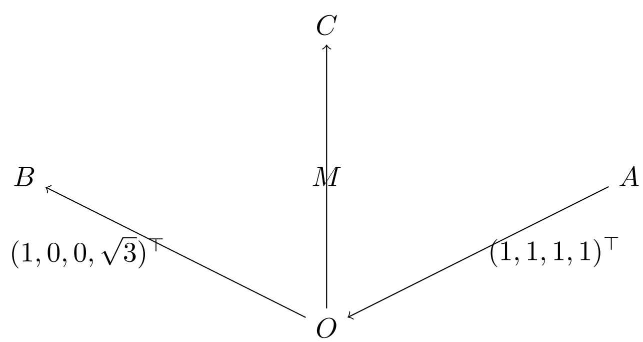
矩阵理论 作业汇总

刘彦铭 ID: 122033910081

Last Edited: 2023 年 2 月 20 日

1. Page 5 习题 1

从镜面反射变换的几何意义来看:



镜面方向上的 α 与 $\overrightarrow{OC} = \overrightarrow{OB} + \overrightarrow{OA} = \overrightarrow{OB} - \overrightarrow{AO} = (0,-1,-1,\sqrt{3}-1)^\top$ 同方向

故单位向量 $\alpha = \dfrac{\overrightarrow{OC}}{|\overrightarrow{OC}|} = \dfrac{1}{\sqrt{6-2\sqrt{3}}}(0,-1,-1,\sqrt{3}-1)^\top$

所以

$$B = E - 2\alpha\alpha^T = \frac{1}{3-\sqrt{3}}\begin{bmatrix} 3-\sqrt{3} & 0 & 0 & 0 \\ 0 & 2-\sqrt{3} & -1 & \sqrt{3}-1 \\ 0 & -1 & 2-\sqrt{3} & \sqrt{3}-1 \\ 0 & \sqrt{3}-1 & \sqrt{3}-1 & \sqrt{3}-1 \end{bmatrix}$$

2. Page 11 习题 1

可将该矩阵分解为行变换矩阵和分块对角阵的乘积:

$$\begin{bmatrix} 0 & B & 0 \\ E & 0 & 0 \\ 0 & 0 & A \end{bmatrix} = \begin{bmatrix} 0 & E & 0 \\ E & 0 & 0 \\ 0 & 0 & E \end{bmatrix} \times \begin{bmatrix} E & 0 & 0 \\ 0 & B & 0 \\ 0 & 0 & A \end{bmatrix}$$

故其逆矩阵为

$$\begin{bmatrix} 0 & B & 0 \\ E & 0 & 0 \\ 0 & 0 & A \end{bmatrix}^{-1} = \begin{bmatrix} E & 0 & 0 \\ 0 & B^{-1} & 0 \\ 0 & 0 & A^{-1} \end{bmatrix} \times \begin{bmatrix} 0 & E & 0 \\ E & 0 & 0 \\ 0 & 0 & E \end{bmatrix} = \begin{bmatrix} 0 & E & 0 \\ B^{-1} & 0 & 0 \\ 0 & 0 & A^{-1} \end{bmatrix}$$

行列式:

$$\begin{vmatrix} 0 & B & 0 \\ E & 0 & 0 \\ 0 & 0 & A \end{vmatrix} = \begin{vmatrix} 0 & E & 0 \\ E & 0 & 0 \\ 0 & 0 & E \end{vmatrix} \times \begin{vmatrix} E & 0 & 0 \\ 0 & B & 0 \\ 0 & 0 & A \end{vmatrix} = (-1)^n \cdot \det(A) \cdot \det(B)$$

3. Page 11 习题 3

注意到对该方阵进行行变换可以得到

$$\begin{bmatrix} E & 0 \\ -C & E \end{bmatrix} \times \begin{bmatrix} A^{-1} & 0 \\ 0 & E_m \end{bmatrix} \times \begin{bmatrix} A & B \\ C & D_m \end{bmatrix} = \begin{bmatrix} E & A^{-1}B \\ 0 & D_m - CA^{-1}B \end{bmatrix}$$

上式中左边两个行变换矩阵均满秩，故 $H = \begin{bmatrix} A & B \\ C & D_m \end{bmatrix}$ 可逆的充要条件为

$$\begin{vmatrix} E & A^{-1}B \\ 0 & D_m - CA^{-1}B \end{vmatrix} = \det(D_m - CA^{-1}B) \neq 0$$

(1) 由上述讨论知，充要条件是 m 阶方阵 $D_m - CA^{-1}B$ 可逆；

(2) 简单起见，令 $D' = D_m - CA^{-1}B$, H 可逆时，D'^{-1} 唯一存在。注意到，运用行变换有:

$$\begin{bmatrix} E & -A^{-1}B \\ 0 & E \end{bmatrix} \times \begin{bmatrix} E & 0 \\ 0 & D'^{-1} \end{bmatrix} \times \begin{bmatrix} E & A^{-1}B \\ 0 & D' \end{bmatrix} = E_{n+m}$$

所以

$$\begin{aligned} H^{-1} &= \begin{bmatrix} E & -A^{-1}B \\ 0 & E \end{bmatrix} \times \begin{bmatrix} E & 0 \\ 0 & D'^{-1} \end{bmatrix} \times \begin{bmatrix} E & 0 \\ -C & E \end{bmatrix} \times \begin{bmatrix} A^{-1} & 0 \\ 0 & E_m \end{bmatrix} \\ &= \begin{bmatrix} A^{-1} + A^{-1}B(D_m - CA^{-1}B)^{-1}CA^{-1} & -A^{-1}B(D_m - CA^{-1}B)^{-1} \\ -(D_m - CA^{-1}B)^{-1}CA^{-1} & (D_m - CA^{-1}B)^{-1} \end{bmatrix} \end{aligned}$$

非常丑，但验证了一下应该是对的。左上角似乎和 Woodbury 公式的形式是一致的。

4. Page 12 习题 6

考虑当 $A \in \mathbb{R}^{n\times n}$ 时，对任意列向量 $x' = \left[x^\top; y\right]^\top \in \mathbb{R}^{(n+1)\times 1}$, 其中 $x \in \mathbb{R}^{n\times 1}, y \in \mathbb{R}$:

计算得到

$$\begin{aligned} f(x') = x'^\top \begin{bmatrix} A & k\alpha \\ k\alpha^\top & 1 \end{bmatrix} x' &= \left[x^\top; y\right] \begin{bmatrix} A & k\alpha \\ k\alpha^\top & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \\ &= x^\top Ax + 2ky(x^\top\alpha) + y^2 \\ &= \big(y + k(x^\top\alpha)\big)^2 + x^\top Ax - k^2(x^\top\alpha x^\top\alpha) \\ &= \big(y + k(x^\top\alpha)\big)^2 + x^\top(A - k^2\alpha\alpha^\top)x \end{aligned}$$

- 若 $A-k^2\alpha\alpha^\top$ 是正定矩阵，那么 $\begin{bmatrix} A & k\alpha \\ k\alpha^\top & 1 \end{bmatrix}$ 正定。因为:
  - 一方面 $f(x')\geq 0$ 恒成立;
  - 另一方面, $f(x')=0$ 可以推出 $x=0,y=0,x'=0$.
- 若 $A-k^2\alpha\alpha^\top$ 是半正定矩阵，那么 $\begin{bmatrix} A & k\alpha \\ k\alpha^\top & 1 \end{bmatrix}$ 半正定。因为:
  - 一方面 $f(x')\geq 0$ 恒成立;
  - 另一方面, 由于存在 $x\neq 0$ 使得 $x^\top(A-k^2\alpha\alpha^\top)x=0$, 取 $y=-k(x^\top\alpha)$ 即得到非零的 x' 使得 $f(x')=0$.
- 若 $A-k^2\alpha\alpha^\top$ 是不定矩阵，那么 $\begin{bmatrix} A & k\alpha \\ k\alpha^\top & 1 \end{bmatrix}$ 也是不定的。因为: 此时 $A-k^2\alpha\alpha^\top$ 存在小于 0 的特征值，取 x 为其对应的特征向量，取 $y=-k(x^\top\alpha)$ 即构造得到 x', $f(x')<0$.

5. Page 13 习题 7

   方便起见，令 $C=AB\in\mathbb{F}^{m\times r}$, 直接计算验证：对于任一 $1\leq k\leq r$,

$$\begin{aligned}\sum_{1\leq i\leq m} C_{ik} &= \sum_{1\leq i\leq m}\sum_{1\leq j\leq n} A_{ij}\times B_{jk}\\ &= \sum_{1\leq j\leq n}\sum_{1\leq i\leq m} A_{ij}\times B_{jk}\\ &= \sum_{1\leq j\leq n} B_{jk}\times\left(\sum_{1\leq i\leq m} A_{ij}\right)\\ &= \sum_{1\leq j\leq n} B_{jk}\times a\\ &= ab\end{aligned}$$

6. Page 16 习题 1

   课上已经讲过解法:

   设 $f(x),g(x)\in\mathbb{F}[x]$, $(f(x),g(x))=1\Rightarrow\exists\, u(x),v(x)\in\mathbb{F}[x]$ 使得 $f(x)\cdot u(x)+g(x)\cdot v(x)=1$.

   将上述多项式的 x 代换为 x^n 即得: $f(x^n)\cdot u(x^n)+g(x^n)\cdot v(x^n)=1$.

   容易验证 $u(x^n),v(x^n)\in\mathbb{F}[x]$, 这就证明了 $(f(x^n),g(x^n))=1$.

7. Page 16 习题 4

   (1) 假设 $p(x)\in\mathbb{Q}[x]$ 也是满足 $p(\alpha)=0$ 的最低次的首一多项式.

   由于都是最低次的，所以有 $\deg m_\alpha=\deg p$.

   作带余除法：存在多项式 $u,v\in\mathbb{Q}[x]$ 使得 $m_\alpha=u\cdot p+v$, 其中 $v=0$ 或者 $\deg v<\deg p$.

   注意到 $v(\alpha)=m_\alpha(\alpha)-u(\alpha)\cdot p(\alpha)=0$, 所以有 $v=0$; 否则存在非零的多项式 $v\in\mathbb{Q}[x]$ 使得 $v(\alpha)=0$ 且 $\deg v<\deg p$，这与 p 最低次的假设矛盾。

   所以 $m_\alpha=u\cdot p$. 因为 $\deg u=\deg m_\alpha-\deg p=0$ 且 m_α,p 均首一, 所以 $u=1$

   因此 $m_\alpha=p$, 这就说明了 m_α 的唯一性

   (2) 只需说明 $\{1,\alpha,\alpha^2,\cdots,\alpha^{m-1}\}$ 是 $\mathbb{Q}[\alpha]$ 的一组基:

- 线性无关：

  对任意的 $c_0, c_1, c_2, \cdots, c_{m-1} \in \mathbb{Q}$, 若 $f(\alpha) = \sum_{0 \le i < m} c_i \alpha^i = 0$, 由于 $\deg f = m - 1 < \deg m_\alpha$, 所以由 m_α 的定义知 $f = 0$, 即 $c_i = 0, \forall\ 0 \le i < m$. 这就证明了 $\{\alpha^i\}, 0 \le i < m$ 的线性无关性。

- 可表示性：

  对 $\mathbb{Q}[\alpha]$ 上的任意一个元素 β, 由 $\mathbb{Q}[\alpha]$ 的生成方式可以知道，存在多项式 $f \in \mathbb{Q}[x]$, 使得 $\beta = f(\alpha)$ 考虑带余除法 $f = q \cdot m_\alpha + r$ 其中 $q, r \in \mathbb{Q}[x]$, $r = 0$ 或 $\deg r < \deg m_\alpha = m$.

  于是 $\beta = f(\alpha) = q(\alpha) \cdot m_\alpha(\alpha) + r(\alpha) = r(\alpha) = \sum_{0 \le i < m} c_i \alpha^i$. 这就说明了 $\mathbb{Q}[\alpha]$ 上的任一元素都能由 $\{\alpha^i\}, 0 \le i < m$ 线性表示

8. Page 16-17 习题 5 (尝试做一下)

(1) 设 p 是 R 上的任意一个素元。对于任意的非零的 $p_1, p_2 \in R$, 如果 $p = p_1 p_2$, 那么有 $p \mid p_1 p_2$. 由于 p 是素元，所以 $p \mid p_1$ 或者 $p \mid p_2$ 。不失一般性，假设 $p \mid p_1$ ，于是存在 $k \in R$, 使得 $p_1 = kp = kp_1p_2 = (kp_2)p_1$ (运用 R 上的乘法交换律和结合律)。由于 R 是一个整环（这里略去证明）没有零因子，所以 $p_1 = (kp_2)p_1 \Rightarrow (kp_2 - 1)p_1 = 0 \Rightarrow kp_2 = 1$, 这就证明了 p_2 是可逆元。

(2) **命题 1** 主理想整环 R 上的不可分解元都是素元。

证明. 假设 $c \in R$ 是一个不可分解元。对于主理想整环 R 上的任意理想 (a), 如果 $(c) \subset (a) \subset R$, 那么存在 $k \in R, c = ka$. 由于 c 是不可分解的，所以 k 是可逆元即 $(c) = (a)$ 或者 a 是可逆元即 $(a) = R$ 。这就验证了 (c) 是一个极大理想。

假设不可分解元 c 不是素元，那么存在非零的 $a, b \in R$ 使得 $c \mid ab$ 但 $c \nmid a$, $c \nmid b$.

$c \nmid a \Rightarrow a \notin (c) \Rightarrow (c) \subset (a, c) \subset R$ 且 $(c) \ne (a, c)$. 其中 (a, c) 表示由 a, c 生成的理想。由于 (c) 是极大的，所以 $(a, c) = R$. 所以存在 $x, y \in R$ 使得 $ax + cy = 1$; 同理，存在 $n, m \in R$ 使得 $bn + cm = 1$. 稍做变换可以得到 $ab \cdot xn + c \cdot (y + m - ymc) = 1$. 说明 ab 与 c 生成的理想 $(ab, c) = (1) = R$. 但由于 $c \mid ab$ 所以 $(ab, c) = (c)$. 这就导出了 c 是单位元的平凡情形。所以 c 不是素元的假设不成立。 □

命题 2 欧几里得整环都是主理想整环。

证明. 设 I 是一欧几里得整环 R 上的理想，设 $\phi : R \to \mathbb{N}$ 是定义在这一欧几里得环上的度量。可以从 I 中选取出度量最小的元素 $a \in I$. 对于任意的 $b \in I$, 由于在欧几里得环上存在 $q, r \in R$ 使得 $b = q \cdot a + r$, 其中 $r = 0$ 或者 $\phi(r) < \phi(a)$. 显然 $r = b - q \cdot a \in I$, 所以 $\phi(r) < \phi(a)$ 不能成立，因此 $r = 0$。这就说明了 $I \subset (a) \subset I$, 即 $I = (a)$ 是可由 a 生成的主理想。 □

由命题 1、2 知，只需要验证 $R \in \{\mathbb{Z}, \mathbb{F}[x], \mathbb{Z}[i]\}$ 是欧几里得环。其中 $\mathbb{Z}, \mathbb{F}[x]$ 是十分常见的欧几里得环，这里略去验证，只验证 $\mathbb{Z}[i]$ 是欧几里得环:

证明. 定义度量 $\phi : \mathbb{Z}[i] \to \mathbb{N}, \phi(a) = |a|^2 = a \cdot \bar{a}$. 对于任意非零的 $a, b \in \mathbb{Z}[i] \subset \mathbb{Q}[i]$, $\dfrac{a}{b} = \dfrac{a\bar{b}}{b\bar{b}} = x + yi$, 其中 $x, y \in \mathbb{Q}$. 取距离 x, y 最近的整数 m, n, 有 $|m - x| \le 0.5, |n - y| \le 0.5$. 构造 $q = m + ni \in \mathbb{Z}[i]$, $r = a - qb \in \mathbb{Z}[i]$, 使得 $a = qb + r$, 且其中 $r = 0$ 或者 $\phi(r) = \phi(((x - m) + (y - n)i) \cdot b) = \phi((x - m) + (y - n)i) \cdot \phi(b) \le 0.5 \cdot \phi(b) < \phi(b)$. □

(3) $(2 + \sqrt{-5}) \nmid 3$ 但 $(2 + \sqrt{-5}) \mid 3 \times 3$. 所以 $2 + \sqrt{-5}$ 不是素元。同理 $2 - \sqrt{-5}$, 3 都不是素元。考虑到在 $\mathbb{Z}[\sqrt{-5}]$ 上复数的模长的相关定义和性质仍然成立，故枚举 $3, 2 + \sqrt{-5}, 2 - \sqrt{-5}$ 可能的因子时，

只需要考虑模长平方小于等于 9 的，即只考虑 $a+b\sqrt{-5}\in\mathbb{Z}[\sqrt{-5}]$ 其中 $a,b\in\mathbb{Z},a^2+5b^2\leq 9$. 简单的穷举即可验证他们都是不可分解元。

9. 2.1 - 习题 3

对矩阵 A 模拟高斯消元过程可以知道，需要将原第 2 行置于第 1 行，将原第 1 行置于第 2 行，于是可以得到可行的置换矩阵 $P=\begin{pmatrix}0&1&0\\1&0&0\\0&0&1\end{pmatrix}$

对 PA 模拟高斯消元过程可得:

$$\begin{aligned}PA=\begin{pmatrix}1&2&3\\0&1&1\\1&2&4\end{pmatrix}&=\begin{pmatrix}1&0&0\\0&1&0\\-1&0&1\end{pmatrix}^{-1}\times\begin{pmatrix}1&2&3\\0&1&1\\0&0&1\end{pmatrix}\\&=\begin{pmatrix}1&0&0\\0&1&0\\1&0&1\end{pmatrix}\times\begin{pmatrix}1&2&3\\0&1&1\\0&0&1\end{pmatrix}\end{aligned}$$

于是 $L=\begin{pmatrix}1&0&0\\0&1&0\\1&0&1\end{pmatrix},U=\begin{pmatrix}1&2&3\\0&1&1\\0&0&1\end{pmatrix}$.

10. 2.2 - 习题 1

交换 1,3 两列,并选取前两列作为列向量的极大无关组。有:$P=\begin{pmatrix}0&0&1&0\\0&1&0&0\\1&0&0&0\\0&0&0&1\end{pmatrix}$ $AP=\begin{pmatrix}1&0&2&1\\-1&-1&-1&-1\\0&-1&1&0\\-1&-2&0&-1\end{pmatrix}$

对前两列做 Schmidt 正交化，再单位化，得到:

$$AP=\begin{pmatrix}\frac{1}{\sqrt{3}}&\frac{1}{\sqrt{3}}\\-\frac{1}{\sqrt{3}}&0\\0&\frac{1}{\sqrt{3}}\\-\frac{1}{\sqrt{3}}&\frac{1}{\sqrt{3}}\end{pmatrix}\times\begin{pmatrix}\sqrt{3}&\sqrt{3}&\sqrt{3}&\sqrt{3}\\0&-\sqrt{3}&\sqrt{3}&0\end{pmatrix}=QR$$

11. 2.2 - 习题 2

$A=\begin{pmatrix}2&0&1&1\\-1&-1&-1&-1\\1&-1&0&0\\0&-2&-1&-1\end{pmatrix}$. 对于 A 的第一列，$A_{:,1}=(-2,-1,1,0)^\top$, 我们希望构造酉矩阵 U_1 使得 $U_1A_{:,1}=(||A_{:,1}||,0,0,0)^\top$. 根据镜面反射矩阵的相关性质，设单位向量 $\beta=\frac{A_{:,1}}{||A_{:,1}||}$，$\epsilon=(1,0,0,0)^\top$, 则

可构造 $U_1 = E - \dfrac{2}{||\beta-\epsilon||^2}(\beta\beta^\top - \epsilon\beta^\top - \beta\epsilon^\top + \epsilon\epsilon^\top)$. 计算过程较繁，这里给出化简结果:

$$U_1 = \frac{1}{6-2\sqrt{6}} \times \begin{pmatrix} -4+2\sqrt{6} & 2-\sqrt{6} & -2+\sqrt{6} & 0 \\ 2-\sqrt{6} & 5-2\sqrt{6} & 1 & 0 \\ -2+\sqrt{6} & 1 & 5-2\sqrt{6} & 0 \\ 0 & 0 & 0 & 6-2\sqrt{6} \end{pmatrix}$$

计算得到

$$U_1A = \begin{pmatrix} \sqrt{6} & 0 & \sqrt{6}/2 & \sqrt{6}/2 \\ 0 & -1 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & -1 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & -2 & -1 & -1 \end{pmatrix}$$

完全类似地，取 $\beta = \dfrac{1}{\sqrt{6}}(-1,-1,-2)^\top$, $\epsilon = (1,0,0)^\top$, 计算得到 3 阶的 U_2，

$$U_2 = \frac{1}{6+\sqrt{6}} \times \begin{pmatrix} -1-\sqrt{6} & -1-\sqrt{6} & -2-2\sqrt{6} \\ -1-\sqrt{6} & 5+\sqrt{6} & -2 \\ -2-2\sqrt{6} & -2 & 2+\sqrt{6} \end{pmatrix}$$

那么 $U_2A_{1:3,1:3} = \begin{pmatrix} \sqrt{6} & \frac{\sqrt{6}}{2} & \frac{\sqrt{6}}{2} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$，已经是上三角矩阵。这就找到了 A 的第二广义 QR 分解，$A = QR$，其中 $R = \begin{pmatrix} \sqrt{6} & 0 & \sqrt{6}/2 & \sqrt{6}/2 \\ 0 & \sqrt{6} & \sqrt{6}/2 & \sqrt{6}/2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$, $Q = \left(\begin{pmatrix} 1 & 0 \\ 0 & U_2 \end{pmatrix} U_1\right)^{-1} = U_1^{-1}\begin{pmatrix} 1 & 0 \\ 0 & U_2^{-1} \end{pmatrix}$.

12. 2.2 - 习题 3

定理 2.2.3 可推广至非方阵的情形:

任一矩阵 $A_{n\times m}$ 具有 QR-分解，其中 Q 是 n 阶酉矩阵，而 R 是 $n\times m$ 的上三角矩阵，且主对角线元素是非负实数。

证明过程，可以完全仿照定理 2.2.3，即不断运用引理 2.2.2 逐一消去矩阵 $A_{n\times m}$ 的列。矩阵是否是方阵，完全不影响引理 2.2.2 的使用。

唯一的区别在于，当矩阵 $A_{n\times m}$ 消去 $\min\{n,m\}$ 列后，不能再继续像定理 2.2.3 证明中那样对剩余的子矩阵分块，所以至多只能消掉前 $\min\{n,m\}$ 列中对角线以下的部分，但这不影响 Q 是酉矩阵以及 R 是非方形的上三角矩阵且对角线元素是非负实数。

13. 2.3 - 习题 2

$\Leftarrow$ 存在实正交矩阵 $P \in \mathbb{R}^{n\times n}$ 使得 $P^\top BP = \begin{bmatrix} -1 & 0 \\ 0 & E_{n-1} \end{bmatrix}$, 所以 $B = P\left(E_n - 2\cdot\begin{bmatrix} 1 & 0 \\ 0 & O_{n-1} \end{bmatrix}\right)P^\top$

展开即可得到：$B = E - 2p_1p_1^\top$, 其中 p_1 是实正交矩阵的第 1 列，是单位向量，这就验证了 B 是镜面反射矩阵。

$\Rightarrow$ 由于 B 是镜面反射矩阵，所以存在某个单位长度的 $\delta \in \mathbb{C}^{n\times 1}$ 使得 $B = E - 2\delta\delta^\star$。

由于 B 是实方阵，所以 $\delta\delta^\star$ 也必须是实矩阵，即 $\delta_i\delta_j \in \mathbb{R}, \forall 1 \le i, j \le n$，这要求 δ 中的各分量的辐角彼此相差 π 的整数倍，所以 δ 可以拆分作 $\delta = \mathsf{e}^{i\theta}\cdot\delta'$, 其中 δ' 是实单位向量。

仿照引理 2.2.2 的推导可知，存在实正交矩阵 P 使得 $P^\top\delta' = [1, 0, 0, \cdots, 0]^\top$. (注：引理 2.2.2 是针对复数域的，但由于这里的 δ' 是实向量，故可以按照完全相同的方法构造出实的镜面反射矩阵 P)

故 $P^\top BP = P^\top(E-2\delta\delta^\star)P = P^\top\left(E - 2(\mathsf{e}^{i\theta}\delta')(\mathsf{e}^{-i\theta}\delta'^\top)\right)P = E-2(P^\top\delta')(P^\top\delta')^\top = \mathtt{diag}\{-1, E\}$

14. 2.4 - 习题 5

(1) 由于 A 是正规矩阵, 根据正规矩阵基本定理,A 酉相似于对角阵, 即存在酉矩阵 U, $U^\star AU = \mathtt{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_n\}$, 且根据 Schur 引理的推导过程知，对角线上元素 λ_i 为 A 的特征值，在本题中他们两两互异。

所以 $A = U\Lambda U^\star$, 其中 $\Lambda = \mathtt{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_n\}$. $AB = BA \Rightarrow U\Lambda U^\star B = BU\Lambda U^\star \Rightarrow$

$\Lambda U^\star BU = U^\star BU\Lambda$. 考虑左右两边的 i, j 位置元素, 可得 $\lambda_i(U^\star BU)_{ij} = \lambda_j(U^\star BU)_{ij}$, 再由 λ_i 两两互异可知，$(U^\star BU)$ 对角线以外的元素均为 0，即 B 酉相似于对角阵，所以 B 是正规矩阵。

(2) 由 (1) 的推导可知, 任意 $B \in \mathbb{C}(A), AB = BA$, 都有 $B = U\Lambda_B U^\star$, 且对不同的 B, 对应的酉矩阵 U 都相同，都由 A 决定。$U = (u_1, u_2, \cdots, u_n)$, 则 $u_1u_1^\star, u_2u_2^\star, \cdots, u_nu_n^\star$ 构成 $\mathbb{C}(A)$ 的一组基。因为对于任意的 B, $B = \sum_i \Lambda_{Bi}u_iu_i^\star$, 且 $u_iu_i^\star$ 之间彼此正交, $\Lambda_{Bi} \in \mathbb{F}$。所以 $\mathbb{C}(A)$ 是数域 $\mathbb{F}$ 上的 n 维线性空间。加法、数乘的封闭性显然。乘法的封闭性由 $(u_iu_i^\star)(u_ju_j^\star) = \delta_{ij}u_iu_i^\star$ 保证。

15. 2.4 - 习题 6

先证明一个性质 $tr(AB) = tr(BA)$. $tr(AB) = \sum_i(AB)_{ii} = \sum_i\sum_k A_{ik}B_{ki} = \sum_k\sum_i B_{ki}A_{ik} = \sum_k(BA)_{kk} = tr(BA)$.

(1) A_n 是 Hermite 矩阵 $\Leftrightarrow tr(A^2) = tr(A^\star A)$.

$\Rightarrow$: $A^\star = A$, 故 $tr(AA) = tr(A^\star A)$ 显然成立。

$\Leftarrow$: 据 Schur 引理，存在酉矩阵 Q 使得 $QAQ^\star = U$, 其中 U 是上三角矩阵, $A = Q^\star UQ$.

于是有 $tr(AA) = tr(Q^\star UQQ^\star UQ) = tr(Q^\star UUQ) = tr(UUQQ^\star) = tr(UU)$.

类似地，$tr(A^\star A) = tr(Q^\star U^\star QQ^\star UQ) = tr(Q^\star U^\star UQ) = tr(U^\star UQQ^\star) = tr(U^\star U)$.

注意到 U 是上三角矩阵，$U^\star$ 是下三角矩阵，$tr(UU) = \sum_k U_{kk}U_{kk}$. 而 $tr(U^\star U) = \sum_{i,j}\overline{U_{ij}}U_{ij}$

由于 $tr(U^\star U)$ 得到的是实数，现考虑 $tr(UU) = \sum_k U_{kk}U_{kk}$ 的实部。容易证明 $\mathtt{Re}(U_{kk}U_{kk}) = \mathtt{Re}(U_{kk})^2 - \mathtt{Im}(U_{kk})^2 \le \mathtt{Re}(U_{kk})^2 + \mathtt{Im}(U_{kk})^2 = \overline{U_{kk}}U_{kk}$. 当且仅当 U_{kk} 为实数时取等号。所以由 $tr(UU) = \sum_k U_{kk}U_{kk} = \sum_{i,j}\overline{U_{ij}}U_{ij} = tr(U^\star U)$ 推出 U_{kk} 都是实数，且 $U_{ij} = 0, \forall i \ne j$.

所以 U 是实对角矩阵，$U^\star = U$. 所以 $A^\star = Q^\star U^\star Q = Q^\star UQ = A$, A 是 Hermite 矩阵。

(2) A, B 都是 Hermite 矩阵, $AB = BA \Leftrightarrow tr((AB)^2) = tr(A^2B^2)$.

$\Rightarrow$: $AB = BA \Rightarrow (AB)^\star = B^\star A^\star = BA = AB$, 所以 AB 是 Hermite 矩阵. 由 (1) 知，$tr((AB)^2) = tr((AB)^\star AB) = tr(BAAB) = tr(AABB) = tr(A^2B^2)$.

$\Leftarrow$: $tr((AB)^2) = tr(A^2B^2) \Rightarrow tr(ABAB) = tr(AABB) \Rightarrow tr(ABAB) = tr(BAAB) \Rightarrow tr(ABAB) = tr((AB)^\star AB)$. 由 (1) 知 AB 是 Hermite 矩阵，所以 $BA = B^\star A^\star = (AB)^\star = AB$.

2.5 习题 2 $\Rightarrow$: A 与 B 酉等价，故而存在酉矩阵 U_m, V_n 使得 $U_mAV_n = B$, $V_n^\star A^\star U_m^\star = B^\star$. 所以存在酉矩阵 $U_1 = \begin{pmatrix} 0 & U_m \\ V_n^\star & 0 \end{pmatrix}$, 使得 $U_1 \begin{pmatrix} 0 & A \\ A^\star & 0 \end{pmatrix} U_1^\star = \begin{pmatrix} 0 & B \\ B^\star & 0 \end{pmatrix}$. 故这两分块矩阵酉相似。

$\Leftarrow$: 考虑到 $A' := \begin{pmatrix} 0 & A \\ A^\star & 0 \end{pmatrix}$ 是 Hermite 矩阵, 故其显然也是正规矩阵, 酉相似于一对角阵, 即存在酉矩阵 U_A, $A' = U_A\Lambda_A U_A^\star$. 同样地, 对于 Hermite 阵 $B' = \begin{pmatrix} 0 & B \\ B^\star & 0 \end{pmatrix}$, 存在酉矩阵 U_B, $B' = U_A\Lambda_B U_B^\star$.

注意到，有 $\mathtt{det}(xE - A') = \mathtt{det}(xE - \Lambda_A)$, $\mathtt{det}(xE - B') = \mathtt{det}(xE - \Lambda_B)$.

由于 A' 酉相似于 B', 即存在酉矩阵 U，$UA'U^\star = B'$, 即 $UU_A\Lambda_A U_A^\star U^\star = U_B\Lambda_B U_B^\star$, 存在酉矩阵 $V = (U_B^\star UU_A)$ 使得 $V\Lambda_A V^\star = \Lambda_B$, 于是有 $\mathtt{det}(xE - \Lambda_A) = \mathtt{det}(xE - \Lambda_B)$.

从而得到 $\mathtt{det}(xE - A') = \mathtt{det}(xE - B')$, 利用行变换将 $xE - A'$ 与 $xE - B'$ 消为下三角分块阵得:

$x^{n-m}\mathtt{det}(x^2E_m - A_{m\times n}A^\star_{n\times m}) = x^{n-m}\mathtt{det}(x^2E_m - B_{m\times n}B^\star_{n\times m})$.

由此知 $\mathtt{det}(xE - AA^\star) = \mathtt{det}(xE - BB^\star)$, 所以 A 和 B 有相同的正奇异值, 即存在酉矩阵 $U_m^A, V_n^A, U_m^B, V_n^B$ 以及一个对角阵 $S_{m\times n}$, 使得 $A = U_m^A S_{m\times n} V_n^A$, $B = U_m^B S_{m\times n} V_n^B$, 得到 $A = (U_m^A U_m^{B\star}) B (V_n^{B\star} V_n^A)$, 这就证明了 A 酉等价于 B.

2.5 习题 3 (1) 对于任一可逆矩阵 A_n, A_n 列满秩，列空间为 $\mathbb{C}^n$，A_n 亦有行满秩性质，故 $A^\star$ 列空间也为 $\mathbb{C}^n$. 故得可逆矩阵都是 EP-阵.

对于任一正规矩阵 A_n, 其酉相似于对角阵, 即存在酉矩阵 U_n 使得 $UA = \mathtt{diag}(\lambda_1, \lambda_2, \cdots, \lambda_r, 0, \cdots, 0)U$, 其中 r 为 A 的列秩。显然地，$UA^\star = \mathtt{diag}(\bar\lambda_1, \bar\lambda_2, \cdots, \bar\lambda_r, 0, \cdots, 0)U$, 故而 $A^\star$ 的列秩也是 r.

为证明它们对应的列向量空间相同，记 $A = (\alpha_1, \cdots, \alpha_n)$, $A^\star = (\beta_1, \cdots, \beta_n)$ 存在可逆矩阵 U_1 和 U_2 使得 $UA = (U\alpha_1, \cdots, U\alpha_n) = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} U_1$, $UA^\star = (U\beta_1, \cdots, U\beta_n) = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} U_2$. 方便起见，记 $E_r = (\epsilon_1, \cdots, \epsilon_r)$. 对 A 的列向量空间上的任一向量 v, 有

$$v = (\alpha_1, \cdots, \alpha_n)\gamma = U^{-1}(U\alpha_1, \cdots, U\alpha_n)\gamma = U^{-1}(U\beta_1, \cdots, U\beta_n)U_2^{-1}U_1\gamma = (\beta_1, \cdots, \beta_n)U_2^{-1}U_1\gamma$$

，其中 $\gamma \in \mathbb{C}^n$. 所以存在 $\gamma' = U_2^{-1}U_1\gamma \in \mathbb{C}^n$, $v = (\beta_1, \cdots, \beta_n)\gamma'$, 所以 v 在 $A^\star$ 的列空间中，这就证明了 A 的列空间是 $A^\star$ 的列空间的子空间。又知二者维数都为 r, 所以两列空间相同。(或者用反之亦然说明 $A^\star$ 列空间是 A 列空间的子空间，进而说明相同)。由此证得，正规矩阵都是 EP-阵.

(2) $\Leftarrow$: 因为 B 是可逆矩阵，故可构造可逆矩阵 $Q_1 = \begin{pmatrix} B & 0 \\ 0 & E \end{pmatrix} Q^\star$, $Q_2 = \begin{pmatrix} B^\star & 0 \\ 0 & E \end{pmatrix} Q^\star$, 有 $Q^\star A = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} Q_1$, $Q^\star A^\star = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} Q_2$. 由此，可完全仿照 (1) 中对正规矩阵的讨论，证明方阵 A 是 EP-阵。

$\Rightarrow$: 对 r 秩方阵 A 作 QR 分解，$A = Q_{n\times r}R_{r\times n}$, 其中 Q 是列酉阵，$r(Q) = r(R) = r$, $Q^\star Q = E_r$. 那么 $A^\star = R^\star Q^\star$. 由 A 是 EP-阵知，$A^\star$ 的列向量可由 Q 的列向量线性组合表示，故 $A^\star = QR'$.

由此得 $R^\star Q^\star = QR'$, 故 $R' = Q^\star QR' = Q^\star R^\star Q^\star$, $A^\star = QR' = QQ^\star R^\star Q^\star$, 于是 $A = QRQQ^\star$. 关于 $x \in \mathbb{C}^n$ 的方程 $Q^\star x = 0$ 其解空间是 $n-r$ 维线性空间，从中取出一组单位正交的基，作为列向量组成矩阵 $Q' = (q_{r+1}, \cdots, q_n)$. 容易验证 $U = (Q; Q')$ 是 n 阶的酉矩阵。于是有

$$A = (QR)(QQ^\star) = \left((Q;Q') \times \begin{pmatrix} R \\ 0 \end{pmatrix}\right) \times \left((Q;0) \times \begin{pmatrix} Q^\star \\ Q'^\star \end{pmatrix}\right) = U \begin{pmatrix} RQ & 0 \\ 0 & 0 \end{pmatrix} U^\star$$

且可以验证 $r(RQ) = r(R) = r(Q) = r$, RQ 是 r 阶可逆矩阵.

2.5 习题 5 (1) 回顾奇异值分解存在性的构造证明可知, V 的后 $n-r$ 列是选取的 $A^\star Ax = 0$ 的解空间的标准正交基。由于 $Ax = 0 \Rightarrow A^\star Ax = 0$, 所以 $Ax = 0$ 的解空间是 $A^\star Ax = 0$ 解空间的子线性空间。又知二者均为 $n-r$ 维，故 $Ax = 0$ 与 $A^\star Ax = 0$ 二者解空间相同，故而 V 的后 $n-r$ 列也是 A 的解空间的一组标准正交基。

(2) $AV = U\Lambda$. 其中 $\Lambda = \begin{pmatrix} \Lambda_r & 0 \\ 0 & 0 \end{pmatrix}$. 故可知 U 的前 r 列是 AV 的列空间的一组基，且由于是奇异值分解所以它是标准正交的。又由于 V 是酉矩阵，故而 AV 和 A 的列空间相同，这就说明了 U 的前 r 列是 A 的列空间的一个标准正交基。

(3) 回顾奇异值分解存在性的构造证明可知，对于 $U = (\alpha_1, \cdots, \alpha_n)$ 的后 $n-r$ 列中的任意一列，比如 $\alpha_i, i > r$, 有 $AA^\star \alpha_i = \lambda_i \alpha_i = 0$ 因为 $\lambda_i = 0, \forall i > r$. 仿 (1) 可以得到 $AA^\star$ 和 $A^\star$ 的解空间相同，故而 U 的后 $n-r$ 列是 $A^\star$ 解空间的子线性空间。再由二者维数相同知，U 的后 $n-r$ 列是该解空间的一组基，且由于 $A = U\Lambda V^\star$ 是 SVD 分解，所以是标准正交的。

(4) $A^\star = V\Lambda^\star U^\star$, 故 $A^\star U = V\Lambda^\star$, 同 (2) 可证。

2.6 习题 1 求 $A = \begin{pmatrix} 1 & 1 & 1 & 0 & 2 \\ 1 & 0 & 1 & 1 & 3 \\ 0 & 1 & 1 & 1 & 4 \end{pmatrix}$ 的 MP 广义逆。经计算得到：有下三角矩阵 $L = \begin{pmatrix} 1 & & \\ 8/7 & 1/7 & \\ 10/7 & 9/70 & 1/10 \end{pmatrix}$ 以及行正交的矩阵 $Q = \begin{pmatrix} 1 & 1 & 1 & 0 & 2 \\ -1 & -8 & -1 & 7 & 5 \\ -13 & 6 & -3 & 1 & 5 \end{pmatrix}$, 使得 $A = LQ$. 可以验证 $QQ^\star = \mathtt{diag}(7, 140, 240)$. 下面验证 $B = Q^\star(QQ^\star)^{-1}L^{-1}$ 是 A 的广义逆：

a. $ABA = LQQ^\star(QQ^\star)^{-1}L^{-1}LQ = LQ = A$

b. $BAB = Q^\star(QQ^\star)^{-1}L^{-1}LQQ^\star(QQ^\star)^{-1}L^{-1} = Q^\star(QQ^\star)^{-1}L^{-1} = B$

c. $AB = LQQ^\star(QQ^\star)^{-1}L^{-1} = E$ 是 Hermite 矩阵

d. $BA = Q^\star(QQ^\star)^{-1}L^{-1}LQ = Q^\star(QQ^\star)^{-1}Q$ 是 Hermite 矩阵，因为 $QQ^\star$ 是实对角阵

带入具体数值即可求得 A 的 MP 广义逆矩阵

$$B = Q^\star(QQ^\star)^{-1}L^{-1} = \begin{pmatrix} 1 & -1 & -13 \\ 1 & -8 & 6 \\ 1 & -1 & -3 \\ 0 & 7 & 1 \\ 2 & 5 & 5 \end{pmatrix} \times \begin{pmatrix} 1/7 & & \\ & 1/140 & \\ & & 1/240 \end{pmatrix} \times \begin{pmatrix} 1 & & \\ -8 & 7 & \\ -4 & -9 & 10 \end{pmatrix}$$

进一步化简得 $B = \dfrac{1}{48} \times \begin{pmatrix} 20 & 21 & -26 \\ 24 & -30 & 12 \\ 12 & 3 & -6 \\ -20 & 15 & 2 \\ -4 & 3 & 10 \end{pmatrix}$

习题 2 $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 5 \end{bmatrix}$, $AA^\top = \begin{bmatrix} 2 & 1 \\ 1 & 26 \end{bmatrix}$，容易验证这是一个正定的 Hermite 矩阵，所以它的 Cholesky 分解存在. 证明了存在性后，Cholesky 分解的具体构造较为容易，这里直接给出结果：

$L = \begin{bmatrix} \sqrt{2} & 0 \\ \sqrt{1/2} & \sqrt{51/2} \end{bmatrix}$, $AA^\top = LL^\top$.

习题 3 (1) 首先，容易验证 $AA^\star$ 是一个半正定的 Hermite 矩阵, 那么存在酉矩阵 $U_m \in \mathbb{C}^{n\times n}$，使得

$AA^\star = U_m \mathtt{diag}(\lambda_1, \ldots, \lambda_r, 0, \ldots, 0)U_m^\star$. 其中 $r = r(A) = r(A^\star) = r(AA^\star)$.

对 A 进行奇异值分解,有 $A = U_m \begin{bmatrix} \Lambda_r & 0 \\ 0 & 0 \end{bmatrix}_{m\times n} V_n$,其中 V_n 是另一个 n 阶酉矩阵, $\Lambda_r = \mathtt{diag}(\lambda_1^{1/2}, \ldots, \lambda_r^{1/2})$.

方便起见，记 $\Lambda = \mathtt{diag}(\lambda_1, \ldots, \lambda_r, 0, \ldots 0)$. 对 SVD 分解中的 $m \times n$ 准对角矩阵进行分块可得:

$$A = U_m \left[\Lambda^{1/2}; O_{m\times(n-m)}\right] V_n = U_m \Lambda^{1/2} \left[E_m; O\right] V_n = U_m \Lambda^{1/2} U_m^\star U_m [E_m; O] V_n = PU_{m\times n}$$

其中 $P = U_m \Lambda^{1/2} U_m^\star = (AA^\star)^{1/2}$, $U_{m\times n} = U_m[E_m; O]V_n = [V_m; O]V_n$, 可以验证 $UU^\star = [V_m; O]V_n V_n^\star [V_m; O]^\star = V_m V_m^\star = E_m$.

(2) $r(A) = m$ 时，$AA^\star$ 是满秩方阵，$P = (AA^\star)^{1/2}$ 也是满秩方阵，故而 $U = P^{-1}A$ 唯一确定。

习题 12 证明关于方阵 $A \in \mathbb{M}_n$ 的下列三个命题的等价性：

(1) 存在正整数 $k \geq 1$, 使得 $A^k = 0$;

(2) 对于任意正整数 $m \geq 1$, $tr(A^m) = 0$;

(3) 对于任意正整数 m, $1 \leq m \leq n$, $tr(A^m) = 0$.

证明. 为方便讨论矩阵的迹, 根据 Schur 引理，对 A 做分解: 存在酉矩阵 $U \in \mathbb{M}_n$, 和上三角矩阵 $R \in \mathbb{M}_n$ 使得 $A = URU^\star$. 从而有 $A^k = UR^kU^\star, \forall k \in \mathbb{N}^*$. 通过简单的数学归纳可以证明 R^k 是上三角矩阵，且 $(R^k)_{ii} = (R_{ii})^k$. 下面按照 (1) $\Rightarrow$ (2) $\Rightarrow$ (3) $\Rightarrow$ (1) 的顺序来证明等价性：

(1) $\Rightarrow$ (2) $A^k = UR^kU^\star = 0$, 其中酉矩阵 U 可逆, 所以 $R^k = 0$. 于是对于任意的 $1 \leq i \leq n, i \in \mathbb{N}$, $(R^k)_{ii} = (R_{ii})^k = 0$, 所以 $R_{ii} = 0$ 即 R 的对角线元素均为 0. 所以对正整数 $m \geq 1$, $tr(R^m) = \sum_i R_{ii}^m = 0$, $tr(A^m) = tr(UR^mU^\star) = tr(R^mU^\star U) = tr(R^m) = 0$.

(2) $\Rightarrow$ (3) 显然

(3) $\Rightarrow$ (1) 对于任意的 $1 \leq m \leq n$, $tr(A^m) = tr(R^m) = \sum_i R_{ii}^m = 0$. 假设 $R_{ii}, 1 \leq i \leq n$ 不全为 0, 则可取出其中不为 0 的项，去重后得到 $r_1, r_2, \cdots, r_t$, $1 \leq t \leq n$. 从而有方程组

$$\begin{bmatrix} r_1 & r_2 & \cdots & r_t \\ r_1^2 & r_2^2 & \cdots & r_t^2 \\ \vdots & \vdots & \ddots & \vdots \\ r_1^t & r_2^t & \cdots & r_t^t \end{bmatrix} \begin{bmatrix} n_1 \\ n_2 \\ \vdots \\ n_t \end{bmatrix} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ r_1^1 & r_2^1 & \cdots & r_t^1 \\ \vdots & \vdots & \ddots & \vdots \\ r_1^{t-1} & r_2^{t-1} & \cdots & r_t^{t-1} \end{bmatrix} \begin{bmatrix} r_1 & & & \\ & r_2 & & \\ & & \ddots & \\ & & & r_t \end{bmatrix} \begin{bmatrix} n_1 \\ n_2 \\ \vdots \\ n_t \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

其中 n_i 表示 r_i 去重前的出现次数，应有 $r_i \neq r_j, \forall i \neq j$ 以及 $n_i > 0, r_i \neq 0, \forall i$. 故而此时方程中的 Vandermonde 矩阵和对角矩阵均可逆，从而 $[n_1, n_2, \cdots, n_t]^\top$ 应为 0 向量，矛盾。所以 R_{ii} 不全为 0 的假设不成立，从而得到 R 是对角线全为 0 的上三角矩阵。即 $R_{ij} = 0, \forall i < j + 1$.

下归纳证明 $(R^k)_{ij} = 0, \forall i < j + k$：(1) 对于 $k = 1$ 成立；(2) $(R^k)_{ij} = 0, \forall i < j + k \Rightarrow (R^{k+1})_{ij} = (R^kR)_{ij} = \sum_t (R^k)_{it} R_{tj}$. 当 $i < j + k + 1$ 时, $i \geq t + k$ 与 $t \geq j + 1$ 不能同时成立，故 $(R^k)_{it}$ 与 R_{tj} 中至少有一个为 0, 从而推出 $(R^{k+1})_{ij} = 0, \forall i < j + (k + 1)$.

对于任意 $1 \leq i, j \leq n$, 有 $i < j + n$, 于是 $(R^n)_{ij} = 0$, 所以 $R^n = 0$, $A^n = UR^nU^\star = 0$.

注：这还说明了，如果存在正整数 k 使得 $A^k = 0$, 那么存在 $k \leq n$ 使得 $A^k = 0$.

□

1. 习题 1

$\sigma : A \to AB^\top + BA$, 注意到 $BA = (A^\top B^\top)^\top = (AB^\top)^\top$, 所以 $AB^\top + BA$ 是 2 阶的实对称矩阵，仍在 V 中。

(1) 对任意的 $\lambda_1, \lambda_2 \in \mathbb{R}$, $A_1, A_2 \in V$, 有 $\sigma(\lambda_1 A_1 + \lambda_2 A_2) = (\lambda_1 A_1 + \lambda_2 A_2)B^\top + B(\lambda_1 A_1 + \lambda_2 A_2) = \lambda_1 \cdot (A_1 B^\top + BA_1) + \lambda_2 \cdot (A_2 B^\top + BA_2) = \lambda_1 \sigma(A_1) + \lambda_2 \sigma(A_2)$.

(2) 由于 $\sigma(E_{11}) = 2E_{11}$, $\sigma(E_{12} + E_{21}) = -2E_{11} + (E_{12} + E_{21})$, $\sigma(E_{22}) = -(E_{12} + E_{21})$. 故 σ 在这组基下，对应于矩阵 $\begin{bmatrix} 2 & -2 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$

(3) 对于任意 $A \in V$, $A = [E_{11}, E_{12} + E_{21}, E_{22}] \cdot [c_1, c_2, c_3]^\top, c_i \in \mathbb{R}$,

$\sigma A = \sigma[E_{11}, E_{12} + E_{21}, E_{22}] \cdot [c_1, c_2, c_3]^\top = [E_{11}, E_{12} + E_{21}, E_{22}] \cdot \begin{bmatrix} 2 & -2 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \cdot [c_1, c_2, c_3]^\top = [E_{11}, E_{12} + E_{21}, E_{22}] \cdot [2c_1 - 2c_2, c_2 - c_3, 0]^\top$, 由此可知 $\{E_{11}, E_{12} + E_{21}\}$ 是像子空间 $\mathtt{im}(\sigma)$ 的一组基。

(4) 延用 (3) 中的记号，由于 $E_{11}, E_{12}+E_{21}, E_{22}$ 线性无关，所以 $\sigma A=0$ 当且仅当 $\begin{bmatrix} 2 & -2 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}=$ 0, 解为 $[c_1,c_2,c_3]^\top=[c_3,c_3,c_3]^\top, c_3\in\mathbb{R}$. 故核子空间 $\mathtt{ker}(\sigma)$ 的一组基是 $\{E_{11}+E_{12}+E_{21}+E_{22}\}$.

(5) 是直和，且 $V=\mathtt{im}(\sigma)\oplus\mathtt{ker}(\sigma)$. 因为 $\mathtt{im}(\sigma)$ 与 $\mathtt{ker}(\sigma)$ 的基不交，且它们基的并是 V 的一组基。

2. 习题 3

设 $\{\eta_1,\eta_2,\cdots,\eta_n\}$ 是 n 维空间 $\mathbb{F}^n$ 的一组基。对于线性变换 σ, 以及任意的 $\alpha\in\mathbb{F}^n$, $\alpha=[\eta_1,\eta_2,\cdots,\eta_n]\cdot c^\top$, 其中 $c=[c_1,c_2,\cdots,c_n]\in\mathbb{F}^n$. $\sigma(\alpha)=\sigma[\eta_1,\eta_2,\cdots,\eta_n]\cdot c^\top$. 由于 $\{\eta_i\}$ 是基，故存在矩阵 A 使得 $\sigma[\eta_1,\cdots,\eta_n]=[\eta_1,\cdots,\eta_n]\cdot A$. 再记 $N=[\eta_1,\eta_2,\cdots,\eta_n]$, 显然 N 可逆，构造 $B=NAN^{-1}\in M_n(\mathbb{F})$, 有 $B\alpha=B(Nc^\top)=NAN^{-1}Nc^\top=NAc^\top=\sigma(\alpha)$.

3. 习题 4

(1) $V=W_1\oplus W_2$, $\sigma:\alpha_1+\alpha_2\to\alpha_1$, 其中 $\alpha_i\in W_i$. 取 W_1 的一组基 $(\alpha_1,\cdots,\alpha_r)$; 取 W_2 的一组基 $(\beta_1,\cdots,\beta_s)$. 那么 $(\alpha_1,\cdots,\alpha_r,\beta_1,\cdots,\beta_s)$ 是 V 的一组基。对于 V 的任意基 $(\gamma,\cdots,\gamma_r,\gamma_{r+1},\cdots,\gamma_{r+s})$, 有 $(\gamma,\cdots,\gamma_r,\gamma_{r+1},\cdots,\gamma_{r+s})=(\alpha_1,\cdots,\alpha_r,\beta_1,\cdots,\beta_s)C$, 其中 $C\in M_{r+s}(\mathbb{F})$. 将 σ 作用于其上:

$$\begin{aligned}\sigma(\gamma,\cdots,\gamma_r,\gamma_{r+1},\cdots,\gamma_{r+s}) &= \sigma(\alpha_1,\cdots,\alpha_r,\beta_1,\cdots,\beta_s)\cdot C \\ &= (\alpha_1,\cdots,\alpha_r,\beta_1,\cdots,\beta_s)\begin{bmatrix} E_r & 0 \\ 0 & 0 \end{bmatrix}C \\ &= (\gamma,\cdots,\gamma_r,\gamma_{r+1},\cdots,\gamma_{r+s})\cdot C^{-1}\begin{bmatrix} E_r & 0 \\ 0 & 0 \end{bmatrix}C\end{aligned}$$

从而得到 $A=C^{-1}\begin{bmatrix} E_r & 0 \\ 0 & 0 \end{bmatrix}C$, 显然有 $A^2=A$.

(2) (a) 即 $\sigma_A:(\alpha_1,\cdots,\alpha_n)\beta\to(\alpha_1,\cdots,\alpha_n)(A\beta)$. 容易验证对任意 $\lambda_1,\lambda_2\in\mathbb{F}$, 以及任意 $v_1,v_2\in V$, 有 $\exists\beta_1, v_1=(\alpha_1,\cdots,\alpha_n)\beta_1$, $\exists\beta_2, v_2=(\alpha_1,\cdots,\alpha_n)\beta_2$，$\sigma(\lambda_1v_1+\lambda_2v_2)=(\alpha_1,\cdots,\alpha_n)(A\lambda_1\beta_1+A\lambda_2\beta_2)=\lambda_1(\alpha_1,\cdots,\alpha_n)(A\beta_1)+\lambda_2(\alpha_1,\cdots,\alpha_n)(A\beta_2)=\lambda_1\sigma(v_1)+\lambda_2\sigma(v_2)$, 所以 σ 是线性变换

(b) $A^2=A\Rightarrow A(A-E)=(A-E)A=0$. 下考察 $r(A-E)$ 与 $r(A)$ 的关系. 设 $r(A)=r$, 则存在可逆矩阵 $P,Q\in M_n(\mathbb{F})$ 使得 $A=P\begin{bmatrix} E_r & 0 \\ 0 & 0 \end{bmatrix}Q$, 由于 $A^2=P\begin{bmatrix} E_r & 0 \\ 0 & 0 \end{bmatrix}QP\begin{bmatrix} E_r & 0 \\ 0 & 0 \end{bmatrix}Q=A$, 所以 QP 具有 $QP=\begin{bmatrix} E_r & C_{12} \\ C_{21} & C_{22} \end{bmatrix}$ 的形式。$A-E=P\begin{bmatrix} E_r & 0 \\ 0 & 0 \end{bmatrix}Q-PP^{-1}Q^{-1}Q=P\left(\begin{bmatrix} E_r & 0 \\ 0 & 0 \end{bmatrix}-\begin{bmatrix} E_r & C_{12} \\ C_{21} & C_{22} \end{bmatrix}^{-1}\right)Q$. 注意到 $\begin{bmatrix} E_r & C_{12} \\ C_{21} & C_{22} \end{bmatrix}^{-1}=\begin{bmatrix} E_r+C_{12}D^{-1}C_{21} & -C_{12}D^{-1} \\ -D^{-1}C_{21} & D^{-1} \end{bmatrix}$, 其中 $D=C_{22}-C_{21}C_{12}$ 是 $(n-r)$ 阶可逆矩阵 (可逆性由 QP 可逆推出). 所以: $\begin{bmatrix} E_r & 0 \\ 0 & 0 \end{bmatrix}-\begin{bmatrix} E_r & C_{12} \\ C_{21} & C_{22} \end{bmatrix}^{-1}=\begin{bmatrix} -C_{12} \\ E_{n-r} \end{bmatrix}D^{-1}[C_{21};-E_{n-r}]$, 从而 $r(A-E)=n-r$.

下面验证 A 的列向量空间和 $A-E$ 的解空间相同，即 $\texttt{col}(A)=(A-E)^{\perp}$. $v\in\texttt{col}(A)\Rightarrow\exists\gamma,v=A\gamma\Rightarrow(A-E)v=(A-E)A\gamma=0\Rightarrow v\in(A-E)^{\perp}$, 故 $\texttt{col}(A)\subset(A-E)^{\perp}$. 又 $\dim(\texttt{col}(A))=r(A)=r=n-r(A-E)=\dim\left((A-E)^{\perp}\right)$, 所以 $\texttt{col}(A)=(A-E)^{\perp}$. 同理可得，$\texttt{col}(A-E)=A^{\perp}$.

令 $W_1=\{(\alpha_1,\cdots,\alpha_n)\beta|\forall\beta\in A^{\perp}\}$, $W_2=\{(\alpha_1,\cdots,\alpha_n)\beta|\forall\beta\in(A-E)^{\perp}\}$ (或等价地，$W_1=\{(\alpha_1,\cdots,\alpha_n)\beta|\forall\beta\in\texttt{col}(A-E)\}$, $W_2=\{(\alpha_1,\cdots,\alpha_n)\beta|\forall\beta\in\texttt{col}(A)\}$), 则 $V=W_1\oplus W_2$ 是满足题意的一个直和分解，验证如下：

(i) 设 $v=(\alpha_1,\cdots,\alpha_n)\beta\in W_1\cap W_2$ 则 $\beta\in A^{\perp}\cap(A-E)^{\perp}$, 于是 $\beta=A\beta=0$, $x=0$. 所以 $W_1\cap W_2=0$.

(ii) $\dim(W_1)+\dim(W_2)=\dim(A^{\perp})+\dim((A-E)^{\perp})=(n-r)+r=n=\dim(V)$, 结合 (i) 知，$V=W_1\oplus W_2$.

(iii) $\forall\alpha=(\alpha_1,\cdots,\alpha_n)\beta\in W_1$, 有 $\sigma\alpha=(\alpha_1,\cdots,\alpha_n)(A\beta)=0$ 因为 $\beta\in A^{\perp},A\beta=0$. 类似地，$\forall\alpha=(\alpha_1,\cdots,\alpha_n)\beta\in W_2$, 有 $\sigma\alpha=(\alpha_1,\cdots,\alpha_n)(A\beta)=\alpha$ 因为 $\beta\in(A-E)^{\perp},A\beta=\beta$.

(好像 W_1,W_2 和题目中的顺序反了，算了不改了)

4. 习题 5

设线性变换在基 $\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$ 下的对应的矩阵为 A, 在基 $\{\beta_1,\beta_2,\cdots,\beta_n\}$ 下对应的矩阵为 B. 根据定义有：$\sigma(\alpha_1,\cdots,\alpha_n)=(\alpha_1,\cdots,\alpha_n)A$, $\sigma(\beta_1,\cdots,\beta_n)=(\beta_1,\cdots,\beta_n)B$. 考虑到同一线性空间的基之间能互相表示，故存在可逆矩阵 Q 使得 $(\beta_1,\cdots,\beta_n)=(\alpha_1,\cdots,\alpha_n)Q$, 从而有

$$\begin{aligned}\sigma(\beta_1,\cdots,\beta_n)\quad&=(\beta_1,\cdots,\beta_n)B=(\alpha_1,\cdots,\alpha_n)QB\\&=\sigma[(\alpha_1,\cdots,\alpha_n)Q]=(\sigma\alpha_1,\cdots,\sigma\alpha_n)Q=(\alpha_1,\cdots,\alpha_n)AQ\end{aligned}$$

故而 $(\alpha_1,\cdots,\alpha_n)(QB-AQ)=0$, 由于基线性无关，故 $QB-AQ=0\Rightarrow A=QBQ^{-1}$, 即 A 与 B 相似。

1. 习题 2

$\Rightarrow$: 不妨设 $\sigma:V=W\oplus W^{\perp}\to W$, 其中 $(w_1,\cdots,w_r)$ 是 W 子线性空间的一组标准正交基, $(w_{r+1},\cdots,w_n)$ 是 $W^{\perp}$ 上的一组标准正交基，显然 $(w_1,\cdots,w_r,w_{r+1},\cdots w_n)$ 构成 $V=W\oplus W^{\perp}$ 的一组标准正交基。对于 V 上的任意一组标准正交基 $\alpha_1,\cdots,\alpha_n$, 存在可逆矩阵 Q 使得 $(\alpha_1,\cdots,\alpha_n)=(w_1,\cdots,w_n)Q$, 由于两组都是标准正交基，故而 Q 是酉矩阵。考虑设 σ 在标准正交基 $\{\alpha_i\}$ 下对应矩阵 A, 则有

$$\begin{aligned}(\alpha_1,\cdots,\alpha_n)A=\sigma(\alpha_1,\cdots,\alpha_n)\quad&=\sigma(w_1,\cdots,w_r,w_{r+1},\cdots,w_n)Q\\&=(w_1,\cdots,w_r,w_{r+1},\cdots,w_n)\begin{bmatrix}E_r&0\\0&0\end{bmatrix}Q\\&=(\alpha_1,\cdots,\alpha_n)Q^{\star}\begin{bmatrix}E_r&0\\0&0\end{bmatrix}Q\end{aligned}$$

所以 $A=Q^{\star}\begin{bmatrix}E_r&0\\0&0\end{bmatrix}Q$ 是 Hermite 矩阵。又 $A^2=Q^{\star}\begin{bmatrix}E_r&0\\0&0\end{bmatrix}QQ^{\star}\begin{bmatrix}E_r&0\\0&0\end{bmatrix}Q=Q^{\star}\begin{bmatrix}E_r&0\\0&0\end{bmatrix}Q=A$, 所以 A 是幂等的。

$\Leftarrow$: 不妨设 σ 在某组标准正交基 $\alpha_1,\cdots,\alpha_n$ 下对应矩阵 A, A 是幂等的 Hermite 矩阵。对于 Hermite 矩阵 A, 存在酉矩阵 $U=(u_1,\cdots,u_n)$ 以及实对角矩阵 Λ, 使得 $A=U\Lambda U^{\star}$. 由于是幂等的,$A^2=U\Lambda^2U^{\star}=$

$A=U\Lambda U^{\star}$, 即 $\Lambda^2=\Lambda$. 所以 Λ 对角线元素为 1 或 0. 不妨设 $\Lambda=\begin{bmatrix} E_r & 0 \\ 0 & 0 \end{bmatrix}$, 则 $A=\sum_{i=1}^{r}u_iu_i^{\star}$.
容易验证 $\forall i\le r, Au_i=u_i, \forall i>r, Au_i=0$.

构造 $W=\{(\alpha_1,\cdots,\alpha_n)x|x=\sum_{i=1}^{r}c_iu_i, c_i\in\mathbb{F}\}$, 从而 $W^{\perp}=\{(\alpha_1,\cdots,\alpha_n)x|x=\sum_{i=r+1}^{n}c_iu_i, c_i\in\mathbb{F}\}$, 容易验证 $V=W\oplus W^{\perp}$, 且 σ 是从 V 到 W 的一个正交投影变换, 即 $\forall v\in W, \sigma v=v$, $\forall v\in W^{\perp}, \sigma v=0$.

2. 习题 5

(1) $[A,B]:=tr(A^{\star}B)=\sum_i(A^{\star}B)_{ii}=\sum_i\sum_k(A^{\star})_{ik}B_{ik}=\sum_i\sum_k\overline{A_{ki}}B_{ki}$. 容易验证:

* 对称性：$[B,A]=\sum_i\sum_k\overline{B_{ki}}A_{ki}=\sum_i\sum_k\overline{\overline{A_{ki}}B_{ki}}=\overline{[A,B]}$.
* 线性性：$[A,\alpha C+\beta D]=\sum_i\sum_k\overline{A_{ki}}(\alpha C_{ki}+\beta D_{ki})=\alpha\sum_i\sum_k\overline{A_{ki}}C_{ki}+\beta\sum_i\sum_k\overline{A_{ki}}D_{ki}=\alpha[A,C]+\beta[A,D]$.
* 正定性：$[A,A]=\sum_i\sum_k\overline{A_{ki}}A_{ki}=\sum_i\sum_k|A_{ki}|_2^2\ge 0$, 当且仅当 $A=0$ 时取等号。

(2) $E_{11}=\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, ee^{\top}=\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$. 记夹角为 θ, 有 $\cos\theta=[E_{11},ee^{\top}]/\sqrt{[E_{11},E_{11}][ee^{\top},ee^{\top}]}=1/\sqrt{1\times 16}=1/4$, 即 $\theta=\arccos(1/4)$.

该内积空间的一个标准正交基 $\{E_{ij}|i,j\in\{1,2,3,4\}\}$.

3. 习题 6

(1) 容易验证 $E_{11},E_{12}+E_{21},E_{22}$ 是 W 的一组基。所以 $B\in W^{\perp}$, 当且仅当 $[E_{11},B]=0,[E_{12}+E_{21},B]=0,[E_{22},B]=0$. 设 $B=b_{11}E_{11}+b_{12}E_{12}+b_{21}E_{21}+b_{22}E_{22}$, 则有

$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} b_{11} \\ b_{12} \\ b_{21} \\ b_{22} \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$, 所以 $[b_{11},b_{12},b_{21},b_{22}]^{\top}=[0,t,-t,0]^{\top}, t\in\mathbb{R}$.

所以正交补子空间 $W^{\perp}=\{t(E_{12}-E_{21})|t\in\mathbb{R}\}$.

(2) 由于 $\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}=A+B=\begin{bmatrix} 1 & 0.5 \\ 0.5 & 0 \end{bmatrix}+\begin{bmatrix} 0 & 0.5 \\ -0.5 & 0 \end{bmatrix}$, 其中 $A\in W, B\in W^{\perp}$, 所以在 W 上的正交投影为 A, 即 $\begin{bmatrix} 1 & 0.5 \\ 0.5 & 0 \end{bmatrix}$.

4. 补充例 4.13

对于 $\forall\beta\in\texttt{im}(g)$, 存在 $\gamma\in V$, $g\gamma=v$. 所以 $f\beta=f(g\gamma)=(fg)\gamma=0$, 即 $\beta\in\texttt{ker}(f)$, 故 $\texttt{im}(g)\subseteq\texttt{ker}(f)$.

根据课上讲到的结论 (或者考查 g 在某组基下对应的矩阵 A 并运用作业 7 中证明的 55 页习题 4 的结论) 有: $r(g)+r(I_V-g)=n$. 容易验证 $\texttt{im}(I_V-g)\subseteq\texttt{ker}(g)$, 因为对任意 $(I_V-g)\gamma, \gamma\in V$ 有 $g(I_V-g)\gamma=0$. 而 $\dim\texttt{ker}(g)=n-r(g)=r(I_V-g)=\dim\texttt{im}(I_V-g)$, 所以有 $\texttt{ker}(g)=\texttt{im}(I_V-g)$.

考虑任意 $x \in \mathtt{im}(g) \cap \mathtt{im}(I_V - g) = \mathtt{im}(g) \cap \mathtt{ker}(g)$, 有 $x = g\gamma$, 且 $0 = gx = g^2\gamma = g\gamma = x$. 所以 $\mathtt{im}(g) \cap \mathtt{im}(I_V - g) = \{0\}$. 所以有 $V = \mathtt{im}(g) \oplus \mathtt{im}(I_V - g)$.

(1) 对于任意 $v \in V - \mathtt{ker}(f)$, 由于 $v \notin \mathtt{ker}(f)$, $\mathtt{im}(g) \subseteq \mathtt{ker}(f)$, 所以 $v \notin \mathtt{im}(g)$. 又 $V = \mathtt{im}(g) \oplus \mathtt{im}(I_V - g)$, 所以 $v \in \mathtt{im}(I_V - g) = \mathtt{ker}(g)$. 这就验证了 $V = \mathtt{ker}(f) + \mathtt{ker}(g)$.

(2) $\Rightarrow$: 若 $V = \mathtt{ker}(f) \oplus \mathtt{ker}(g)$, 则 $n = \dim \mathtt{ker}(f) + \dim \mathtt{ker}(g) = n - r(f) + n - r(g)$, 故 $r(f) + r(g) = n$.

$\Leftarrow$: 若 $r(f) + r(g) = n$, 则 $\dim \mathtt{ker}(f) = n - r(f) = r(g) = \dim \mathtt{im}(g)$. 又 $\mathtt{im}(g) \subseteq \mathtt{ker}(f)$, 故此时 $\mathtt{ker}(f) = \mathtt{im}(g)$. 又因为 $\mathtt{ker}(g) = \mathtt{im}(I_V - g)$, 所以 $V = \mathtt{im}(g) \oplus \mathtt{im}(I_V - g) = \mathtt{ker}(f) \oplus \mathtt{ker}(g)$.

1. 59 页习题 4

Lemma 0.1. 正交矩阵可以由若干镜面反射矩阵相成得到

证明. 设 $Q \in M_n(\mathbb{R})$, $Q = (q_1, q_2, \cdots, q_n)$. 由于 $(q_1, q_1) = 1$, 故存在镜面反射矩阵 U_1 使得 $U_1 q_1 = e_1 := (1, 0, \cdots, 0)^\top$. 对于任意 $j \neq 1$, $(U_1 q_j, U_1 q_1) = q_j^\top U_1 U_1 q_1 = q_j^\top q_1 = 0$, 所以 $U_1 q_j = (0, *, \cdots, *)^\top$, 即第一个分量必然为 0. 从而有

$$U_1 Q = (U_1 q_1, U_1 q_2, \cdots, U_1 q_n) = \begin{bmatrix} 1 & 0 \\ 0 & Q' \end{bmatrix}$$

由于镜面反射变换不改变内积，即 $(U_1 q_i, U_1 q_j) = (q_i, q_j)$，故 Q' 是 $n-1$ 阶的正交矩阵。归纳地进行下去即可得到 $U_{n-1} \cdots U_2 U_1 Q = E_n$, 即 $Q = U_1 U_2 \cdots U_{n-1}$. $\square$

对于正交变换 σ 它在 V 的一组标准正交基 $\eta_1, \cdots, \eta_n$ 下对应于矩阵 Q, 容易验证 Q 是正交矩阵。

那么 $\sigma(\eta_1, \cdots, \eta_n) = (\eta_1, \cdots, \eta_n) Q = (\eta_1, \cdots, \eta_n) U_1 U_2 \cdots U_{n-1}$, 其中 U_i 是镜面反射矩阵

在标准正交基 $\eta_1, \cdots, \eta_n$ 下, 镜面反射矩阵 U_1 对应于镜面反射变换 σ_1, 于是 $(\eta_1, \cdots, \eta_n) U_1 U_2 \cdots U_{n-1} = \sigma_1(\eta_1, \cdots, \eta_n) U_2 \cdots U_{n-1}$. 由于镜面反射变换不改变内积, 故而 $\sigma_1(\eta_1, \cdots, \eta_n)$ 仍是一组标准正交基, 不妨设 U_2 在 $\sigma_1\{\eta_i\}$ 下对应于镜面反射变换 σ_2, 则有 $(\eta_1, \cdots, \eta_n) U_1 U_2 \cdots U_{n-1} = (\sigma_2 \sigma_1)(\eta_1, \cdots, \eta_n) U_3 \cdots U_{n-1}$. 归纳地进行下去，假设镜面反射矩阵 U_{i+1} 在 $(\sigma_i \cdots \sigma_1)(\eta_1, \cdots, \eta_n)$ 这一标准正交基下对应于镜面反射变换 σ_{i+1}, 则最终得到

$$\sigma(\eta_1, \cdots, \eta_n) = (\sigma_{n-1} \sigma_{n-2} \cdots \sigma_1)(\eta_1, \cdots, \eta_n)$$

这就构造出了 $\sigma = \sigma_{n-1} \sigma_{n-2} \cdots \sigma_1$.

似乎证复杂了,实际上只需要假设 U_i 在 $(\eta_1, \cdots, \eta_n)$ 下对应于镜面反射变换 σ_i, 即可得到 $\sigma = \sigma_1 \sigma_2 \cdots \sigma_{n-1}$

2. 60 页习题 8

方便起见，这里先证明第 2 问的结论，以说明 τ 的存在性，再回到第 1 问，补充证明其唯一性:

σ 在基 $V = (v_1, v_2, \cdots, v_n)$ 下对应于矩阵 A, 取 τ 为在这组基下 $A^\star$ 对应的线性变换,任取 $v = (v_1, \cdots, v_n)\alpha$, $w = (v_1, \cdots, v_n)\beta$, 则 $\sigma v = (v_1, \cdots, v_n) A \alpha$, $\tau w = (v_1, \cdots, v_n) A^\star \beta$.

计算得到 $[\sigma v, w] = \alpha^\star A^\star V^\star V \beta$, $[v, \tau w] = \alpha^\star V^\star V A^\star \beta$.

实际上应该需要增加 $(v_1, \cdots, v_n)$ 是标准正交基的条件，以保证 $V^\star V = E$, 从而有 $[\sigma v, w] = [v, \tau w]$.

下面验证这样的 τ 的唯一性：假设 τ_1, τ_2 都满足 $[\sigma v, w] = [v, \tau_i w], \forall v, w \in V, i = 1, 2$.

构造线性变换 $\tau' : x \to \tau_1 x - \tau_2 x$, 则有 $\forall v, w \in V, [v, \tau' w] = 0$. 取 $v = \tau' w$ 即有对任意的 $w \in V$, $[\tau' w, \tau' w] = 0$, $\tau' w = 0$, 这就验证了 $\tau' = \tau_1 - \tau_2$ 是零线性变换, 故 $\tau_1 = \tau_2$.

若 $\sigma\sigma^\star = \sigma^\star\sigma$ 则称 σ 是正规线性变换，这一定义当然与正规矩阵的概念和谐。因为正规矩阵是指使得 $AA^\star = A^\star A$ 成立的矩阵，而在标准正交基下，线性变换与矩阵是对应的。

3. 63 页习题 3

$\Rightarrow$: 由教材 62 页命题 3.4.3 直接可得;

$\Leftarrow$:

Lemma 0.2. *如果 $V = A \oplus W_A = B \oplus W_B$, 且有 $B \subseteq W_A$, 那么 $W_A = B \oplus (W_A \cap W_B)$.*

证明. 显然有 $W_A \cap W_B \subseteq W_B$ 与 B 的交集为 $\{0\}$, 所以 $B + (W_A \cap W_B) = B \oplus (W_A \cap W_B)$.
考虑到 $B \subseteq W_A, (W_A \cap W_B) \subseteq W_A$, 所以 $B \oplus (W_A \cap W_B) \subseteq W_A$. 下验证 $W_A \subseteq B \oplus (W_A \cap W_B)$:
对于任意 $v \in W_A \subseteq V = B \oplus W_B$, $v = b + w_b$, 其中 $b \in B, w_b \in W_B$. 假设 w_b 在 $V = A \oplus W_A$ 下表示为 $w_b = a + w_a$, $a \in A, w_a \in W_A$, 那么有 $v = b + a + w_a$. 由于 $b \in B \subseteq W_A, w_a \in W_A$, W_A 是线性空间, 所以 $a = v - b - w_a \in W_A$, 又 $a \in A$, 所以 $a = 0$. 从而 $w_b = w_a \in W_A$, 即 $w_b \in (W_A \cap W_B)$.
这就验证了 W_A 中任意的 v 可以表示为 $v = b + w_b$, 其中 $b \in B, w_b \in (W_A \cap W_B)$. □

假设 λ 是 σ 在 V 下的某一个特征值，设 $A = \{v \in V | \sigma v = \lambda v\}$.
如果 $\dim A = \dim V$, 那么 σ 在某组基下对应于对角阵 λE.
如果 $0 < \dim A < \dim V$, 显然 A 是一个 σ-子空间，根据题设，它存在 σ-子空间直和补 W_A, $V = A \oplus W_A$. 由 Lemma 0.2 可知 W_A 也满足"每个 σ-子空间都有一个 σ-子空间直和补", 因为对于 W_A 的 σ-子空间 B, 在 V 上存在 σ-子空间直和补 W_B, 从而可以构造 W_A 上的 σ-子空间直和补 $W_A \cap W_B$.
由于 $\dim W_A < \dim V$, 归纳地进行下去，即可证得 σ 在 W_A 的某个基下对应于对角阵 $\Lambda_{\dim W_A}$, 从而证得 σ 在 V 的某组基下对应于对角阵 $\begin{bmatrix} \lambda E_{\dim A} & \\ & \Lambda_{\dim W_A} \end{bmatrix}$

4. 63 页习题 4

(1) 这里利用 Jordan 标准型的唯一性来说明：假设对于 V 的某个非平凡的 σ-子空间 W，存在 σ-子空间直和补 W'，那么假设 $w_1, \cdots w_r$ 是 W 的一组基，$0 < r < n$, $w_{r+1}, \cdots, w_n$ 是 W' 的一组基，则有 $\sigma(w_1, \cdots, w_r, w_{r+1}, \cdots, w_n) = (w_1, \cdots, w_r, w_{r+1}, \cdots, w_n) \begin{bmatrix} A & \\ & B \end{bmatrix}$. 其中 A 是 r 阶方阵，B 是 $n - r$ 阶方阵。于是 $J = \lambda E_n + E_{12} + \cdots + E_{n-1,n}$ 相似于 $\begin{bmatrix} A & \\ & B \end{bmatrix}$ 以及相似于它的 Jordan 标准型 $\begin{bmatrix} J_A & \\ & J_B \end{bmatrix}$. 这与方阵的 Jordan 标准型唯一相矛盾，所以对于 V 的任意非平凡 σ-子空间不存在 σ-子空间直和补。

(2) 不妨将这组基显式地设出来 $\{\alpha_i\}$, $\sigma(\alpha_1, \cdots, \alpha_n) = (\alpha_1, \cdots, \alpha_n)J$. 下面归纳地证明:
若 W 是 V 的一个维数不少于 k 的 σ-子空间，那么 $\alpha_i \in W, \forall i \leq k$.

证明.

i. 对于 $k=1$, 由于 W 维数至少为 1, 故存在 $W \ni v=\sum_i c_i\alpha_i$, 使得至少有一个 $c_i \neq 0$. 设 j 是使 $c_j \neq 0$ 成立的最大下标。考虑到 W 也是 $\tau=\sigma-\lambda I_V$ 不变的，又有 $\tau\alpha_1=0, \tau\alpha_{i+1}=\alpha_i, i\geq 1$, 故 $\tau^{j-1}v=c_j\alpha_1 \in W$, 从而得到 $\alpha_1 \in W$.

ii. 若 W 是 V 的一个维数不少于 $k+1$ 的 σ-子空间，那么存在 $W \ni v=\sum_i c_i\alpha_i$, 使得至少有 $k+1$ 个 $c_i \neq 0$ (否则 W 的维数不超过 k). 设 j 是使得 $c_j \neq 0$ 成立的最大下标, 显然有 $j \geq k+1$。同样地, 构造 $\tau=\sigma-\lambda I_V$, 显然 W 也是 τ 不变的。由于 $\tau^{j-k-1}v=c_j\alpha_{k+1}+\sum_{i>j-k-1} c_i\alpha_{i-(j-k-1)} = c_j\alpha_{k+1}+v' \in W$, 其中 $v' \in \mathtt{span}\{\alpha_1,\cdots,\alpha_k\}$. 由归纳假设知，$\mathtt{span}\{\alpha_1,\cdots,\alpha_k\} \subseteq W$ (由 W 维数不少于 k 推知), 故而 $\alpha_{k+1} \in W$.

□

由上述命题可知，σ-子空间有且仅有 $\{0\}$ 以及 $\mathtt{span}\{\alpha_1,\cdots,\alpha_i\}$ 其中 $1 \leq i \leq n$.

事实上由这一命题也能推出本题 (1) 问中的结论。

5. 63 页习题 6

由本次作业第 2 题 (60 页习题 8) 可知，σ 是正规变换，当且仅当它在某一组标准正交基 $\alpha_1,\cdots,\alpha_n$ 下对应的矩阵 A 是正规矩阵。而由正规矩阵基本定理可知，正规矩阵 A 可以酉对角化，即存在酉矩阵 U, $UAU^\star=\Lambda$. 所以这等价于 σ 在某组标准正交基, 即 $(\alpha_1,\cdots,\alpha_n)U^\star$ 下, 对应于对角矩阵 $\Lambda=UAU^\star$.

那么由教材 61 页引理 3.4.2 可知，设 $\lambda_1,\cdots,\lambda_r$ 是 σ 的互异的特征值，则 $V=\bigoplus_{i=1}^r V_i$, 其中 $V_i=\{v\in V|\sigma v=\lambda_i v\}$. 再仿照教材 62 页命题 3.4.3 的方法, 可将 W 分解为 $W=\bigoplus_i(W\cap V_i)$. 考虑到对于 $W\cap V_i$, 其在 V_i 中存在正交补 $(W\cap V_i)^\perp$ (对 V_i 中的特征向量做正交化即可构造), 显然 $(W\cap V_i)^\perp$ 也是 σ 不变的。于是可以构造 $W^\perp=\bigoplus_i(W\cap V_i)^\perp$, 且它是 σ-子空间。

1. 习题 1

$A=\begin{bmatrix} 3 & -1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 3 & 0 & 5 & -3 \\ 4 & -1 & 3 & -1 \end{bmatrix}$ 首先求其 Jordan 标准型：容易计算 $\det(xE-A)=(x-2)^4$. 考虑

$B=A-2E=\begin{bmatrix} 1 & -1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 3 & 0 & 3 & -3 \\ 4 & -1 & 3 & -3 \end{bmatrix}$, $B^2=O$. 求得 $Bx=0$ 的解空间上的一组基 $\alpha_1=[-1,-1,0,-1]^\top, \alpha_2=[0,0,3,3]^\top$. 容易构造 $\beta_1=[0,1,0,0]^\top, \beta_2=[0,0,1,0]^\top$ 使 $B\beta_1=\alpha_1, B\beta_2=\alpha_2$. 容易验证 $\beta_1,\beta_2,\alpha_1,\alpha_2$ 线性无关，故构造得到 $S=[\alpha_1,\beta_1,\alpha_2,\beta_2]=\begin{bmatrix} -1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 3 & 1 \\ -1 & 0 & 3 & 0 \end{bmatrix}$. 求得 $S^{-1}=\begin{bmatrix} -1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ -1/3 & 0 & 0 & 1/3 \\ 1 & 0 & 1 & -1 \end{bmatrix}$. 从而 $S^{-1}AS=J=\begin{bmatrix} 2 & 1 & & \\ & 2 & & \\ & & 2 & 1 \\ & & & 2 \end{bmatrix}=2E+B_1$. 其中 B_1 是幂零的 Jordan 阵, $B_1^2=O$.

(1) $\sin(2+x)=\sin(2)+\cos(2)\cdot x-\dfrac{\sin(2)}{2}x^2+\cdots$, $\sin(J)=\sin(2E+B_1)=\sin(2)E+\cos(2)B_1$.

$$\sin(A)=S\cdot\sin(J)\cdot S^{-1}=S\cdot\begin{bmatrix}\sin(2)&\cos(2)&&\\&\sin(2)&&\\&&\sin(2)&\cos(2)\\&&&\sin(2)\end{bmatrix}\cdot S^{-1}$$

由于 $\sin A=(\sin 2)SES^{-1}+(\cos 2)SB_1S^{-1}=(\sin 2)E+(\cos 2)(A-2E)$,

具体计算可得 $\sin(A)=\begin{bmatrix}\sin 2+\cos 2&-\cos 2&0&0\\\cos 2&-\cos 2+\sin 2&0&0\\3\cos 2&0&3\cos 2+\sin 2&-3\cos 2\\4\cos 2&-\cos 2&3\cos 2&\sin 2-3\cos 2\end{bmatrix}$.

(2) $\mathrm{e}^J=\mathrm{e}^2(E+B_1)=\mathrm{e}^2\begin{bmatrix}1&1&&\\&1&&\\&&1&1\\&&&1\end{bmatrix}$, $\mathrm{e}^A=S\cdot\mathrm{e}^J\cdot S^{-1}$.

由于 $\mathrm{e}^A=\mathrm{e}^2\cdot S(E+B_1)S^{-1}=\mathrm{e}^2(E+A-2E)=\mathrm{e}^2(A-E)$,

具体计算可得 $\mathrm{e}^A=\mathrm{e}^2\cdot\begin{bmatrix}2&-1&0&0\\1&0&0&0\\3&0&4&-3\\4&-1&3&-2\end{bmatrix}$.

1. 94 页习题 1

为避免符号混淆，将新定义的范数记为 $||-||_m$, $||\alpha||_m:=||A\alpha||$, A 列满秩.

正定性: 由于 $||-||$ 是向量范数，故而 $||\alpha||_m=||A\alpha||\geq 0$. A 可作分解 $A=QR$, 其中 $Q\in\mathbb{F}^{m\times n}$ 是列酉阵 $Q^{\star}Q=E_n$, R 是满秩上三角矩阵。故而 $||\alpha||_m=0\Rightarrow||A\alpha||=0\Rightarrow A\alpha=QR\alpha=0\Rightarrow\alpha=R^{-1}Q^{\star}QR\alpha=0$.

齐次性: $||k\alpha||_m=||A(k\alpha)||=|k|\cdot||A\alpha||=|k|\cdot||\alpha||_m$.

三角不等式: $||\alpha+\beta||_m=||A(\alpha+\beta)||\leq||A\alpha||+||A\beta||=||\alpha||_m+||\beta||_m$.

2. 94 页习题 2

$||\alpha||_p=(\sum_i^n\ \alpha_i{}^p)^{1/p}$. 设 $m=\max_i|\alpha_i|=||\alpha||_\infty$, 则 $1^{1/p}\leq\dfrac{||\alpha||_p}{||\alpha||_\infty}=\left(\sum_i^n\left(\dfrac{\alpha_i}{m}\right)^p\right)^{1/p}\leq n^{1/p}$.

考虑到 $\lim_{p\to\infty}1^{1/p}=1,\lim_{p\to\infty}n^{1/p}=1$, 所以 $\lim_{p\to\infty}\dfrac{||\alpha||_p}{||\alpha||_\infty}=1$, 所以 $\lim_{p\to\infty}||\alpha||_p=||\alpha||_\infty$.

3. 94 页习题 3

(1) 首先验证 $||\alpha||_A=\sqrt{\alpha^{\star}A\alpha}$ 是一个向量范数，其中 A 是正定 Hermite 矩阵

正定性: 由 A 是 Hermite 矩阵可知 $||\alpha||_A\geq 0$ 且 $||\alpha||_A=0\to\alpha=0$.

齐次性: $||k\alpha||_A=\sqrt{k^2\alpha^{\star}A\alpha}=|k|\cdot\sqrt{\alpha^{\star}A\alpha}=|k|\cdot||\alpha||_A$.

三角不等式: 对正定的 Hermite 矩阵 A 作 Cholesky 分解，$A = LL^\star$. 由 Cauchy-Schwarz 不等式, $\forall x, y$,

$[L^\star x, L^\star y] \le \sqrt{[L^\star x, L^\star x]}\sqrt{[L^\star y, L^\star y]}$, 展开即得 $x^\star Ay = x^\star LL^\star y \le \sqrt{x^\star LL^\star xy^\star LL^\star y} = \sqrt{x^\star Axy^\star Ay}$.

所以 $(||x||_A + ||y||_A)^2 - ||x+y||_A^2 = 2\left(\sqrt{x^\star Ax}\sqrt{y^\star Ay} - x^\star Ay\right) \ge 0$, 即 $||x||_A + ||y||_A \ge ||x+y||_A$.

(2) 为说明"当 A 遍历全部 n 阶 Hermite 正定矩阵时，$||-||_A$ 遍历全部由 V 上内积确定的范数", 只需要说明，对于任意一个 V 上内积 $[-,-]$，存在对应的 Hermite 正定矩阵 A 使得 $\forall \beta \in V, [\beta, \beta] = ||\beta||_A^2$. 考虑 $V = \mathbb{C}^{n\times 1}$ 的一组基 $e_1, e_2, \cdots, e_n$, 其中 e_i 表示第 i 个分量为 1 其余分量全为 0 的向量。对于任意的 $\beta \in V$, 有 $\beta = (e_1, e_2, \cdots, e_n)\beta$. 对于内积 $[-,-]$ 有 $[\beta, \beta] = [\sum_i \beta_i e_i, \sum_i \beta_i e_i] = \sum_i \sum_j \beta_i^\star [e_i, e_j]\beta_j = \beta^\star A\beta$, 其中 $A = \{[e_i, e_j]\}_{ij}$，即 A 是由基 $\{e_i\}$ 之间的内积取值组成的矩阵。由内积的正定性可知 A 是正定矩阵，由内积的对称性可知 A 是 Hermite 矩阵。

4. 101 页习题 1

显然，由于 $||A||_{M_1}$ 相当于矩阵拉平后的向量的 $1-$ 范数，所以 $||-||_{M_1}$ 满足向量范数的要求.

且有 $||AB||_{M_1} = \sum_{i,j} |(AB)_{ij}| = \sum_i \sum_j \sum_k |A_{ij}||B_{jk}| \le \sum_i \sum_j |A_{ij}|||B||_{M_1} = ||A||_{M_1}||B||_{M_1}$.

5. 101 页习题 2

$0 \le ||A^k|| \le ||A||^k$. 又知 $||A|| < 1$, 所以 $\lim_{k\to\infty} ||A||^k = 0$, 所以 $\lim_{k\to\infty} ||A^k|| = 0$, 故 $\lim_{k\to\infty} A^k = 0$.

6. 101 页习题 3

(1) $||UA||_F = \sqrt{\mathtt{tr}\left((UA)^\star(UA)\right)} = \sqrt{\mathtt{tr}(A^\star U^\star UA)} = \sqrt{\mathtt{tr}(A^\star A)} = ||A||_F$;

$||AU||_F = \sqrt{\mathtt{tr}\left((AU)^\star(AU)\right)} = \sqrt{\mathtt{tr}(U^\star A^\star AU)} = \sqrt{\mathtt{tr}(UU^\star A^\star A)} = \sqrt{\mathtt{tr}(A^\star A)} = ||A||_F$.

(2) 正规矩阵 N 可酉相似对角化，$N = U^\star \mathtt{diag}\{\lambda_1, \cdots, \lambda_n\}U$. 从而有 $||N||_F = ||U^\star \mathtt{diag}\{\lambda_1, \cdots, \lambda_n\}U||_F = ||\mathtt{diag}\{\lambda_1, \cdots, \lambda_n\}||_F = \sqrt{\sum_i \bar{\lambda}_i\lambda_i} = \sqrt{\sum_i |\lambda_i|^2}$.

7. 101 页习题 4

(1) 对于任一矩阵 A, $||A||_2 = \sqrt{\rho(A^\star A)}$. 其中 $\rho(-)$ 表示谱半径。所以

$||UA||_2 = \sqrt{\rho(A^\star U^\star UA)} = \sqrt{\rho(A^\star A)} = ||A||_2$

$||AU||_2 = \sqrt{\rho(U^\star A^\star AU)} = \sqrt{\rho(A^\star A)} = ||A||_2$, 因为 $\det(xE - U^\star A^\star AU) = \det(xE - A^\star A)$.

所以 $||-||_2$ 是酉不变的。

(2) $N = U^\star \mathtt{diag}\{\lambda_1, \cdots, \lambda_n\}U$, $||N||_2 = ||\mathtt{diag}\{\lambda_1, \cdots, \lambda_n\}||_2 = \sqrt{\max_i |\lambda_i|^2} = \max_i |\lambda_i|$.

1. 习题 6

(1) $||A||_{M_1} := \sum_{1\le i,j\le n} |a_{ij}|$. 由例 4.2.8(1) 可知 $||A||_1 = \max_{1\le j\le n} \sum_{i=1}^n |a_{ij}|$,

从而有 $||A||_1 \le \sum_{1\le j\le n} \sum_{1\le i\le n} |a_{ij}| = ||A||_{M_1}$.

所以对于任意的 $A \in M_n(\mathbb{F}), v \in \mathbb{F}^n$, $||Av||_1 \le ||A||_1||v||_1 \le ||A||_{M_1}||v||_1$, 即 $||-||_{M_1}$ 与 $||-||_1$ 相容。

(2) 由例 4.2.8(2) 知, $||A||_2 = \sqrt{\rho(A^\star A)} = \sqrt{\lambda_{\mathtt{max}}(A^\star A)}$. 而对 Frobenius 范数 $||A||_F := \sqrt{\sum_{i,j} \bar{a}_{ij}a_{ij}} = \sqrt{\mathtt{tr}(A^\star A)} = \sqrt{\sum_i \lambda_i(A^\star A)} \ge \sqrt{\lambda_{\mathtt{max}}(A^\star A)} = ||A||_2$.

所以 $||Av||_2 \le ||A||_2||v||_2 \le ||A||_F||v||_2$, 即 $||-||_F$ 与 $||-||_2$ 相容。

(3) $||A||_{M_\infty} := n \cdot \max_{i,j} |a_{ij}|$.

$p=1$: 由例 4.2.8(1) 知, $||A||_1 = \max_j \sum_i |a_{ij}| = \sum_i |a_{ij'}| \le n \cdot \max_i |a_{ij'}| \le n \cdot \max_{i,j} |a_{ij}| = ||A||_{M_\infty}$, 同 (1)(2) 的步骤可知，$||-||_{M_\infty}$ 与 $||-||_1$ 相容

$p=2$: 由 (2) 知, $||A||_2 \le ||A||_F = \sqrt{\sum_{i,j} |a_{ij}|^2} \le \sqrt{n^2 \cdot (\max_{i,j} |a_{ij}|)^2} = ||A||_{M_\infty}$. 再同 (1)(2) 可得 $||-||_{M_\infty}$ 与 $||-||_2$ 相容

$p=\infty$: 由例 4.2.8(3) 知, $||A||_\infty = \max_i \sum_j |a_{ij}|$, 完全仿照 $p=1$ 的情况即得 $||-||_{M_\infty}$ 与 $||-||_\infty$ 相容

2. 习题 7

考虑对 $A \in M_n(\mathbb{C})$ 作奇异值分解，即 $A = USV$, 其中 U, V 是酉矩阵，$S = \mathtt{diag}(s_A)$, $s_A(1), \cdots, s_A(n)$ 是 A 的全体奇异值。由于 $||-||$ 是酉不变的，所以 $||A|| = ||USV|| = ||S||$.

按如下方式定义 $\mathbb{R}^n$ 上的范数 N: $N(v) := ||\mathtt{diag}(v)||$. 由于矩阵范数 $||-||$ 满足向量范数的各个要求，所以这样定义出来的 N 也满足向量范数的要求，且显然有 $||A|| = ||S|| = ||\mathtt{diag}(s_A)|| = N(s_A)$.

3. 习题 8

对于任意的 $x, y \in \mathbb{R}^+, \lambda \in [0,1]$, $N(\lambda x + (1-\lambda)y) = ||A + (\lambda x + (1-\lambda)y)B|| = ||\lambda A + \lambda x B + (1-\lambda)A + (1-\lambda)yB|| \le ||\lambda A + \lambda x B|| + ||(1-\lambda)A + (1-\lambda)yB|| = \lambda N(x) + (1-\lambda)N(y)$, 这就验证了 $N(x)$ 是凸函数