线性代数习题

问题1

如果U,V,W是线性空间L的子空间

$(U+V)\cap(U+W)\subset U+(U+V)\cap W$：假设$x=u+v=u'+w'\in(U+V)\cap(U+W)$，这里$u.u'\in U,v\in V,w'\in W$那么$u-u'+v=w\in(U+V)\cap W$，所以$x=u'+(u-u'+v)\in U+(U+V)\cap W$。

$(U+V)\cap(U+W)\supset U+(U+V)\cap W$：由于$U\subset(U+V),(U+W)$有$U\subset(U+V)\cap(U+W)$，$(U+V)\cap W\subset W\subset(U+W)$，$(U+V)\cap W\subset(U+V)$同理$(U+V)\cap W\subset(U+V)\cap(U+W)$，后者是子空间那么

$$U+(U+V)\cap W\subset(U+V)\cap(U+W)$$

问题2(Frobenius不等式)

考虑$C:F^q\to F^p,B:F^p\to F^n,A:F^m\to F^n$是三个矩阵，考虑其作为线性映射的如下关系：

$$\begin{aligned}\mathrm{Im}(AB)&=A(\mathrm{Im}B/\mathrm{Im}B\cap\ker(A))\\\mathrm{Im}(BC)&=B(\mathrm{Im}C/\mathrm{Im}C\cap\ker(B))\\\mathrm{Im}(ABC)&=A(\mathrm{Im}(BC)/\mathrm{Im}(BC)\cap\ker(A))\end{aligned}$$

注意到A在$\mathrm{Im}(BC)/\mathrm{Im}(BC)\cap\ker(A),\mathrm{Im}B/\mathrm{Im}B\cap\ker(A)$是单射，从而取维数有：

$$\mathbf{rank}(ABC)+\mathbf{rank}(B)\geq\mathbf{rank}BC-(m-\mathbf{rank}(A))+\mathbf{rank}(B)=\mathbf{rank}BC+(\mathbf{rank}(A)+\mathbf{rank}(B)-m)\geq\mathbf{rank}BC+\mathbf{rank}A$$

也就是

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

推论，如果$\mathbf{rank}(B)=\mathbf{rank}AB$，那么$\mathbf{rank}BC\leq\mathbf{rank}ABC\leq\mathbf{rank}BC$，从而$\mathbf{rank}BC=\mathbf{rank}ABC$。

问题3

如果$P_1,..,P_k,Q_1,..,Q_k$满足Q_i和$P_1,..,P_k$交换，P_i和$Q_1,..,Q_k$交换而且$\mathbf{rank}P_i=\mathbf{rank}P_iQ_i$对任何的$1\leq i\leq k$成立，那么：

$$P_1\cdots P_kQ_1\cdots Q_k=P_1Q_1P_2Q_2\cdots P_kQ_k=Q_1P_1\cdots Q_kP_k$$

根据上面的推论也就是：

$$\begin{aligned}\mathbf{rank}(B)=\mathbf{rank}AB\quad&\Leftrightarrow\quad\mathbf{rank}BC=\mathbf{rank}ABC\\\mathbf{rank}(B)=\mathbf{rank}BC\quad&\Leftrightarrow\quad\mathbf{rank}AB=\mathbf{rank}ABC\end{aligned}$$

$$\mathbf{rank}P_1\cdots P_kQ_1\cdots Q_k=$$
$$\mathbf{rank}P_1Q_1\cdots P_{k-1}Q_{k-1}P_k=\mathbf{rank}P_1Q_1\cdots P_{k-1}Q_{k-1}P_k=\mathbf{rank}P_1Q_1\cdots P_{k-2}Q_{k-2}P_{k-1}P_k=\cdots=\mathbf{rank}P_1\cdots P_k$$

也就是

$$\mathbf{rank}P_1\cdots P_k=\mathbf{rank}P_1\cdots P_kQ_1\cdots Q_k$$

问题4

考虑$V=\mathrm{Span}\{e_1,..,e_n\}$是n-维线性空间，置换群$S_n$在$V$上有作用：$\{\pi(v)\}_i=v_{\pi(i)}$，考虑如下的线性映射：

$$A=\sum_{\pi=t\pi'\in S_n}\mathrm{sgn}(\pi)\pi=\sum_{\pi'\in S_n}\mathrm{sgn}(t\pi')\pi'=-\sum_{\pi'\in S_n}\mathrm{sgn}(\pi')\pi'=-A$$

取t是一个对换，符号为-1。于是得到 $A=0$。

问题5

考虑 A 是实方阵，取

$$B=\frac{1}{2}(A+A^T),C=\frac{1}{2}(A-A^T)$$

注意到B是实对称矩阵，C 是实反对称矩阵，所以 B 的特征值都是实数，C 的特征值都是纯虚数，更重要的是全空间可以分解成B的特征子空间的复正交直和，也可以分解成 C 的特征子空间的正交直和，而且

$$A=B+C$$

设 B 的最大特征值和最小特征值分别是 μ_1 和 μ_n，那么假设 v 是 A 关于 λ 的特征向量，有

$$v^*Av=\lambda v^*v=v^*Bv+v^*Cv$$

设 v 分解成 B 的相互复正交单位特征向量的和是，对应的特征值是 $\mu_1,..,\mu_n,\quad\mu_i\in\mathbb{R}$

$$v=b_1w_1+\ldots+b_nw_n$$

分解成 C 的相互复正交单位特征向量的和是，对应的特征值是 $\sqrt{-1}\nu_1,..,\sqrt{-1}\nu_n, \quad \nu_i\in\mathbb{R}$

$$v=c_1u_1+\ldots+c_nu_n$$

代入上式，又注意到正交分解中如下关系

$$\lambda v^*v=\sum_{i=1}^{i=n}\mu_i|b_i|^2+\sqrt{-1}\sum_{i=1}^{i=n}\nu_i|c_i|^2$$
$$v^*v=\sum_{i=1}^{i=n}|b_i|^2=\sum_{i=1}^{i=n}|c_i|^2$$

所以取实部得到

$$\mu_n\le\frac{\sum_{i=1}^{i=n}\mu_i|b_i|^2}{v^*v}=\mathrm{Re}\lambda\le\mu_1$$

问题6(实方阵的实相似标准型)

设 $A\in\mathbb{R}^{n\times n}$，考虑其极小多项式 $f(\lambda)=(\lambda^2+a\lambda+b)^k,k\in\mathbb{N}$ 没有实根， 当 $k=1$ 时根据极小多项式的定义，能够取到向量 $v\in\mathbb{R}^n$ 满足

$$A^2v=A(Av),A^2v+aAv+bv=\mathbf{0}$$

而且 v,Av 线性无关，从而考虑如下关系：

$$A[v,Av]=[Av,A^2v]=[Av,-aAv-bv]=[v,Av]\begin{bmatrix}0&-b\\1&-a\end{bmatrix}$$

这些向量生成了空间的一组基，在实相似的情况下 n 只能是2的倍数。不断取这样的向量 v_i，则在这样的向量构成的基底 $v_1,Av_1,v_2,Av_2,v_3,Av_3,..,v_{n/2},Av_{n/2}$ 下矩阵 A 可以写成：

$$A[v_1,Av_1,\ldots,v_{n/2},Av_{n/2}]=[Av_1,-aAv_1-bv_1,\ldots,Av_{n/2},-aAv_{n/2}-bv_{n/2}]=[v_1,Av_1,\ldots,v_{n/2},Av_{n/2}]\begin{bmatrix}0&-b&&&\\1&-a&&&\\&&\ddots&&\\&&&0&-b\\&&&1&-a\end{bmatrix}$$

$$A=[v_1,Av_1,\ldots,v_{n/2},Av_{n/2}](\bigoplus_{1\le i\le n/2}\begin{bmatrix}0&-b\\1&-a\end{bmatrix})[v_1,Av_1,\ldots,v_{n/2},Av_{n/2}]^{-1}$$

这时具有实相似标准型

$$\bigoplus_{1\le i\le n/2}\begin{bmatrix}0&-b\\1&-a\end{bmatrix}$$

当分歧指数高于1时，考虑 $k=2$ 的情况，存在循环向量 v 满足 $(A^2+aA^2+b)^2v=\mathbf{0}$，这时设 $r(\lambda)=\lambda^2+a\lambda+b$考虑

$$A[v,Av,r(A)v,r(A)Av]=[v,Av,r(A)v,r(A)Av]\begin{bmatrix}0&-b&&\\1&-a&&\\&1&0&-b\\&&1&-a\end{bmatrix}$$

注意到上述的 $v,Av,r(A)v,r(A)Av$ 是线性无关的，于是推广到任意指数的形式就是

$$A[v,Av,r(A)v,r(A)Av,\ldots,r(A)^{k-1},r(A)^{k-1}Av]=[v,Av,r(A)v,r(A)Av,\ldots,r(A)^{k-1},r(A)^{k-1}Av]\begin{bmatrix}0&-b&&&&&\\1&-a&&&&&\\&1&0&-b&&&\\&&1&-a&&&\\&&&&\ddots&&\\&&&&1&0&-b\\&&&&&1&-a\end{bmatrix}$$

$$A=[v,Av,r(A)v,r(A)Av,\ldots,r(A)^{k-1},r(A)^{k-1}Av]\begin{bmatrix}0&-b&&&&&\\1&-a&&&&&\\&1&0&-b&&&\\&&1&-a&&&\\&&&&\ddots&&\\&&&&1&0&-b\\&&&&&1&-a\end{bmatrix}[v,Av,r(A)v,r(A)Av,\ldots,r(A)^{k-1},r(A)^{k-1}Av]$$

有相似标准型

$$\bigoplus_{1\le i\le n/2k}\begin{bmatrix}0 & -b & & & & & & \\ 1 & -a & & & & & & \\ & 1 & 0 & -b & & & & \\ & & 1 & -a & & & & \\ & & & & \ddots & & & \\ & & & & & 1 & 0 & -b \\ & & & & & & 1 & -a\end{bmatrix}_{2k}$$

如果对 A^T 用上面结果就得到如下的标准型：

$$\bigoplus_{1\le i\le n/2k}\begin{bmatrix}0 & 1 & & & & & \\ -b & -a & 1 & & & & \\ & & 0 & 1 & & & \\ & & -b & -a & & & \\ & & & & \ddots & 1 & \\ & & & & & 0 & 1 \\ & & & & & -b & -a\end{bmatrix}_{2k}$$

事实上对不同的二次实不可约的零化因子取上述分块，就是实方阵的实相似标准型非约当块的部分。假设一个一般的时方阵的极小多项式可以写成 $(\lambda-x_1)^{r_1}\dots(\lambda-x_m)^{r_m}(\lambda^2+a_1\lambda+b_1)^{k_1}\dots(\lambda^2+a_l\lambda+b_l)^{k_l}$，那么其标准型就是一些关于实特征值的约当块 $J_i(x_k), i, k$ 和上述标准型 $Q_j(a_{k'}, b_{k'})$ 的直和。

问题7(内积的直和)

考虑 V_1 上内积 f_1 和 V_2 上内积 f_2，可以在 $V_1\oplus V_2$ 上构造如下内积 $f=f_1\oplus f_2$ 满足 V_1 和 V_2 正交；$f|_{V_1}=f_1, f|_{V_2}=f_2$

根据第一条

$$\begin{gathered}f((\mathbf{0},\beta_1),(\alpha_2,\mathbf{0}))=0\\ f((\alpha_1,\mathbf{0}),(\mathbf{0},\beta_2))\\ \text{对所有向量}\alpha_1,\alpha_2,\beta_1,\beta_2\text{成立}\end{gathered}$$

注意唯一性和第二条：

$$\begin{gathered}f((\alpha_1,\beta_1),(\alpha_2,\beta_2))=f((\alpha_1,\mathbf{0}),(\alpha_2,\beta_2))+f((\mathbf{0},\beta_1),(\alpha_2,\beta_2))\\ =f((\alpha_1,\mathbf{0}),(\alpha_2,\mathbf{0}))+f((\alpha_1,\mathbf{0}),(\mathbf{0},\beta_2))+f((\mathbf{0},\beta_1),(\alpha_2,\mathbf{0}))+f((\mathbf{0},\beta_1),(\mathbf{0},\beta_2))\\ =f_1(\alpha_1,\alpha_2)+f_2(\beta_1,\beta_2)\end{gathered}$$

根据上面计算可知唯一性，如果定义

$$f((\alpha_1,\beta_1),(\alpha_2,\beta_2))=f_1(\alpha_1,\alpha_2)+f_2(\beta_1,\beta_2)$$

验证这是一个内积即可。

问题8(极化恒等式和牛顿恒等式)_解答来自知乎用户Lociatte的解答

假设$x_1,..,x_n$ 是 n 个未定元，在多项式环$F(x_1,\dots,x_n)$ 中，有

$$x_ix_j=\frac{1}{2}(x_i+x_j)^2-\frac{1}{2}x_i^2-\frac{1}{2}x_j^2$$

那么

$$x_1\cdots x_n=\frac{1}{n!}\sum_{k=1}^{k=n}(-1)^{n-k}\sum_{1\le i_1<i_2<\dots<i_k\le n}(x_{i_1}+\dots+x_{i_k})^n$$

牛顿恒等式是对于未定元 $x_1,..,x_n$ 考虑 $e_0,..,e_n$ 是它们的基本对称多项式， $p_0,..,p_n..$ 是 $x_1,..,x_n$ 的幂和，那两者可以互相表示

证明：不妨利用幂级数来比较系数，这是代数学中非常常见的方法，神奇地定义如下两个 $F[x_1,..,x_n]$-系数的幂级数如

$$\begin{gathered}f(y):=\sum_{i=1}^{i=n}\frac{x_i}{1-x_iy}=\sum_{k\ge0}p_{k+1}y^k\\ g(y):=\prod_{i=1}^{i=n}(1+x_iy)=\sum_{0\le k\le n}e_ky^k\end{gathered}$$

那么就有幂级数环 $F[x_1,..,x_n][[y]]$ 上的相等如 $f(-y)g(y)=g'(y), f(-y)=(\ln g(y))'$ 比较系数给出

$$\begin{gathered}\sum_{i+j=m}(-1)^ip_{i+1}e_j=(m+1)e_{m+1},\qquad m+1\le n\\ \sum_{i=1}^{i=m}(-1)^ip_ie_{m-i}=0,\qquad\qquad m>n\end{gathered}$$

上面的第一个等式可以修改成更明确的形式

$$ke_k+\sum_{i=1}^{i=k}(-1)^i p_i e_{k-i}=0, \qquad k\le n$$
$$\sum_{i=1}^{i=k}(-1)^i p_i e_{k-i}=0, \qquad k>n$$

由此不难将 p_k 用 $e_0,..,e_n$ 表示，反之亦然(上面第一式已经做到)。如(注意到上面的 $e_{>n}=0$ 这一关系)

$$p_k=(-1)^{k-1}ke_k+\sum_{i=1}^{i=k-1}(-1)^{k-1-i}p_i e_{k-i}, \qquad k\le n$$
$$p_k=\sum_{i=1}^{i=k-1}(-1)^{k-1-i}p_i e_{k-i}=\sum_{i=k-n}^{i=k-1}(-1)^{k-1-i}p_i e_{k-i}, \qquad k>n$$

下面考虑如何证明极化恒等式，首先注意到两边都是 n-次对称多项式，采用算子方法能够十分巧妙地得到这个恒等式，考虑多项式环 $F[x]$ 包含在多项式环 $F[x,x_1,..,x_n]$ 中考虑平移算子 L_{x_i} 定义为 $L_{x_i}f(x)=f(x+x_i)$，那么就有如下算子的恒等式

$$\prod_{i=1}^{i=n}(L_{x_i}-1)=\sum_{k=1}^{k=n}(-1)^{n-k}\sum_{1\le i_1<\cdots<i_k\le n}L_{x_{i_1}}\cdots L_{x_{i_k}}$$

取 $f(x)=x^n$，用上面的算子作用于 f 上，右边已经是和恒等式右侧十分相似的样子

$$\sum_{k=1}^{k=n}(-1)^{n-k}\sum_{1\le i_1<\cdots<i_k\le n}f(x+x_{i_1}+\cdots+x_{i_k})=\sum_{k=1}^{k=n}(-1)^{n-k}\sum_{1\le i_1<\cdots<i_k\le n}(x+x_{i_1}+\cdots+x_{i_k})^n$$

乐观地期望左边可以是 $n!x_1\cdots x_n$。那么每一次 $L_{x_i}-1$ 的作用会使得 f 关于 x 的次数降低一次，可以想象到 $\prod_{i=1}^{i=n}(L_{x_i}-1)f$ 是一个 $F[x_1,..,x_n]$ 中的元素，而且由于 L_{x_i} 这些算子彼此交换，$\prod_{i=1}^{i=n}(L_{x_i}-1)f$ 必然是一个对称多项式。归纳法可知

$$\prod_{i=1}^{i=n}(L_{x_i}-1)f=\prod_{i=1}^{i=n-1}(L_{x_i}-1)nx_nx^{n-1}=n(n-1)\prod_{i=1}^{i=n-2}(L_{x_i}-1)x_nx_{n-1}x^{n-2}=\cdots=n!x_1\cdots x_n$$

所以令人惊讶的是这样居然说明了一个更强的极化的信息

$$x_1\cdots x_n=\frac{1}{n!}\sum_{k=1}^{k=n}(-1)^{n-k}\sum_{1\le i_1<i_2<\cdots<i_k\le n}(x+x_{i_1}+\ldots+x_{i_k})^n \tag{8.1}$$

取 $x=0$ 就给出题目里面的极化恒等式。

问题9(半正定矩阵的同时相合)

假设实对称(复自伴)半正定矩阵A,B，则存在可逆矩阵 P 满足 P^*AP 和 P^*BP 都是对角矩阵，特别地可以取 $P^*AP=I_r\oplus\mathbf{0}_{n-r}$，其中 r 是 A 的秩，n 是 A 的阶。

对称矩阵可以正交对角化，进一步可以相合于标准的对角元是±1的矩阵，由于 A 半正定，存在可逆矩阵 S 满足

$$S^*AS=I_r\oplus\mathbf{0}_{n-r}$$

考虑

$$C=S^*BS=\begin{bmatrix}B_{11} & B_{12}\\ B_{12}{}^* & B_{22}\end{bmatrix}, B_{11}\in\mathbb{R}^{r\times r}(\mathbb{C}^{r\times r})$$

作为半正定的对称(自伴)矩阵，如果 $B_{11}y=0$，那么考虑 $x=y\oplus\mathbf{0}$ 就有 $x^*Cx=0$，不妨假设 C 是对角矩阵，对角元都是正数，这样只能有 $Cx=0$ 从而 $B_{12}{}^*y=0$，从而有

$$\mathbf{ker}B_{11}\subset\mathbf{ker}B_{12}^* \quad\Rightarrow\mathbf{rank}B_{11}^*=\mathbf{rank}B_{11}\ge\mathbf{rank}B_{12}$$

从而 B_{12} 的列向量可以写成 B_{11} 的列向量的线性组合，存在 $X\in M^{r\times n-r}$ 满足 $B_{12}=B_{11}X$，于是

$$R=\begin{bmatrix}I_r & -X\\ & I_{n-r}\end{bmatrix}\Rightarrow R^*(S^*BS)R=\begin{bmatrix}B_{11} & \\ & B_{22}-X^*B_{11}X\end{bmatrix}$$

取酉矩阵 U_1,U_2 满足 $U_1{}^*B_{11}U_1=\Lambda_1$ 和 $U_2{}^*(B_{22}-X^*B_{11}X)U_2=\Lambda_2$，这里 Λ_1,Λ_2 都是对角矩阵，于是取 $U=U_1\oplus U_2$，考虑

$$P=SRU$$

计算

$$P^*AP=U^*R^*(I_r\oplus\mathbf{0}_{n-r})RU=I_r\oplus\mathbf{0}_{n-r}$$
$$P^*BP=U^*R^*CRU=\Lambda_1\oplus\Lambda_2$$

问题10(半正定矩阵平方根保序)

如果自伴(对称)矩阵 A,B 半正定而且 $A\le B$，那么 $A^{\frac{1}{2}}\le B^{\frac{1}{2}}$，（证明参考矩阵分析第二版第七章第七节）

首先假设 B 正定，那么根据问题9可以找到可逆矩阵 S^* 满足 $S^*BS=I$，$S^*AS=\Lambda$，其中 Λ 是对角元非负的对角矩阵，而相合保持半正定矩阵的序，有 $I\ge\Lambda$，所以 Λ 的对角元的绝对值都小于一，从而谱半径(特征值绝对值的最大值)

$$\rho(B^{-1}A)=\rho(SS^*S^{-*}\Lambda S^{-1})=\rho(S\Lambda S^{-1})=\rho(\Lambda)\leq 1$$

记 $\sigma_{max}(\bullet)$ 是矩阵的最大奇异值，考虑

$$1\geq\rho(B^{-1}A)=\rho(B^{\frac{1}{2}}(B^{-1}A^{\frac{1}{2}}A^{\frac{1}{2}})B^{-\frac{1}{2}})=\rho((B^{-\frac{1}{2}}A^{\frac{1}{2}})(B^{-\frac{1}{2}}A^{\frac{1}{2}})^*)=\sigma_{max}(B^{-\frac{1}{2}}A^{\frac{1}{2}})^2\geq\rho(B^{-\frac{1}{2}}A^{\frac{1}{2}})^2$$
$$\Rightarrow\quad 1\geq\rho(B^{-\frac{1}{2}}A^{\frac{1}{2}})$$

注意到 $\sigma_{max}(A)=\max_{|x|=1}|Ax|\geq\max_{Ax=\lambda x}|\lambda|=\rho(A)$ 对所有方阵 A 成立，这样考虑在某一个可逆矩阵 P 的相合下

$$B^{\frac{1}{2}}=PP^*,A^{\frac{1}{2}}=P\Lambda_1P^*$$

（Λ_1是对角元非负的对角矩阵）就有

$$1\geq\rho(B^{-\frac{1}{2}}A^{\frac{1}{2}})=\rho(P^{-*}\Lambda_1P^*)=\rho(\Lambda_1)$$

得到

$$\Lambda_1\leq I\quad\Rightarrow\quad A^{\frac{1}{2}}\leq B^{\frac{1}{2}}$$

如果 B 是半正定的，考虑一系列矩阵

$$B_k=B+\frac{1}{k}I\quad k=1,2,\ldots..$$

是正定矩阵，于是有 ${B_k}^{\frac{1}{2}}\geq A^{\frac{1}{2}}$，而且根据平方根函数的性质，考虑 B 的酉相似型 $B=U\mathrm{diag}(\lambda_1,..,\lambda_n)U^*$，有

$$B_k=U\mathrm{diag}(\lambda_1+\frac{1}{k},..,\lambda_n+\frac{1}{k})U^*$$
$${B_k}^{\frac{1}{2}}=U\mathrm{diag}(\sqrt{\lambda_1+\frac{1}{k}},..,\sqrt{\lambda_n+\frac{1}{k}})U^*$$
$$B^{\frac{1}{2}}=U\mathrm{diag}(\sqrt{\lambda_1},..,\sqrt{\lambda_n})U^*$$

根据平方根的连续性有 $\lim_{k\to+\infty}{B_k}^{\frac{1}{2}}=B^{\frac{1}{2}}$，这里极限的意义是每一个矩阵元的极限都对应相等。所以如果 $B^{\frac{1}{2}}\ngeq A^{\frac{1}{2}}$，存在向量 x_0 满足

$$x_0^*B^{\frac{1}{2}}x_0<x_0^*A^{\frac{1}{2}}x_0$$

根据 $\lim_{k\to+\infty}{B_k}^{\frac{1}{2}}=B^{\frac{1}{2}}$ 存在正整数 k_0 满足

$$|x_0^*({B_{k_0}}^{\frac{1}{2}}-B^{\frac{1}{2}})x_0|\leq\frac{1}{2}|x_0^*(A^{\frac{1}{2}}-B^{\frac{1}{2}})x_0|$$

于是

$$x_0^*B^{\frac{1}{2}}x_0\geq x_0^*{B_{k_0}}^{\frac{1}{2}}x_0-\frac{1}{2}|x_0^*(A^{\frac{1}{2}}-B^{\frac{1}{2}})x_0|\geq x_0^*A^{\frac{1}{2}}x_0-\frac{1}{2}|x_0^*(A^{\frac{1}{2}}-B^{\frac{1}{2}})x_0|=\frac{1}{2}x_0^*(A^{\frac{1}{2}}+B^{\frac{1}{2}})x_0\Rightarrow x_0^*(A^{\frac{1}{2}}-B^{\frac{1}{2}})x_0\leq 0$$

和假设矛盾，所以有

$$A^{\frac{1}{2}}\leq B^{\frac{1}{2}}$$

问题11(复正交矩阵的实正交等价)

取 $A=B+\sqrt{-1}C$ 是复正交矩阵，B,C 是实分量和虚分量，为此首先对 B 进行奇异值分解

$$B=O_3^t\begin{bmatrix}\mathrm{M} & \\ & \end{bmatrix}O_4^t$$

其中 O_3,O_4 是实正交矩阵，$\mathrm{M}=\mathrm{diag}(\mu_1,..,\mu_s)$ 是 B 的（从大到小的）奇异值组成的对角矩阵。

在讨论实正交等价时不妨在 A 左右乘以 O_3,O_4 不改变其复正交性，进而考虑 C 是 O_3CO_4 的情况，这一操作也不改变 C 的秩。注意到有如下运算关系

$$AA^t=(BB^t-CC^t)+\sqrt{-1}(BC^t+CB^t)=I\Rightarrow CC^t=BB^t-I,BC^t+CB^t=0$$
$$A^tA=(B^tB-C^tC)+\sqrt{-1}(C^tB+B^tC)=I\Rightarrow C^tC=B^tB-I,C^tB+B^tC=0$$

不妨按照 $B=\begin{bmatrix}\mathrm{M} & \\ & \end{bmatrix}$ 处理，于是

$$C^tC=\begin{bmatrix}\mathrm{M}^2-I_s & \\ & -I_{n-s}\end{bmatrix}$$

而 CC^t 半正定，从而 B 的奇异值只能大于等于1，于是取 $\mu_1,..,\mu_m$ 是(从小到大排列的)B 的大于 1 的奇异值值，从而设

$$B=\begin{bmatrix}T & \\ & I\end{bmatrix},T=\mathrm{diag}(\mu_1,..,\mu_m)$$

计算 CC^t,C^tC 得到

$$
C^tC = \begin{bmatrix} \mu_1^2-1 & & & \\ & \ddots & & \\ & & \mu_m^2-1 & \\ & & & \mathbf{0}_{n-t} \end{bmatrix} = BB^t - I = B^tB - I = CC^t
$$

于是可知 C 是正规矩阵，而且上述格拉姆矩阵可以说明 C 的列向量彼此正交而且模长是$\sqrt{\mu_1^2-1},..,\sqrt{\mu_m^2-1}$，

利用斜对称条件 $BC^t + CB = C^tB + BC = 0$ 得到 $C^t = -B^{-1}CB, CB^2 = B^2C$对 C 进行和 B 共形的分块计算，得到

$$
C = \begin{bmatrix} C_1 & \\ & \end{bmatrix} \Rightarrow (C_1T) + (C_1T)^t = \mathbf{0}, (TC_1) + (C_1T)^t = \mathbf{0}
$$

注意到 C_1 和 T 均可逆，C_1T, TC_1 斜对称，假设其特征多项式是

$$
f(x) = \det(C_1T - xI) = \det((C_1T)^t - xI) = \det(-C_1T - xI) = (-1)^m f(-x)
$$

从而考虑其零点可知 f 只有 x 或者 $(x^2+b^2)(b \in \mathbb{R})$ 两种因子，进一步由于 C_1T 可逆，从而 0 不是特征值得到 $m = 2t$，于是 C_1T 具有特征值

$$
\pm\sqrt{-1}b_1, .., \pm\sqrt{-1}b_t, b_i > 0
$$

甚至对应到模长 $(C_1T)^t(C_1T) = \mathrm{diag}(b_1^2, b_1^2, .., b_t^2, b_t^2) = \mathrm{diag}(\mu_1^2(\mu_1^2-1), .., \mu_m^2(\mu_m^2-1))$，不难说明 $g(x) = x^4 - x^2$ 在 $x > 1$ 时是严格单调递增的，从而 g 是单射，那么从 $\mu_i^2(\mu_i^2-1)$ 可以两两配成相等的一对可知 μ_i 也能两两配成相等的一对，于是至少可以知道 m 个奇异值中至少有 t 对相等，于是不妨在之前的假设中取成 $\mu_1, \mu_1, .., \mu_t, \mu_t$，且 $b_i = \mu_i\sqrt{\mu_i^2-1}$。

C_1T 作为斜对称矩阵，$\sqrt{-1}C_1T$ 作为自伴矩阵存在酉对角化，从而 C_1T也存在酉对角化，这意味着其特征值对应的复特征向量可以选择成彼此共轭正交的，假设我们做出了这样的选择比如 $v_1, w_1.., v_t, w_t$，且满足 $C_1Tv_i = \sqrt{-1}b_iv_i, C_1Tw_i = -\sqrt{-1}b_iw_i$

如果把向量 v_i 分解成实部和虚部 $v_i = v_i^R + \sqrt{-1}v_i^I$ 我们得到

$$
A(v_i^R + \sqrt{-1}v_i^I) = (\sqrt{-1}b_i)(v_i^R + \sqrt{-1}v_i^I)
$$
$$
Av_i^R = -b_iv_i^I, Av_i^I = b_iv_i^R
$$
$$
\langle v_i^R, v_i^I\rangle = \frac{1}{b_i}\langle Av_i^I, v_i^I\rangle = -\frac{1}{b_i}\langle v_i^I, Av_i^I\rangle = 0 \quad i = 1, \dots, t
$$

如果能够说明对 $i \neq j$ v_i^I, v_i^R 和 v_j^I, v_j^R彼此正交，就可以得到斜对称矩阵的正交标准型，作如下计算

$$
0 = v_i^*v_j = \langle v_i, v_j\rangle = (\langle v_i^R, v_j^R\rangle + \langle v_i^I, v_j^I\rangle) + \sqrt{-1}(\langle v_i^R, v_j^I\rangle - \langle v_i^I, v_j^R\rangle)
$$
$$
\langle v_i^R, v_j^R\rangle = \frac{b_i}{b_j}\langle v_i^I, v_j^I\rangle, \quad \langle v_i^R, v_j^I\rangle = -\frac{b_i}{b_j}\langle v_i^I, v_j^I\rangle
$$

由于 $1 + \frac{b_i}{b_j}$ 非零，从而 $\langle v_i^R, v_j^R\rangle = \langle v_i^I, v_j^R\rangle = \langle v_i^R, v_j^I\rangle = \langle v_i^I, v_j^I\rangle = 0$，这样在正交向量组 $v_1^R, v_1^I, .., v_t^R, v_t^I$ 下 C_1T 可写成

$$
F = \mathrm{diag}(\begin{bmatrix} & \mu_1\sqrt{\mu_1^2-1} \\ -\mu_1\sqrt{\mu_1^2-1} & \end{bmatrix}, .., \begin{bmatrix} & \mu_t\sqrt{\mu_t^2-1} \\ -\mu_t\sqrt{\mu_t^2-1} & \end{bmatrix})
$$

对上述正交向量组作适当的标准化可知 C_1T 正交相似于上述矩阵，事实上上述结论对一般斜对称实方阵都成立，特别的如果取

$$
(C_1T^{-1}) + (C_1T^{-1})^t = T^{-1}((TC_1) + (TC_1)^t)T^{-1} = \mathbf{0}
$$

注意到 C_1T^{-1} 作为斜对称实矩阵的特征值是 $\pm\sqrt{-1}\sqrt{\mu_i^2-1}$，其正交相似于

$$
F' = \mathrm{diag}(\begin{bmatrix} & \frac{\sqrt{\mu_1^2-1}}{\mu_1} \\ -\frac{\sqrt{\mu_1^2-1}}{\mu_1} & \end{bmatrix}, .., \begin{bmatrix} & \frac{\sqrt{\mu_t^2-1}}{\mu_t} \\ -\frac{\sqrt{\mu_t^2-1}}{\mu_t} & \end{bmatrix})
$$

那么假设 $C_1T^{-1} = O_5F'O_5^t$，$T = \mathrm{diag}(\mu_1, \mu_1, .., \mu_t, \mu_t)$，其中矩阵$O_3$ 实正交，设 $O_6 = T^{-1}O_5^tT$ 也是实正交矩阵，就有 $T = O_5TO_6, O_5^tT = TO_6$，从而得到

$$
C_1 = C_1T^{-1}T = O_5F'O_5^tT = O_5F'TO_6 = O_5\mathrm{diag}(\begin{bmatrix} & \sqrt{\mu_1^2-1} \\ -\sqrt{\mu_1^2-1} & \end{bmatrix}, .., \begin{bmatrix} & \sqrt{\mu_t^2-1} \\ -\sqrt{\mu_t^2-1} & \end{bmatrix})O_6
$$

这样有了前期这些铺垫，考虑

$$A=O_3(\begin{bmatrix}T & \\ & I\end{bmatrix}+\sqrt{-1}\begin{bmatrix}C_1 & \\ & \end{bmatrix})O_4$$

$$=O_3\begin{bmatrix}O_5 & \\ & I\end{bmatrix}(\begin{bmatrix}T & \\ & I\end{bmatrix}+\sqrt{-1}\mathrm{diag}(\begin{bmatrix} & \sqrt{\mu_1^2-1}\\ -\sqrt{\mu_1^2-1} & \end{bmatrix},\ldots,\begin{bmatrix} & \sqrt{\mu_t^2-1}\\ -\sqrt{\mu_t^2-1} & \end{bmatrix},\mathbf{0}_{n-2t}))\begin{bmatrix}O_6 & \\ & I\end{bmatrix}O_4=$$

$$O_3\begin{bmatrix}O_5 & \\ & I\end{bmatrix}P\mathrm{diag}(\mathbf{1}_{n-2t},\begin{bmatrix}\mu_1 & \sqrt{-1}\sqrt{\mu_1^2-1}\\ -\sqrt{-1}\sqrt{\mu_1^2-1} & \mu_1\end{bmatrix},\ldots,\begin{bmatrix}\mu_2 & \sqrt{-1}\sqrt{\mu_t^2-1}\\ -\sqrt{-1}\sqrt{\mu_t^2-1} & \mu_2\end{bmatrix}))P\begin{bmatrix}O_6 & \\ & I\end{bmatrix}O_4$$

这里 P 是把前 $2t$ 位置和后 $n-2t$ 位置颠倒的置换矩阵，自然是实正交矩阵，那么取

$$O_1^t=O_3\begin{bmatrix}O_5 & \\ & I\end{bmatrix}P,\quad O_2^t=P\begin{bmatrix}O_6 & \\ & I\end{bmatrix}O_4$$

为实正交矩阵，得到

$$O_1AO_2=\mathrm{diag}(\mathbf{1}_{n-2t},\begin{bmatrix}\mu_1 & \sqrt{-1}\sqrt{\mu_1^2-1}\\ -\sqrt{-1}\sqrt{\mu_1^2-1} & \mu_1\end{bmatrix},\ldots,\begin{bmatrix}\mu_2 & \sqrt{-1}\sqrt{\mu_t^2-1}\\ -\sqrt{-1}\sqrt{\mu_t^2-1} & \mu_2\end{bmatrix}))$$

问题12(复正交矩阵的实正交相抵不变量)

简言之，$A_1=A_1^R+\sqrt{-1}A_1^I$，其中 A_1^R,A_1^I 是复方阵 A_1 的实部和虚部，他们都是实方阵，如果 $A_2=O_1A_1O_2$，而且 O_1,O_2实正交，那么下降到实部就有

$$A_2^R=O_1A_1^RO_2$$

而奇异值是 $A_2^RA_2^{R^t}=O_1A_1^RA_1^{R^t}O_1^t$ 的非零特征值的平方根，那么就有

$$\sigma(A_2^R)=\rho(A_2^RA_2^{R^t})=\rho(O_1A_1^RA_1^{R^t}O_1^t)=\rho(A_1^RA_1^{R^t})=\sigma(A_1^R)$$

所以 $\sigma(-)$ 确实是正交等价下的不变量。

问题13(半定矩阵的特征值估计)

设 $J=(1)_{ij}$ 是矩阵元全为1的 n 阶方阵，A 是所有列和全为0的半正定对称矩阵，那么根据列和为零的条件和对称条件直接得到所有行和也为零

$$AJ=JA=\mathbf{0}$$

取 A 的特征值的一个降序排列 $\lambda_1\geq\lambda_2\geq\cdots\geq\lambda_{n-1}\geq\lambda_n\geq 0$考虑对称方阵

$$A-\lambda_{n-1}(I_n-\frac{1}{n}J)$$

如果 $\lambda_{n-1}=0$，不等式自然成立，否则 $\mathrm{rank}A=n-1$ 考虑 A 将全空间分解成上述特征值对应的互相正交的特征子空间 $v_1,..,v_n$,对于非零的特征值 $\lambda>0$，考虑 $Av_\lambda=\lambda v_\lambda$,有

$$\sum_k a_{ik}v_{\lambda k}=\lambda v_{\lambda i}\quad i=1,..,n$$

$$\sum_i\sum_k a_{ik}v_{\lambda k}=\sum_k(\sum_i a_{ik})v_{\lambda k}=\lambda\sum_i v_{\lambda k}=0\Rightarrow\sum_i v_{\lambda k}=0,Jv_\lambda=\mathbf{0}$$

于是 $Jv_i=0,i\leq n-1$，那么考虑任一个向量 $v=\sum_i x_iv_i$，有

$$[A-\lambda_{n-1}(I_n-\frac{1}{n}J)]v=\sum_i(\lambda_i-\lambda_{n-1})x_iv_i+\frac{\lambda_{n-1}x_n}{n}Jv_n$$

$$v^*[A-\lambda_{n-1}(I_n-\frac{1}{n}J)]v=\sum_i(\lambda_i-\lambda_{n-1})x_i^2||v_i||^2+\frac{x_n^2v_n^*Jv_n}{n}\lambda_{n-1}$$

由于A 奇异，$\lambda_n=0$，那么 $\sum_i a_{ik}=\sum_i a_{ik}v_{nk}=0$，由于 $\ker A$ 是一维的，那么 $v_n\in\mathbb{C}\mathbf{1}$，即 $v_{n1}=..=v_{nn}$

$$\frac{x_n^2v_n^*Jv_n}{n}=x_n^2\frac{(v_{n1}+\ldots+v_{nn})^2}{n}=x_n^2||v_n||^2$$

从而对任何向量 $v=\sum_i x_iv_i$

$$v^*[A-\lambda_{n-1}(I_n-\frac{1}{n}J)]v=\sum_i(\lambda_i-\lambda_{n-1})x_i^2||v_i||^2+\frac{x_n^2v_n^*Jv_n}{n}\lambda_{n-1}=\sum_{i\leq n-2}(\lambda_i-\lambda_{n-1})x_i^2||v_i||^2\geq 0$$

从而 $A-\lambda_{n-1}(I_n-\frac{1}{n}J)$ 半正定，根据西尔维斯特惯性定理或取 $e_i^*(A-\lambda_{n-1}(I_n-\frac{1}{n}J))e_i\geq 0$ 可以得到这个矩阵的对角元都是非负的，于是

$$0\leq\min_i\{a_{ii}-\lambda_{n-1}(\frac{n-1}{n})\}=\min_i\{a_{ii}\}-\lambda_{n-1}\frac{n-1}{n}$$

$$\Rightarrow\lambda_{n-1}\leq\frac{n}{n-1}\min_i\{a_{ii}\}$$

问题14(投影算子的正交相似型)

考虑 $P^2=P$ 是有限维 k-线性空间上的投影算子，而且 V 上有非退化的内积 $\langle -,-\rangle$ 则 P 正交相似于如下的分块矩阵

$$P\simeq_{Q(-)Q^t}\begin{bmatrix}1 & \sigma_1\\ & \end{bmatrix}\oplus\cdots\oplus\begin{bmatrix}1 & \sigma_k\\ & \end{bmatrix}\oplus I_m\oplus 0_s$$

证明来自于对投影算子的刻画，一般而言通过准素分解和正交分解有 $V=\ker P\oplus\ker(1-P)=\ker P\oplus \mathrm{Im}P=\mathrm{Im}P^{\perp}\oplus\ker P$。那么 $\mathrm{Im}P=\ker(1-P)$ 说明子空间 $\mathrm{Im}P$ 的正交分解共享了特征值为1的部分，希望 $\mathrm{Im}P^{\perp}$ 贡献剩余的部分，首先计算维数

$$\begin{aligned}\dim\mathrm{Im}P&=k+m&&=\dim\ker P^{\perp}\\ \dim\ker P&=k+s&&=\dim\mathrm{Im}P^{\perp}\\ \dim V&=2k+m+s\end{aligned}$$

投影算子会消灭一些正交于投影空间 $\mathrm{Im}P$ 的子空间，这些向量是投影算子比较"像"正交投影的部分，记

$$V_{\text{orthogonal}}=\mathrm{Im}P^{\perp}\cap\ker P\qquad \dim V_{\text{orthogonal}}=s$$

而投影算子未消灭的正交补中向量一一对应了那些和投影空间 $\mathrm{Im}P$ 斜交的部分也就是 $\ker P/V_{\text{orthogonal}}$。这个空间同构于

$$P^{-1}(\mathrm{Im}P|_{\mathrm{Im}P^{\perp}})$$

事实上可以记斜交部分的空间是

$$V_{\text{oblique}}=\mathrm{Im}P|_{\mathrm{Im}P^{\perp}}\qquad \dim V_{\text{oblique}}=k$$

首先取单位化后的 $V_{\text{orthogonal}}$ 的一组正交基 $\{v_1,..,v_s\}$ 使得 $Pv_i=0,i=1,..,s$，考虑这些向量在 $\mathrm{Im}P^{\perp}$ 的正交补维数是k。那么 P 限制在 $\mathrm{Im}P^{\perp}\cap V_{\text{orthogonal}}^{\perp}$ 上是满射，也有 $P|_{\mathrm{Im}P^{\perp}\cap V_{\text{orthogonal}}^{\perp}}\mathrm{Im}P^{\perp}\cap V_{\text{orthogonal}}^{\perp}=V_{\text{oblique}}$。则利用 P^*P 的特征子空间的正交分解，也是导致 P 的奇异值分解的本质方法，可以找到 $\mathrm{Im}P^{\perp}\cap V_{\text{orthogonal}}^{\perp}$ 的一组正交基使得 P 把它们半单地映射到 V_{oblique} 的一组正交基，分别将两者单位化，记

$$\begin{gathered}\mathrm{Im}P^{\perp}\cap V_{\text{orthogonal}}^{\perp}=\mathrm{Span}\{v_{s+1},..,v_{s+k}\}\\ V_{\text{oblique}}=\mathrm{Span}\{v_{s+k+1},..,v_{s+2k}\}\\ Pv_{s+i}=\sigma_i v_{s+k+i},\quad \sigma_i>0\text{作为}P\text{的奇异值}\end{gathered}$$

P 在 v_{s+k+i},v_{s+i} 下的矩阵形式就是

$$\begin{bmatrix}1 & \sigma_i\\ & \end{bmatrix}$$

最后将 $\{v_{s+k+1},..,v_{s+2k}\}$ 扩充成 $\mathrm{Im}P$ 的一组标准正交基(可采用Schmit正交化，$fv_{n+1}=v_{n+1}-P_{fv_1,..,fv_n}(v_{n+1})$，$P_{..}$ 是正交投影)如 $\{v_{s+k+1},..,v_{s+2k+m}\}$。那么在标准正交基 $\{v_{s+k+1},v_{s+1},..,v_{s+2k},v_{s+k},v_{s+2k+1},...,v_{s+2k+m},v_1,..,v_s\}$ 下 P 的作用写成矩阵形式就是

$$\begin{bmatrix}1 & \sigma_1\\ & \end{bmatrix}\oplus\cdots\oplus\begin{bmatrix}1 & \sigma_k\\ & \end{bmatrix}\oplus I_m\oplus 0_s$$

问题15(trace和determinant的唯一性)

如果在 $\mathrm{Hom}_k(V,V)$ 上定义一个满足如下条件的函数，

$$tr(AB)=tr(BA),tr(1_V)=\dim_k V,tr(A+kB)=tr(A)+ktr(B)$$

则一定有 $tr=\mathrm{tr},\quad \mathrm{tr}\,(A)=\sum_i\langle Av_i,v_i\rangle$，这里 $\{v_i\}_i$ 是任取的正交基。

首先说明上述定义不依赖于正交基的选择，假设 $v_i=Qw_i$ 这里 Q 是正交(酉)矩阵，那么计算

$$\begin{aligned}\mathrm{tr}\,(A)=\sum_i\langle Av_i,v_i\rangle=\sum_i\langle AQw_i,Qw_i\rangle&=\sum_{i,j,k}\langle AQ_{ij}w_j,Q_{ik}w_k\rangle=\sum_{i,j,k}Q_{ij}\overline{Q_{ik}}\langle Aw_j,w_k\rangle\\&=\sum_{j,k}\delta_{kj}\langle Aw_j,w_k\rangle=\sum_i\langle Aw_i,w_i\rangle\end{aligned}$$

接下来说明这样的函数一定是迹函数，首先根据定义 tr 是共轭不变的，而且

$$tr(E_{ij}E_{jk})=tr(E_{ik})=tr(E_{jk}E_{ij})=\delta_{ik}tr(E_{jj})$$

给出 $tr(E_{11})=\cdots=tr(E_{nn}),tr(E_{ij})=0,i\neq j$，所以计算 $tr(1)=n$ 有 $tr(E_{ii})=1$ 推出 $tr=\mathrm{tr}$。

类似有行列式的唯一性，$\det$ 可以由乘性完全确定 $f(AB)=f(A)f(B)$ 和 $f(I)=1,f(kA)=k^nA$。上三角化可以说明代数闭域的情况，一般的域不妨使用初等矩阵。

问题16(迹零矩阵写成换位子)

对一个数域 F 上的迹零方阵，其总能写成 $AB-BA$ 的形式。

这样的一个命题来自于 $\mathrm{tr}\,(AB)=\mathrm{tr}\,(BA)$ 的事实，自然想象 $AB-BA$ 是不是可以表示所有的迹零矩阵。这一般只在有限维情况下讨论，例如在无穷维时考虑光滑函数空间 $C^{\infty}(\mathbb{R})$ 中的数乘 $x\cdot$ 算子和求导算子 ∂_x 则

$$[x,\partial_x]=1$$

不是一个迹零算子。换位子构成的集合会大于迹零算子构成的集合。

很有趣的例子是在有限域的时候，例如 $\mathrm{M}_2(\mathbb{F}_2)$ 中有 $AB-BA=I$，这在特征零的情况下是不存在的。如

$$I_2=\begin{bmatrix}1&1\\&1\end{bmatrix}\begin{bmatrix}1&\\1&1\end{bmatrix}-\begin{bmatrix}1&\\1&1\end{bmatrix}\begin{bmatrix}1&1\\&1\end{bmatrix}$$

问题17(Lie定理的想法)

考虑复方阵如 A,B，如果 $[A,[A,B]]=[B,[A,B]]=0$ 则 $[A,B]$ 是幂零矩阵。

我们这里从解题的想法出发，如果 $C=[A,B]$ 是幂零矩阵，那存在正整数 r 使得 $C^r=0$，那么 C 的特征值一定都是零，从而考虑上三角化得到 $\operatorname{tr}\left(C^i\right)=\sum_{\lambda\in\sigma(C)}\lambda^i=0$ 对每一个指数 i 都成立。反过来如果 $\operatorname{tr}\left(C^i\right)=0\quad\forall i>0$，那么 特征多项式的根的 i 次幂和都是零，这让我们想到经典的牛顿恒等式(此时我们只需要次数小于等于 n 的那部分幂和)。牛顿恒等式告诉我们一个多项式的系数和根的幂和可以互相表示，最为关键的是高次项的系数可以表示为低次项的幂和，反之亦然。参考第八题，那么根据特征多项式代入 C 得到零，我们已经知道此时的特征多项式是 $f(x)=x^n$ 了，那么代入得到 $C^n=0$。也就是说

$$C\in\mathrm{M}_n(F)\text{幂零}\Leftrightarrow\operatorname{tr}\left(C^i\right)=0,\quad i=1,2,..\Leftrightarrow\operatorname{tr}\left(C^i\right)=0,\quad i=1,..,n$$

于是问题可以简化成计算 $[A,B]^i$ 的迹，验证它们都是零。此时 $\operatorname{tr}(C)=0$ 而且利用归纳的思想说明 C^k 和 A,B 交换

$$C^{k+1}=(AB-BA)C=ABC^k-BAC^k=ABC^k-BC^kA\Rightarrow\operatorname{tr}\left(C^{k+1}\right)=\operatorname{tr}\left(A\cdot BC^k\right)-\operatorname{tr}\left(BC^k\cdot A\right)=0$$

所以这里已经说明了 $\operatorname{tr}\left(C^i\right)=0,\quad i=1,2,...$。按照上面的论述来说，$C$ 就一定是一个幂零矩阵。

在证明过程中也可以看到，只需要 C 和 A 交换或者 C 和 B 交换这两个条件之一便可推出幂零的结论。