2022 秋高等代数习题课

2022-11-18 第六次习题课

第一题 (习题 4.3 第 2 题 (3)) 设 A 是方阵，$A^k=0$ 对某个正整数 k 成立。求证下列方阵可逆，并分别求他们的逆

$$I+A+\frac{1}{2!}A^2+\cdots+\frac{1}{(k-1)!}A^{k-1}.$$

解：我们知道函数 $f(x)=e^x$ 在 $x=0$ 附近有展开式

$$f(x)=1+x+\frac{1}{2!}x^2+\cdots+\frac{1}{(k-1)!}x^{k-1}+\cdots$$

那么我们可以类比考虑 $f(-A)=I+(-A)+\frac{1}{2!}(-A)^2+\cdots+\frac{1}{(k-1)!}(-A)^{k-1}$, 看它与 $I+A+\frac{1}{2!}A^2+\cdots+\frac{1}{(k-1)!}A^{k-1}$ 相乘是否等于单位阵 I. 我们有

$$\begin{aligned}\sum_{i=0}^{k-1}\frac{1}{i!}A^i\sum_{j=0}^{k-1}\frac{1}{j!}(-A)^j=\frac{1}{0!0!}I+\frac{1}{0!1!}A+\frac{1}{0!2!}A^2+\frac{1}{0!3!}A^3+\cdots+\frac{1}{0!(k-1)!}A^{k-1}\\ +\frac{-1}{1!0!}A+\frac{-1}{1!1!}A^2+\frac{-1}{1!2!}A^3+\cdots+\frac{-1}{1!(k-2)!}A^{k-1}\\ +\frac{1}{2!0!}A^2+\frac{1}{2!1!}A^3+\cdots+\frac{1}{2!(k-3)!}A^{k-1}\\ \vdots\\ +\frac{(-1)^{k-1}}{(k-1)!0!}A^{k-1}\end{aligned}$$

可以看到，当 $1\leqslant n\leqslant k-1$ 时，上式右边 A^n 的系数等于

$$\sum_{i=0}^{n}\frac{(-1)^i}{i!(n-i)!}=\frac{1}{n!}\sum_{i=0}^{n}C_n^i1^{n-i}(-1)^i=\frac{1}{n!}(1-1)^n=0.$$

于是有 $\sum\limits_{i=0}^{k-1}\frac{1}{i!}A^i\sum\limits_{j=0}^{k-1}\frac{1}{j!}(-A)^j=\frac{1}{0!0!}I=I$. 所以 $\sum\limits_{i=0}^{k-1}\frac{1}{i!}A^i$ 是可逆的，它的逆就是 $\sum\limits_{j=0}^{k-1}\frac{1}{j!}(-A)^j$.

事实上，我们有所谓的方阵函数 $f:M_n(\mathbb{F})\to M_n(\mathbb{F})$, 可以从多项式扩展定义到一般的（复）函数（由所谓的代表多项式定义），例如 $\sin,\cos,\exp,\log$ 等。这些方阵函数保留了很多原来函数的

性质。这部分内容在学习完方阵的 Jordan 标准形之后比较好理解，所以这里暂时不展开讲了。这题的关键就在于利用实值函数的展开式，猜测一个逆，然后用矩阵的乘法去验证。

第二题 (习题 4.7 第 5 题) 设 $A \in \mathbb{F}^{m\times n}, \operatorname{rank} A = r$.

(1) 从 A 中任意取出 s 行组成 $s\times n$ 矩阵 B, 证明：$\operatorname{rank} B \geqslant r+s-m$;

(2) 从 A 中任意指定 s 个行和 t 个列，这些行和列的交叉位置的元组成的 $s\times t$ 矩阵记为 D, 求证：$\operatorname{rank} D \geqslant r+s+t-m-n$.

证明：(1) 我们记这 s 行的下标为 $i_1,\ldots,i_s$, 余下的行的下标为 $i_{s+1},\ldots,i_m$.

方法一：考虑 $m-s+1$ 个行向量组 $\{v_{i_1},\ldots,v_{i_s}\},\{v_{i_1},\ldots,v_{i_s},v_{i_{s+1}}\},\ldots,\{v_{i_1},\ldots,v_{i_s},\ldots,v_{i_m}\}$, 其中 v 为 A 的行向量。记这些向量组的秩为 $r_0,\ldots,r_{m-s}$, 那么

$$r_0=\operatorname{rank} B, r_{m-s}=\operatorname{rank} A=r, \text{ 并且有} r_{k+1}-r_k\leqslant 1,\ k=0,\ldots,m-s-1,$$

其中等号成立当且仅当 v_{s+k+1} 不落在前一个向量组张成的空间中。所以有

$$r-\operatorname{rank} B=r_{m-s}-r_0=\sum_{k=0}^{m-s-1}(r_{k+1}-r_k)\leqslant\sum_{k=0}^{m-s-1}1=m-s.$$

方法二：令 P 为 m 阶单位阵删去第 $i_{s+1},\ldots,i_m$ 行组成的 $s\times m$ 矩阵，那么有 $PA=B$, 且容易看出 $\operatorname{rank} P=s$. 根据之前证明的矩阵乘积的秩关系

$$\operatorname{rank} PA\geqslant\operatorname{rank} P+\operatorname{rank} A-m$$

即有

$$\operatorname{rank} B\geqslant s+r-m.$$

(2) 令 B 为从 A 中取出这 s 行构成的 $s\times n$ 矩阵，那么根据 (1) 中结论，有

$$\operatorname{rank} B\geqslant r+s-m.$$

从矩阵 B 中取对应题设的 t 列，可以得到矩阵 D. 那么将 (1) 中结论应用到 B^T 与 D^T, 会有

$$\operatorname{rank} D=\operatorname{rank} D^T\geqslant\operatorname{rank} B^T+t-n=\operatorname{rank} B+t-n\geqslant s+r-m+t-n.$$

第三题 (习题 4.7 第 9 题) 设 $A\in\mathbb{F}^{m\times n}$, 求证：$\operatorname{rank}(I_m-AA^T)-\operatorname{rank}(I_n-A^TA)=m-n$.

证明：容易想到 I_m-AA^T 与 I_n-A^TA 都是某一个（同一个）分块矩阵进行了不同的“初等行列变换”得到的某些位置上的矩阵。考虑分块矩阵 $M=\begin{pmatrix}A & I_m\\ I_n & A^T\end{pmatrix}$. 一方面我们以 I_n 为“中心”保持不变做“初等行列变换”有

$$\begin{pmatrix}I_m & -A\\ 0 & I_n\end{pmatrix}\begin{pmatrix}A & I_m\\ I_n & A^T\end{pmatrix}\begin{pmatrix}I_n & -A^T\\ 0 & I_m\end{pmatrix}=\begin{pmatrix}I_m & -A\\ 0 & I_n\end{pmatrix}\begin{pmatrix}A & I_m-AA^T\\ I_n & 0\end{pmatrix}=\begin{pmatrix}0 & I_m-AA^T\\ I_n & 0\end{pmatrix},$$

从而有 $\operatorname{rank} M = \operatorname{rank}\begin{pmatrix} 0 & I_m - AA^T \\ I_n & 0 \end{pmatrix} = n + \operatorname{rank}(I_m - AA^T)$.

另一方面以 I_m 为 “中心” 保持不变做 “初等行列变换” 有

$$\begin{pmatrix} I_m & 0 \\ -A^T & I_n \end{pmatrix}\begin{pmatrix} A & I_m \\ I_n & A^T \end{pmatrix}\begin{pmatrix} I_n & 0 \\ -A & I_m \end{pmatrix} = \begin{pmatrix} I_m & 0 \\ -A^T & I_n \end{pmatrix}\begin{pmatrix} 0 & I_m \\ I_n - A^TA & A^T \end{pmatrix} = \begin{pmatrix} 0 & I_m \\ I_n - A^TA & 0 \end{pmatrix},$$

从而有 $\operatorname{rank} M = \operatorname{rank}\begin{pmatrix} 0 & I_m \\ I_n - A^TA & 0 \end{pmatrix} = m + \operatorname{rank}(I_n - A^TA)$, 进而有

$$\operatorname{rank} M = m + \operatorname{rank}(I_n - A^TA) = n + \operatorname{rank}(I_m - AA^T),$$

即有

$$\operatorname{rank}(I_m - AA^T) - \operatorname{rank}(I_n - A^TA) = m - n.$$

类似的从一个 “中间分块矩阵” 通过不同的 “初等行列变换” 组合得到我们想要的不同形式的矩阵的方法，在之前我们证明 Sherman–Morrison formula，以及 Woodbury matrix identity 的时候已经介绍过了，这里又应用了一次。

第四题 (习题 4.7 第 10 题) 矩阵的广义逆。

(1) 对任意矩阵 $A \in \mathbb{F}^{m\times n}$, 存在矩阵 $A^- \in \mathbb{F}^{n\times m}$ 满足条件 $AA^-A = A$. 什么条件下 A^- 由 A 唯一决定?（A^- 称为 A 的广义逆（generalized inverse matrix)。)

(2) 设 $A \in \mathbb{F}^{m\times n}, \beta \in \mathbb{F}^{m\times 1}, A^- \in \mathbb{F}^{n\times m}$ 满足条件 $AA^-A = A$. 求证：
线性方程组 $AX = \beta$ 有解的充分必要条件是 $AA^-\beta = \beta$;
方程组有解时的通解为 $X = A^-\beta + (I_n - A^-A)Y, \forall Y \in \mathbb{F}^{n\times 1}$.

证明：(1) 我们先证明广义逆 A^- 的存在性。考虑 A 的相抵标准形，即设 P, Q 分别为 m 阶与 n 阶可逆阵，使得 $A = P\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}_{m\times n} Q$, 其中 $r = \operatorname{rank} A$. 假设 A^- 存在，并将其做对应的划分

$$A^- = Q^{-1}\begin{pmatrix} E_0 & E_1 \\ E_2 & E_3 \end{pmatrix}_{n\times m} P^{-1},$$

E_0, E_1, E_2, E_3 分别是大小为 $r\times r, r\times(m-r), (n-r)\times r$, 以及 $(n-r)\times(m-r)$ 的块。那么根据 $A = AA^-A$, 即

$$P\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}_{m\times n} Q = P\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}_{m\times n} Q\cdot Q^{-1}\begin{pmatrix} E_0 & E_1 \\ E_2 & E_3 \end{pmatrix}_{n\times m} P^{-1}\cdot P\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}_{m\times n} Q = P\begin{pmatrix} E_0 & 0 \\ 0 & 0 \end{pmatrix}_{m\times n} Q$$

知，只需要 $E_0 = I_r$ 即可使上式恒成立。故 A 的广义逆 A^- 是一定存在的，其一般形式为

$$A^- = Q^{-1}\begin{pmatrix} I_r & E_1 \\ E_2 & E_3 \end{pmatrix}_{n\times m} P^{-1},$$

E_1, E_2, E_3 是大小为 $r\times(m-r), (n-r)\times r$, 以及 $(n-r)\times(m-r)$ 的块，块中元素可以任取。从以上广义逆 A^- 的形式可以看出，一个矩阵 A 的广义逆并不唯一。要使得 A^- 由 A 唯一决定，只有让 E_1, E_2, E_3 都不存在，即 $m - r = n - r = 0$, 即 A 为可逆方阵。

广义逆这个名字指的是“像是逆”的一个矩阵：

$$(AA^-)A = A = A(A^-A)$$

AA^- 与 A^-A 都“像是”单位阵（左、右）作用于 A 上。他们分别在 A 的列空间与行空间上是恒等变换。这个虽然弱于在全空间上是恒等变换，但已经是比较强的条件了（比不变子空间强很多）。

(2) $\Rightarrow$: 若线性方程组 $AX=\beta$ 有解，任取一个解，记为 X_0, 那么有

$$AA^-\beta = AA^-(AX_0) = (AA^-A)X_0 = AX_0 = \beta.$$

$\Leftarrow$: 若 $AA^-\beta=\beta$, 那么令 $X_0 = A^-\beta$, 我们会有

$$AX_0 = A(A^-\beta) = AA^-\beta = \beta,$$

即知 $X_0 = A^-\beta$ 为线性方程组 $AX=\beta$ 的一个解。

这一问比较好理解，就是我们知道 $AX=\beta$ 有解，当且仅当 β 属于 A 的列空间 $\mathrm{col}(A)$, 而上面已经提到了，AA^- 在 A 的列空间上是恒等变换，自然会把 β 映成 β.

要证明线性方程组 $AX=\beta$ 有解时的通解为 $X = A^-\beta + (I - A^-A)Y, \forall Y \in \mathbb{F}^{m\times 1}$, 只要证明对应的齐次线性方程组 $AX=0$ 的解的全体为 $X = (I_n - A^-A)Y, \forall Y \in \mathbb{F}^{n\times 1}$.

由于广义逆 A^- 满足 $AA^-A = A$, 即 $A(I_n - A^-A) = 0$, 所以

$$\mathrm{col}(I_n - A^-A) \subseteq \mathrm{Null}(A),$$

其中 $\mathrm{col}(I_n - A^-A)$ 为矩阵 $I_n - A^-A$ 的列空间，即由 $I_n - A^-A$ 的列张成的线性空间，$\mathrm{Null}(A)$ 为矩阵 A 的零化空间，即 $AX=0$ 的解空间。要证明以上包含关系实际上是相等的关系，我们需要考察他们的维数。

沿用第 (1) 问的记号，我们有

$$A^-A = Q^{-1}\begin{pmatrix} I_r & E_1 \\ E_2 & E_3 \end{pmatrix}_{n\times m} P^{-1}P\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}_{m\times n} Q = Q^{-1}\begin{pmatrix} I_r & 0 \\ E_2 & 0 \end{pmatrix}_{n\times n} Q.$$

所以

$$\dim\left(\mathrm{col}(I_n - A^-A)\right) = \mathrm{rank}(I_n - A^-A) = \mathrm{rank}\left(Q^{-1}\begin{pmatrix} 0 & 0 \\ -E_2 & I_{n-r} \end{pmatrix}_{n\times n} Q\right) = n - r = \dim\left(\mathrm{Null}(A)\right).$$

也就是说，我们实质上有

$$\{(I_n - A^-A)Y \mid Y \in \mathbb{F}^{n\times 1}\} = \mathrm{col}(I_n - A^-A) = \mathrm{Null}(A).$$

所以齐次线性方程组 $AX=0$ 的解的全体为 $X = (I_n - A^-A)Y, \forall Y \in \mathbb{F}^{n\times 1}$.

以上是当我们取定了某一个广义逆 A^- 时，对线性方程组 $AX=\beta$ 解集的刻画。实际上，当 $\beta \neq 0$ 时，我们还有另一种刻画，即非齐次线性方程组 $AX=\beta$ 解集等于

$$\{A^-\beta \mid A^- \text{ 是 } A \text{ 的广义逆}\}.$$

我们已经知道了形如 $A^-\beta$ 的向量是 $AX=\beta$ 的解，这里的 A^- 是矩阵 A 的任何一个广义逆。我们只要证明对 $AX=\beta$ 的任何一个解 X_0 都存在某个广义逆 A^-, 使得 $X_0 = A^-\beta$.

我们利用 A 的相抵标准形以及对应的 A^- 的一般形式。设 E_1, E_2, E_3 分别为元素未定的大小为 $r\times(m-r), (n-r)\times r$, 以及 $(n-r)\times(m-r)$ 的矩阵。我们想要求解下列的方程

$$Q^{-1}\begin{pmatrix} I_r & E_1 \\ E_2 & E_3 \end{pmatrix} P^{-1}\beta = X_0.$$

我们已知的是 $AX_0=\beta$, 即 $P\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} QX_0=\beta$. 记 $X_0'=QX_0, \beta'=P^{-1}\beta$, 那么我们有

$$\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} X_0'=\beta', \quad \textbf{要求解矩阵方程} \begin{pmatrix} I_r & E_1 \\ E_2 & E_3 \end{pmatrix}\beta'=X_0'.$$

即求解

$$X_0'=\begin{pmatrix} I_r & E_1 \\ E_2 & E_3 \end{pmatrix}\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} X_0'=\begin{pmatrix} I_r & 0 \\ E_2 & 0 \end{pmatrix} X_0'.$$

令 $X_0'=\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$, 其中 $v_1\in\mathbb{F}^{r\times 1}, v_2\in\mathbb{F}^{(n-s)\times 1}$, 有

$$\beta'=\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}=\begin{pmatrix} v_1 \\ 0 \end{pmatrix}, \quad \textbf{相应矩阵方程化为} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}=\begin{pmatrix} I_r & 0 \\ E_2 & 0 \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}=\begin{pmatrix} v_1 \\ E_2v_1 \end{pmatrix}.$$

由于我们假设了 $\beta\neq 0$, 即有 $\beta'\neq 0$, 那么 $v_1\neq 0$. 设 v_1 的第 i 位元素等于 $a\in\mathbb{F}, a\neq 0$, 那么取

$$E_2=(0,\ldots,0,\underbrace{\frac{1}{a}v_2}_{\text{第 } i \text{ 列}},0,\ldots,0),$$

即可以满足 $E_2v_1=v_2$. 其余的 E_1, E_3 任取，即可满足

$$\begin{pmatrix} I_r & E_1 \\ E_2 & E_3 \end{pmatrix}\beta'=\begin{pmatrix} I_r & E_1 \\ E_2 & E_3 \end{pmatrix}\begin{pmatrix} v_1 \\ 0 \end{pmatrix}=\begin{pmatrix} v_1 \\ E_2v_1 \end{pmatrix}=\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}=X_0'.$$

也就是说，对于非齐次线性方程组 $AX=\beta$ 的任取的一个解 X_0, 我们都找到了相应的 A 的广义逆

$$A^-=Q^{-1}\begin{pmatrix} I_r & E_1 \\ E_2 & E_3 \end{pmatrix}P^{-1}$$

使得 $X_0=A^-\beta$.

由于广义逆并不唯一，我们可以在 $AA^-A=A$ 以外多加一些条件（$A^-AA^-=A^-, AA^-$ 与 A^-A 都对称），得到满足唯一性的 Moore–Penrose 广义逆（或伪逆，pseudoinverse）A^+. A^+ 的构造可以由 A 的奇异值分解引出。A^+ 可以用于求解最小二乘问题。

第五题 **(习题 5.1 第 4 题) (综合除法) 设 $f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$ 是数域 $\mathbb{F}$ 上的多项式，$c\in\mathbb{F}$. 求证：$x-c$ 除 $f(x)$ 的商 $q(x)=b_{n-1}x^{n-1}+\cdots+b_1x+b_0$ 和余式 r 可以用如下的算法得出**

$$\begin{array}{c|ccccccc} c & & a_n & a_{n-1} & \cdots & a_i & \cdots & a_1 & a_0 \\ & +) & & cb_{n-1} & \cdots & cb_i & \cdots & cb_1 & cb_0 \\ \hline & & b_{n-1} & b_{n-1} & \cdots & b_{i-1} & \cdots & b_0 & r \end{array}$$

其中 $b_{n-1} = a_n, b_{i-1} = a_i + cb_i \ (\forall 1 \leqslant i \leqslant n), r = a_0 + cb_0.$

证明：
考虑 $f(x) = (x-c)g(x) + r$, **即**

$$\begin{aligned} a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 &= (x-c)(b_{n-1}x^{n-1} + \cdots + b_1 x + b_0) + r \\ &= b_{n-1}x^n + \cdots + b_1 x^2 + b_0 x + r \\ &\quad - cb_{n-1}x^{n-1} - \cdots - cb_1 x - cb_0 \end{aligned}$$

移项之后即有

$$a_n x^n + (a_{n-1} + cb_{n-1})x^{n-1} + \cdots + (a_1 + cb_1)x + (a_0 + cb_0) = b_{n-1}x^n + \cdots + b_1 x^2 + b_0 x + r,$$

对应次项的系数相等，从而有 $b_{n-1} = a_n, b_{n-2} = a_{n-1} + cb_{n-1}, \ldots, b_0 = a_1 + cb_1, r = a_0 + cb_0.$