据Gram矩阵的意义, g到V的距离d满足

$$d^2=\frac{\det|G(f_1,f_2,\cdots,f_n,g)|}{\det|G(f_1,f_2,\cdots,f_n)|}$$

这就立即可以得到Müntz定理的一个弱形式:
如果互不相同的正数列$\lambda_1<\lambda_2<\cdots<\lambda_n<\cdots$满足$\lambda_n\to\infty$且

$$\sum_{n=1}^{\infty}\frac{1}{\lambda_n}=\infty$$

那么

$$\{x^{\lambda_1},x^{\lambda_2},\cdots\}$$

就是$L^2[0,1]$的一个完备序列(在L^2范数下稠密).
证明: 由于多项式在$L^2[0,1]$中稠密, 只要证明对于非负整数p, x^p到$\{x^{\lambda_1},x^{\lambda_2},\cdots,x^{\lambda_n}\}$ 的距离d_n随着n的增大趋于0. 不妨截取$\{x^{\lambda_i}\}$中满足$\lambda_i>p$的后半截子列,那么根据前面的公式

$$d_n=\frac{1}{\sqrt{2p+1}}\frac{|p-\lambda_1|\cdots|p-\lambda_n|}{(p+\lambda_1+1)\cdots(p+\lambda_n+1)}=\frac{1}{\sqrt{2p+1}}\prod_{i=1}^{n}|1-\frac{2p+1}{p+\lambda_i+1}|$$

由于

$$\sum_{n=1}^{\infty}\frac{2p+1}{p+\lambda_i+1}=\infty$$

所以无穷乘积趋于0, 得证.
由于

$$\sum_{p\text{是素数}}^{\infty}\frac{1}{p}=\infty$$

立刻有**所有素数次幂多项式在$L^2[0,1]$中稠密.**

1.7 不能被有限个真子空间覆盖

定理: 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$V\subseteq V_1\cup V_2\cup\cdots\cup V_m$.

注意题中的V不一定是有限维的.
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, 其次如果两个不同的λ_1, 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$\alpha$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$\{\lambda\alpha+\beta\}$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.
对于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
$$\alpha_i = 1 + i\varepsilon_1 + i^2\varepsilon_2 + \cdots + i^n\varepsilon_n$$
由Vandermonde行列式的性质知道无穷向量序列$\alpha_1, \alpha_2, \cdots, \alpha_n, \cdots$中任何$n$个都是$V$的一组基, 然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.

1.7.1 例题

例题 1: (2005南开大学) 设V是数域F上的向量空间, 则存在一个V中的无穷向量序列$\{\alpha_n\}_{n=1}^{\infty}$使得$\{\alpha_n\}_{n=1}^{\infty}$中任何$n$个向量都是$V$的一组基.
证明: 递归构造. 先取一组基$\alpha_1, \alpha_2, \cdots, \alpha_n$, 任取其中$n-1$个张成的子空间$\{M_k\}$至多有$C_n^{n-1}$个, 所以可以取$\alpha_{n+1}$使得$\alpha_{n+1}$不在任何$M_k$内, 从而$\alpha_1, \alpha_2, \cdots, \alpha_n, \alpha_{n+1}$ 中的任何n个都是V的一组基. 任取其中$n-1$个张成的子空间N_k至多有C_{n+1}^{n-1}个, 所以可以取α_{n+2}使得α_{n+2}不在任何N_k内. 一直继续下去就可以得到满足要求的序列.

例题 2: (2007上海交大) 设$A_1, A_2, \cdots, A_n$是向量空间V上的n个互不相同的线性变换, 求证存在$v \in V$使得$A_1v, A_2v, \cdots, A_nv$互不相同.
证明: 定义
$$V_{ij} = \{v \in V \mid A_iv = A_jv\}, \quad 1 \leqslant i < j \leqslant n$$
那么V_{ij}都是子空间, 且由于$A_i \neq A_j$, V_{ij}还是真子空间. 所以根据定理v的存在是显然的了.

1.7.2 以极小多项式为阶的元素的存在

这个定理可以用来证明在一个向量空间V中必有一个向量v, 其极小多项式等于A的极小多项式.

证明: 根据熟知的结论, 任一元素v的极小多项式$m_v(x)$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, 对应的极小多项式只有有限多个$m_1(x), m_2(x), \cdots, m_k(x)$. 考虑

$$V_i = \{v \in V \mid m_i(A)v = 0\}$$

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$V = V_i$, 那么显然有$m(x) = m_i(x)$, 得证.

1.8 域的变化

这两个问题都可以用不变因子直接说明, 这里给出不依赖于不变因子概念的证明.

1.8.1 相似关系和域无关

A, B是数域$\mathbb{K}$上的两个$n \times n$矩阵, 如果它们在复数域上相似, 那么它们在数域$\mathbb{K}$上也相似.

证明: 设$T = (t_{ij})$是一个n阶矩阵, 把$AT - TB = 0$看成一个由n^2个变元t_{ij}组成的齐次线性方程组, A, B复相似说明这个方程组在复数域内有非零解, 从而在$\mathbb{K}$内也有非零解. 从而可设$T_1, T_2, \cdots, T_s$是一组基础解系, **注意$T_1, T_2, \cdots, T_s$也是该方程组在$\mathbb{C}$内的基础解系.** 这是因为在$\mathbb{K}$中求基础解系的过程与在$\mathbb{C}$中求基础解系的过程是一样的. 那么对任何$x_1, x_2, \cdots, x_s \in \mathbb{K}$, $x_1T_1 + x_2T_2 + \cdots + x_sT_s$仍是$\mathbb{K}$上的矩阵且满足该方程组. 所以只要找合适的$x_1, x_2, \cdots, x_s \in \mathbb{K}$使得$x_1T_1 + x_2T_2 + \cdots + x_sT_s$可逆即可. 考虑函数$f(x_1, x_2, \cdots, x_s) = |x_1T_1 + x_2T_2 + \cdots + x_sT_s|$, 由于存在$a_1, a_2, \cdots, a_s \in \mathbb{C}$使得$f(a_1, a_2, \cdots, a_s) \neq 0$, 所以$f \neq 0$. 从而存在$x_1, x_2, \cdots, x_s \in \mathbb{K}$使得$f(x_1, x_2, \cdots, x_s) \neq 0$, 得证.

1.8.2 最小多项式和域无关

设矩阵A是数域F上的n阶矩阵, 则A在F上的极小多项式与A在$\mathbb{C}$上的极小多项式相同.

证明: 设A在F上的极小多项式为$\varphi(x)$, 在$\mathbb{C}$上的极小多项式为$p(x)$, 显然$p(x)|\varphi(x)$,
只要再证明二者次数相等.
设$p(x) = x^k + a_{k-1}x^{k-1} + \cdots + a_1x + a_0$, 那么根据$p(A) = 0$, 有

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

把$a_{k-1},\cdots,a_0$看成未知元, $A,\cdots,A^k$的各个分量看成已知的系数, 那么上式可以看作一个系数在F上的, 有k个未知元的, 有n^2个方程的齐次方程组, 它在复数域内有解, 那么在F内也有解, 也就是存在F中的数$b_{k-1},\cdots,b_0$ 使得

$$A^k+b_{k-1}A^{k-1}+\cdots+b_1A+b_0=0.$$

然而A在F上的极小多项式$\varphi(x)$应该整除多项式$x^k+b_{k-1}x^{k-1}+\cdots+b_1x+b_0$, 从而其次数不大于$p(x)$的次数$k$, 从而必有$\varphi(x)=p(x)$.

1.9 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. 关键的想法加了黑体以示强调.

1.9.1 变换的角度

如果一个与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? 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

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

是V关于A的不变子空间直和分解, 那么它其实也是关于B的不变子空间直和分解. 换句话说, 一个与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

如果

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

是V关于A的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A_i表示A在V_i上的限制.

若不然, 当k不等于i时定义B在所有的V_k上都是0, 在V_i就是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1.2讨论过的:

引理: 设A_1,A_2分别是复数域上n维向量空间V_1和m维向量空间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.
由于A总是可以分解为循环不变子空间的直和, 每个子空间下呈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:
A是V上的循环变换.

现在我们已经找到了A应该满足的必要条件: A是一个循环变换. 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, 那么**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$m = f(A)v$. 设$g(x)$使得$Bv = g(A)v$. 下面证明$B = g(A)$: 对$V$中任一向量$m$, 设$m = f(A)v$, 那么

$$Bm = Bf(A)v = f(A)Bv = f(A)g(A)v = g(A)f(A)v = g(A)m$$

即对V中任一向量m有$Bm = g(A)m$. 从而必有$B = g(A)$.
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A是一个循环变换.

1.9.2 矩阵的角度

我们要算一算, 与矩阵A交换的矩阵到底有哪些? 首先明确一点, 所有与A交换的矩阵构成一个向量空间, 下面来计算这个空间的维数. 取合适的基以后, 不妨假设A呈Jordan标准形, 即

$$A = \begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_r \end{pmatrix}$$

对于与A交换的矩阵B, 按照A的方式将B分成r^2个子块:

$$B=\begin{pmatrix} B_{11} & B_{12} & \cdots & B_{1r} \\ B_{21} & B_{22} & \cdots & B_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ B_{r1} & B_{r2} & \cdots & B_{rr} \end{pmatrix}$$

于是从$AB=BA$得到r^2个方程

$$J_iB_{ij}=B_{ij}J_j \quad 1\leqslant i,j\leqslant r$$

当J_i对应的特征值λ_i与J_j对应的特征值λ_j不相等的时候, 根据前面的引理1, $B_{ij}=0$. 而当$\lambda_i=\lambda_j$时,

$$N_iB_{ij}=B_{ij}N_j$$

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B_{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, B_{ij}含有$\min\{B_{ij}$的行数,B_{ij}的列数$\}$个参数. 也就是

$$\min\{J_i\text{的阶数},J_j\text{的阶数}\}$$

个参数, 这个数还是J_i, J_j对应的初等因子的最大公因子的次数. 而整个B的参数个数是所有B_{ij}参数个数的和. 如果设$\lambda_1, \cdots, \lambda_s$是$A$的全部互不相同的特征值, 而$f_1(\lambda), \cdots, f_t(\lambda)$是$A$的不变因子, 按照次数的高低从大到小排成$t$行

$$f_1(\lambda) = (\lambda - \lambda_1)^{e_{11}}(\lambda - \lambda_2)^{e_{12}} \cdots (\lambda - \lambda_s)^{e_{1s}}$$

$$f_2(\lambda) = (\lambda - \lambda_1)^{e_{21}}(\lambda - \lambda_2)^{e_{22}} \cdots (\lambda - \lambda_s)^{e_{2s}}$$

$$\cdots\cdots$$

$$f_t(\lambda) = (\lambda - \lambda_1)^{e_{t1}}(\lambda - \lambda_2)^{e_{t2}} \cdots (\lambda - \lambda_s)^{e_{ts}}$$

在这里假定所有的$e_{ij} \geqslant 0$. 此外还假设$f_1, f_2, \cdots, f_t$次数分别是$n_1, n_2 \cdots, n_t$, 那么B的参数个数就是首先对每一列j, 两两求出第i行和第i'行对应的初等因子$(\lambda - \lambda_j)^{e_{ij}}$和$(\lambda - \lambda_j)^{e_{i'j}}$(有零次多项式没有关系) 的最大公因子的次数, 全部加起来. 然后再将各列所得的结果相加. **注意到这一手续可以对各列同时进行, 也就是可以对各行两两求最大公因子的次数, 然后相加, 所得结果与原来的结果相同.** 所以我们得到B的参数个数为

$$\sum_{1 \leqslant i \leqslant j \leqslant t} \deg(f_i(x), f_j(x)) = \sum_{1 \leqslant i \leqslant j \leqslant t} n_j = n_1 + 3n_2 + \cdots + (2t-1)n_t$$

这样我们就算出了与A交换的矩阵组成的空间的维数:

命题2: 与A交换的矩阵构成的向量空间的维数是

$$n_1 + 3n_2 + \cdots + (2t-1)n_t$$

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.

注意由于不变因子是不依赖于域的, 所以与A交换的矩阵构成的子空间维数也不依赖于域.

注意到总是有

$$n_1 + 3n_2 + \cdots + (2t-1)n_t \geqslant n_1 + n_2 + \cdots + n_t = n$$

等号成立当且仅当$t = 1$, 即A仅有一个次数大于0的不变因子, 也就是极小多项式等于它的特征多项式, 而这时A的多项式在$M_n(\mathbb{C})$中生成的空间也是n维的, 以$I, A, A^2, \cdots, A^{n-1}$为一组基, 所以这就说明

$$A\text{的多项式生成的空间} = \text{与}A\text{交换的矩阵构成的空间}$$

从而从矩阵的角度再次证明了当A循环时与A交换的矩阵总能表示为A的多项式.

1.9.3 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: 注意到题中的条件比前面讨论的条件更强, C不仅仅只和A交换, 而且A也不见得是循环的. 那么这个强体现在什么地方呢? 回顾前面的讨论, **有一些与A交换的变换可以建立A的同一特征值的不同Jordan块之间的映射, 这种变换一定不能表示为A的多项式, 关键就在于此!** 我们说C一定不属于这种变换, 那就是

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

$$AP = PA$$

那么根据已知就会有

$$CP = PC$$

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, 因为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, 其中$\{v_1\}$对应的不变因子是A的极小多项式$f_1(x)$, 此外还有$f_r(x)|\cdots|f_2(x)|f_1(x)$. 我们有

引理3: 存在与A交换的变换B使得$Bv_1 = v_2$.
证明: 这个比较好想, 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X使得$Xv_1 = v_2$而且X与A交换: $XA = AX$, 那么就有X与C交换: $CX = XC$. 然而

$$XCv_1 = Xg_1(A)v_1 = g_1(A)Xv_1 = g_1(A)v_2$$

$$CXv_1 = Cv_2$$

所以

$$Cv_2 = g_1(A)v_2 = g_2(A)v_2$$

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.

1.9.4 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:

$$\begin{array}{ccc} u & \longrightarrow & Au \\ f\downarrow & & f\downarrow \\ v & \longrightarrow & Bv \end{array}$$

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

$$\begin{array}{ccc} v & \longrightarrow & Av \\ \pi\downarrow & & \pi\downarrow \\ \bar{v} & \longrightarrow & \bar{A}v \end{array}$$

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$U/Kerf$上的诱导变换与B在Imf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$\mathrm{End}V$的一个子环, 所以可以考虑R在$\mathrm{End}V$内的中心化子

$$D = Hom_R(V,V) = \{B \in \mathrm{End}V : AB = BA\}$$

以及D在$\mathrm{End}V$中的中心化子

$$E = Hom_D(V,V) = \{C \in \mathrm{End}V : CB = BC, \forall B \in D\}$$

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是循环的, 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. 特别注意对于一般的非交换环上的模, 这几个结论可能都不成立. 双中心化子性质在Wedderburn-Artin的半单代数结构定理中扮演了关键角色.

1.10 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. 直观上看, $f(J)$会碎裂成一些更小的Jordan块的和. 下面来讨论这一问题:
设

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

由于Hamilton-Caley定理, 对任何次数大于等于r的多项式$f(x)$,$f(J)$总是$I,J,\cdots,J^{r-1}$的线性组合, 所以只要对次数小于r的多项式定出标准形来.

设$f(x)=a_0+a_1(x-\lambda_0)+\cdots+a_{r-1}(x-\lambda_0)^{r-1}$, 那么

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

(这里有一个小技巧, 设f是一个多项式, 计算$f(J)$的时候不要直接把J带进去硬算, 而是先把$f(x)$在λ_0点作Taylor 展开, 这样直接就能写出$f(J)$来. 还有一个常见的问题是当矩阵A与J交换的时候证明A是J的多项式. 在前面已经看到这样的A必定是上面的上三角分层矩阵, 所以$f(x)$直接就写出来了.)

其$\lambda-$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$(\lambda-a_0)^k, 0\leqslant k\leqslant r$. 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$(r-1)!$项中除了对角线乘积$(-a_1)^{r-1}$以外都含有$\lambda-a_0$, 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$B=f(J)-a_0I, r_i=rankB^i, i=0,1,\cdots,r-1$. **注意$B^i$这种严格”分层“上三角矩阵的秩由最靠近主对角线的非零次对角线决定, 与更高的次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. 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.
其实这一段挺深刻的. 线性代数本质上可以说是研究$F[x]$的有限维表示在相似意义下的分类问题, 最后的结论是Jordan标准形: $\mathbb{C}[x]$的任何有限维表示都是若干个循环的不可分解模的直和. 很有意思, 上面的结果是说当这个表示限制在子代数$\mathbb{C}[f(x)]$上的时候每个不可分解模又继续"碎裂"了, 而且碎的很规则. 这和域的扩张是对偶的, 当基域变大的时候, $F[x]$的每个不可分解模也会变小, 当基域变大成$\mathbb{C}$的时候就稳定下来成为Jordan块了. 另一方面从某个k以后A^k和A具有同样的Jordan形这说明当子代数缩小到一定程度的时候Jordan形也稳定下来了. 能够使得一个代数的所有不可分解模稳定下来的域就叫做分裂域. 复数域是所有多项式的分裂域, 也是其所有不可分解模的分裂域.

2 几个专题

2.1 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

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

证法1:

$$\begin{pmatrix} AB & 0 \\ B & BC \end{pmatrix} \xrightarrow{\text{第二行左乘以}-A\text{加到第一行}} \begin{pmatrix} 0 & -ABC \\ B & BC \end{pmatrix} \xrightarrow{\text{第一列右乘以}-C\text{加到第二列}} \begin{pmatrix} 0 & -ABC \\ B & 0 \end{pmatrix}$$

比较两端的大的矩阵秩就得到结论.
证法2: 把A,B,C都看成线性变换, 就有

$$\mathbb{F}^s \xrightarrow{C} \mathbb{F}^k \xrightarrow{B} \mathbb{F}^n \xrightarrow{A} \mathbb{F}^m$$

记$W=KerA\cap Im(B), W'=KerA\cap Im(BC)$, 那么$W'$是$W$的子空间. 现在来看$A$在$Im(B)$上的限制, 由同态基本定理,

$$Im(B)/W\cong Im(AB)$$

同理对A在$Im(BC)$上的限制有

$$Im(BC)/W'\cong Im(ABC)$$

计算维数就有

dim$W=r(B)-r(AB)$,dim$W'=r(BC)-r(ABC)$, 再由W'是W的子空间知道dim$W'\leqslant$dim$W$. 即得结论.

2.2 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: 给两种做法:

法1: 记$f(x)=x^3-1, g(x)=x-1, h(x)=x^2+x+1$, 那么$f(x)=g(x)h(x)$且$(g,h)=1$,

$$Kerf(A)=Kerg(A)\oplus Kerh(A)$$

于是

$$\begin{aligned}&A^3=I\\ &\Leftrightarrow f(A)=0\\ &\Leftrightarrow Kerf(A)=V\\ &\Leftrightarrow Kerg(A)\oplus Kerh(A)=V\\ &\Leftrightarrow n-r(g(A))+n-r(h(A))=n\\ &\Leftrightarrow r(I-A)+r(I+A+A^2)=n\end{aligned}$$

法2: 用前面的正交幂等元的技巧也能做. 虽然结论用不上, 但是方法可以照搬:

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

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

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

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

证明实际上是Sylvester秩不等式的证明方法的拷贝.

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

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$X, Y\subset Z$. 从而$X+Y\subset Z$. 而且$X\cap Y\subset W$. 由维数公式,

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

从而

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

即

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

2.3 n维欧式空间中两两夹钝角的向量个数

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$n+1$.

证明: 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$\alpha_1$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. 然而从$n=1$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$\beta$是$M^{\perp}$中的非零单位向量,取正数$\lambda$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.

2.4 两个半正定矩阵可以同时合同于对角形

(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.
证明: 以下约定A, B为半正定实矩阵, X, Y为$n \times 1$矩阵.
引理(1): 如果A半正定, 那么$X'AX = 0$等价于$AX = 0$.
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.
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.

下面进入原命题的证明: 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, 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.

2.5 实Jordan块的开平方问题

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. 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. 在下面的讨论中将反复用到两个结论:

结论(1): (1.8.1)实矩阵复相似等价于实相似.

结论(2): 如果实矩阵A复相似于一个实矩阵B的平方, 则A可以在实数域内开平方.

(2)的证明: 根据(1)可知存在实可逆矩阵T使得$T^{-1}B^2T = A$, 即$(T^{-1}BT)^2 = A$, 从而A可以在实数域内开平方.

命题1 当$a > 0$时J可以在实数域内开平方.

证明: 考察

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

根据1.10的结论, B^2的Jordan标准形就是J, 所以B^2与J相似. 从而根据结论(2)J可以开平方.

这个问题还有一个解法: 记$J = aI + N$, 则$N^n = 0$. 考虑$\sqrt{x+a}$的n次Taylor多项式$P(x)$,

则$P(x)^2-(x+a)=x^mQ(x)$, 这里$Q(x)$也是一个多项式. 我们还有$P(x),Q(x)$都是实系数的. 那么

$$P(N)^2=N+aI+N^nQ(N)=N+aI=J$$

所以$B=P(N)$满足要求.
这个证法实际上给出了B的计算方法.
命题2 当特征值为$-a(a>0)$时J不能在实数域内开平方.
证明: 若不然, 设实矩阵$B^2=A$, 那么B的特征值为$\sqrt{a}i$或$-\sqrt{a}i$, 由于实矩阵的复特征根成对出现, 所以必然一半是$\sqrt{a}i$, 一半是$-\sqrt{a}i$. 所以B的Jordan标准形至少含有两个子块. 但是根据1.10 的结论, B^2的Jordan标准形与B的标准形形状完全一样, 仅仅对角线上的元素有所不同, 所以B^2的标准形当中也应该至少有两个子块, 这就导致了矛盾.
命题3 当a等于0时如果A的阶数大于1则A不能在实数域内开平方.
还是用1.10的结论, B的特征值也都是0, 而且一定有阶数大于1的子块(否则$B=0$, 不可能.) **而0对应的Jordan块在平方以后会分解为阶数相等或者相邻的两个子块**, 这就导致了矛盾.

这个命题也可以不用1.10的结论做, 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$B^n=0$, 从而$n\geqslant 2$时有$B^{2n-2}=B^nB^{n-2}=0$, 矛盾!

下面讨论何时一个特征根都是实数的方阵A可以在实数域内开平方. 根据结论(2), 这只要考虑A的Jordan标准形何时可以在实数域内开平方.
先看必要性, 看看要使得A能开平方的话A应该满足什么条件.
必要性之一 A的任一负特征值$-a(a>0)$对应的任一阶Jordan块必定出现偶数次(成对出现).
证明: 如果$B^2=A$的话考察B的标准形: $\sqrt{a}i$对应的i阶Jordan块的个数是

$$rank(B-\sqrt{a}iI)^{i-1}+rank(B-\sqrt{a}iI)^{i+1}-2rank(B-\sqrt{a}iI)^i$$

由于把一个矩阵的元素都取复共轭以后不改变矩阵的秩, 所以这个个数还等于

$$rank(B+\sqrt{a}iI)^{i-1}+rank(B+\sqrt{a}iI)^{i+1}-2rank(B+\sqrt{a}iI)^i$$

所以B的特征值为$\sqrt{a}i$的Jordan块和特征值为$-\sqrt{a}i$的Jordan块是一一对应的, 平方以后就是A的特征值为$-a$的Jordan块成对出现. (再一次用到1.10最后的推论.)
必要性之二 A的0特征值对应的Jordan块可以两两组对为

$$(J_1,J_2),\ (J_3,J_4),\ \cdots,(J_{2k-1},J_{2k}),\ 0,\cdots,0$$

使得每一对(J_{2i-1},J_{2i})中J_{2i-1}和J_{2i}的阶数要么相等, 要么相邻. 这里J_{2i-1},J_{2i}中至少有一个的阶要大于1.
证明: 1.10最后的推论说的就是这个.
我们说以上两个必要性还是充分的, 也就是满足上面两个必要条件的A一定可以在实数域内

开平方:
由于命题1, 只需要考虑划去正特征值对应的Jordan块后剩下的部分的开平方问题. 把0特征值对应的Jordan块和负特征值对应的Jordan块分开来考虑: 如果0对应的Jordan块有划分

$$(J_1, J_2),\ (J_3, J_4),\ \cdots, (J_{2k-1}, J_{2k}),\ 0, \cdots, 0$$

不妨取出(J_1, J_2)来, 如果J_1, J_2的阶都是r, 那么考虑$2r \times 2r$的方阵

$$A = \begin{pmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix}_{2r\times 2r}$$

它平方以后分解为2个阶为r的子块, 也就是分解为J_1和J_2的和. 从而$\mathrm{diag}\{J_1, J_2\}$可以开平方. 如果$J_1, J_2$的阶一个是$r$, 一个是$r+1$的话情况类似. 所以0特征值对应的Jordan块整体$\mathrm{diag}\{J_1, J_2, \cdots, J_{2k-1}, J_{2k}, 0, \cdots, 0\}$可以开平方. 下面只要处理负特征值对应的Jordan块开平方的问题. 由于负特征值对应的Jordan块总是成对出现, 所以只要证明矩阵

$$C = \begin{pmatrix} J & 0 \\ 0 & J \end{pmatrix} \quad 其中 \quad J = \begin{pmatrix} -a & 1 & & \\ & -a & \ddots & \\ & & \ddots & 1 \\ & & & -a \end{pmatrix}$$

可以在实数域内开平方. 考察矩阵

$$B = \begin{pmatrix} J_1 & 0 \\ 0 & J_2 \end{pmatrix} \quad J_1 = \begin{pmatrix} \sqrt{a}i & 1 & & \\ & \sqrt{a}i & \ddots & \\ & & \ddots & 1 \\ & & & \sqrt{a}i \end{pmatrix} \quad J_2 = \begin{pmatrix} -\sqrt{a}i & 1 & & \\ & -\sqrt{a}i & \ddots & \\ & & \ddots & 1 \\ & & & -\sqrt{a}i \end{pmatrix}$$

那么B^2复相似于C, 如果能够证明B复相似于一个实矩阵D, 那么C复相似于实矩阵D^2, 从而C实相似于D^2, C可以开平方, 从而负特征值对应的Jordan块整体可以开平方.
B复相似于一个实矩阵的证明: 记

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

则U是酉矩阵: $UU^* = I$. 直接计算验证

$$U^{-1} \begin{pmatrix} \sqrt{a}iI & 0 \\ 0 & -\sqrt{a}iI \end{pmatrix} U = \begin{pmatrix} 0 & \sqrt{a}I \\ -\sqrt{a}I & 0 \end{pmatrix}$$

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

$$U^{-1}BU=\begin{pmatrix} N & \sqrt{a}I \\ -\sqrt{a}I & N \end{pmatrix}$$

从而B复相似于一个实矩阵.

至此完成了特征值都是实数的实矩阵在实数域内开平方的全部讨论.

3 问题集

问题 1: (Hamilton-Caley 定理) 设A是V上的线性变换, f是A的特征多项式, 那么$f(A)=0$.

问题 2: 若数域$\mathbb{K}$上的方阵A的迹$trA=0$, 求证A相似于一个对角线上都是0的方阵.

问题 3: 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 4: 设A,B是两个n阶方阵, $AB=BA$, 且A是幂零的, 求证$|A+B|=|B|$.

问题 5: 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 6: 对怎样的正整数n, 存在有理数域$\mathbb{Q}$上的$n\times n$矩阵A满足$A^4-A^3+2A+1=0$?

问题 7: 设A,B是n阶正交矩阵且$|A|=-|B|$, 求证$|A+B|=0$.

问题 8: 设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 9: 设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 10: 设n阶方阵A满足$A+A'=I$, 求证A可逆.

问题 11: 求证不存在正交矩阵A,B满足$A^2=AB+B^2$.

问题 12: 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$tr(A)=tr(B)$.

问题 13: 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 14: 设$A,B\in M_n(\mathbb{R})$且A,B对称. 求证

$$tr[(AB)^2]\leqslant tr(A^2B^2)$$

.

问题 15: (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 16: 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$B=AO$.

问题 17: 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 18: 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 19: 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

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

问题 20: 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 21: 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 22: 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, $r+s\leqslant n$.

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: 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 25: 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 26: A,B都是n阶半正定矩阵, 求证AB的特征值都是实数.

问题 27: 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 28: 设A,B是n阶方阵且B可逆, $r(I-AB)+r(I+BA)=n$, 求证A也可逆.

问题 29: 设A,B是两个n阶矩阵满足$r(A+B)=r(A)+r(B)$, 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 30: (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.

4 问题集解答

解答 1: 对空间V的维数n归纳, $n = 1$显然. 设$n-1$的时候成立, 取一个特征向量$A\alpha = \lambda\alpha$, 将α扩充为V的一组基, 在这组基下A的矩阵形如

$$\begin{pmatrix} \lambda & C \\ 0 & A_1 \end{pmatrix}$$

其中A_1是A在商空间$\bar{V} = V/\{\alpha\}$上的诱导变换对应的矩阵. 设A_1的特征多项式为$g(x)$, 那么$f(x) = (x-\lambda)g(x)$. 根据归纳假设, $g(A_1)\bar{V} = 0$, 即$g(A_1)(v + \{\alpha\}) = g(A)v + \{\alpha\} = 0$. 这说明对任何向量$v \in V, g(A)v \in \{\alpha\}$, 然而注意到$(A - \lambda I)\alpha = 0$, 所以对任何$v \in V$,

$$f(A)v = (A - \lambda I)g(A)v \subset (A - \lambda I)\{\alpha\} = 0$$

得证.

解答 2: 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. 注意这个结论不依赖于基域.

解答 3: $\Rightarrow$:

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

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 4: 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

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

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 5: $n = 1, 2, 3$的时候没有, $n \geqslant 4$的时候一定有. 这是因为$f(x) = x^4 - x^3 + 2x + 1$是有理数域上的不可约多项式(作代换$y = x - 1$后用爱森斯坦判别法). 所以$n \leqslant 3$的时候$A$的极小多项式次数也小于等于3, 又要整除$f(x)$, 不可能. $n \geqslant 4$的时候用Frobenius标准形不难构造所要求的矩阵.

解答 6: 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 7: 由于$AB = BA$, 所以A, B可以同时上三角化; 再由于A是幂零的, 所以A的特征值全是0, 从而$|A + B| = |B|$.

解答 8: 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 9: 首先用纯形式的推导找出这个逆来. 把矩阵看成数:

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

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 10: 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 11: 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:

$$tr(AB') = n + tr(A'B)$$

但是

$$trAB' = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}b_{ij} = trA'B$$

(或者$trAB' = trB'A = tr(B'A)' = trA'B$)
从而$0=n$, 矛盾!

解答 12: 只要注意到$(I-A-B)A = B(I-A-B)$ 就可得到A,B相似.

解答 13: 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$tr(A-A')(A-A')' \geqslant 0$且等号成立当且仅当$A=A'$.

解答 14: 令$C=AB-BA$, 则$trCC'\geqslant 0$, 展开整理:

$$tr(ABBA+BAAB)\geqslant tr(ABAB+BABA)$$

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

$$tr(ABBA+BAAB)=2tr(A^2B^2)$$

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

解答 15: 用1.10中讲过的计算技巧, 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 16: 设A的列向量为$A=(\alpha_1,\alpha_2,\cdots,\alpha_n)$, B的列向量为$B=(\beta_1,\beta_2,\cdots,\beta_n)$. 那么$A'A=B'B$说明$(\alpha_i,\alpha_j)=(\beta_i,\beta_j)$. 所以问题的本质暴露出来了:
设$\alpha_1,\alpha_2,\cdots,\alpha_n,\beta_1,\beta_2,\cdots,\beta_n$是两组列向量满足$(\alpha_i,\alpha_j)=(\beta_i,\beta_j)$, 要证明存在正交变换$O$使得$O\alpha_i=\beta_i$. 不妨设$\alpha_1,\alpha_2,\cdots,\alpha_r$是$\alpha_1,\alpha_2,\cdots,\alpha_n$中的线性极大无关组, 那必有$\beta_1,\beta_2,\cdots,\beta_r$是$\beta_1,\beta_2,\cdots,\beta_n$中的线性极大无关组. 记$M=\{\alpha_1,\alpha_2,\cdots,\alpha_r\}$,$N=\{\beta_1,\beta_2,\cdots,\beta_r\}$. 取$M^\perp$和$N^\perp$中的标准正交基$\{x_i\}$和$\{y_j\}$将它们分别扩充为$\mathbb{R}^n$的一组基:

$$\alpha_1,\alpha_2,\cdots,\alpha_r,x_{r+1},\cdots,x_n$$

$$\beta_1,\beta_2,\cdots,\beta_r,y_{r+1},\cdots,y_n$$

定义O为$O\alpha_i=\beta_i, 1\leqslant i\leqslant r, Ox_i=y_i, r+1\leqslant i\leqslant n$. 那么$O$是一个保距变换, 从而是正交变换, 这就证明了结论.

解答 17: 考察正交矩阵$C=AB^{-1}$, 那么$r(A+B)=r(C+I)$, 而且$n-r(C+I)$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 18: 我们有

$$A(A-B)=-(A-B)B$$

如果$A-B$不是0的话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. 甚至条件改为“A,B的特征值的凸包不包含原点” 时结论也成立.

解答 19: 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:

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

从而$r(A+B)\geqslant r(B)+r(A^3)=r(B)+r(A^2)$.

另证: 首先$r(A)=r(A^2)$说明

$$V=KerA\oplus ImA$$

其次不难证明有$Ker(A+B)=KerA\cap KerB, ImA\subset KerB$. 从而

$$KerB=(KerB\cap KerA)\oplus(KerB\cap ImA)=Ker(A+B)\oplus ImA$$

两边计算维数就得结论.

解答 20: 设n维向量空间V上的线性变换A在基$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$下的矩阵为$A$. 那么

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

令$M_i = L(\varepsilon_i, \varepsilon_{n-i+1})$, 则$M_i$为$A$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A可以对角化当且仅当A在每一个M_i上的限制$A|_{M_i}$可以对角化. $A|_{M_i}$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 21: 首先存在一组标准正交基使得A在这组基下的矩阵形如

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

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

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

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

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 22: 周老师的巧妙证明:

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.

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

$$A^k=\begin{pmatrix}0 & A_1B_1^{k-1}\\ 0 & B_1^k\end{pmatrix}$$

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

$$Ap(A)=\begin{pmatrix}0 & A_1p(B_1)\\ 0 & B_1p(B_1)\end{pmatrix}=0$$

所以A的极小多项式的次数至多是$r+1$. 问题得证.

解答 24: 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 25: 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 26: 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 27: 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1.9的命题1的结论即得所证.

解答 28: 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 29: 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 30: 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$f = y_2 + 2y_3^2$. 反解出$y_1, y_2, y_3$来可得正交矩阵为

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

5 后记