Linear Algebra 06

时间：2022年5月12日

注意纠正Linear Algebra 05 中第三题的误区

(方阵可逆，非方阵的逆不存在，所以第二题求解坐标变换矩阵只能用方法三，不可以使用方法一和方法二)

矩阵的秩的相关定义的证明：

[常见的矩阵秩(不)等式及其各种证明 - 知乎 (zhihu.com)]

去年附加题第二题的证明：

[康托洛维奇(Kantorovich)不等式的一种初等证明 - 知乎 (zhihu.com)]

1. $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 分别是 $m \times n$ 和 $n \times p$ 矩阵, 证明: $r(\boldsymbol{AB}) + n \geq r(\boldsymbol{A}) + r(\boldsymbol{B})$ 。

   提示：(1)设A是$m{\times}n$矩阵，B是$n{\times}s$矩阵，若$AB=O$，证明$rank(A)+rank(B){\le}n.$

   (2)设A,B是行数相同的矩阵，(A,B)是由A,B并排组成的矩阵，证明$rank(A,B){\le}rank(A)+rank(B).$

   引理1

   设A是$m{\times}n$矩阵，B是$n{\times}s$矩阵，若$AB=O$，证明$rank(A)+rank(B){\le}n.$

   第一题的证明方法有多种，我们采用方程组的形式给大家讲解

   1、 引理1. 作业 Linear Algebra 01 的证明题.

   证明：思路，对矩阵B分块考虑

   于是 $\vec{B}=\{\vec{\beta_1},\vec{\beta_2},\vec{\beta_3},\cdots,\vec{\beta_s}\}$.

   故 $AB=O \Leftrightarrow [A\vec{\beta_1},A\vec{\beta_2},A\vec{\beta_3},\cdots,A\vec{\beta_s}]=0$

   对于 $\forall i\in[1,s]$， $\vec{\beta_i}$ 为 $A\vec{\beta_i}=\vec{0}$ 的零空间（这个是书上的秩定理）参考23页定理14.

   所以 $rank(\vec{\beta_i})=n-rank(A)$.

   提 对于矩阵B而言 $rank(\vec{\beta_1},\vec{\beta_2},\vec{\beta_3},\cdots\vec{\beta_s})=rank(B)$.

   $\cancel{rank(\vec{\beta_1},\vec{\beta_2},\vec{\beta_3},\cdots\vec{\beta_s})\le rank(\vec{\beta_1})+rank(\vec{\beta_2}}$

   而B的各列$\vec{\beta_i}$都含于齐次线性方程组的$AX=0$的零空间V_A.

   由于 $\dim V_A=n-rank(A)$，而由于V_A的子集$\{\vec{\beta_1},\vec{\beta_2},\cdots,\vec{\beta_s}\}$最多有$n-rank(A)$个线性无关的向量，因此 $rank(B)\le n-rank(A)$ □.

   [证明]

   分析：总体思路：采用矩阵分块的方式求解.题目具有一定难度.

   证明：
   对矩阵B按照列分块，我们记$B=\begin{bmatrix}\beta_1 & \beta_2 & \beta_3 & \cdots & \beta_s\end{bmatrix}$
   那么有

$$
\begin{aligned}
AB &= A\begin{bmatrix}\beta_1 & \beta_2 & \beta_3 & \cdots & \beta_s\end{bmatrix} \\
&= \begin{bmatrix}A\beta_1 & A\beta_2 & A\beta_3 & \cdots & A\beta_s\end{bmatrix} \\
&= \begin{bmatrix}\vec{0} & \vec{0} & \vec{0} & \cdots & \vec{0}\end{bmatrix}
\end{aligned}
$$

   于是我们有$A\beta_j=\vec{0}, j=1,2,3,\cdots,s,$
   所以B的列向量都是齐次方程组$A\vec{x}=\vec{0}$的解，由于方程组$A\vec{x}=\vec{0}$的解向量的$rank(\vec{x})=n-rank(A)$,这里的$\vec{x}$与$\vec{\beta}$含义等价，所以

$$rank(\beta_1,\beta_2,\beta_3,\cdots,\beta_s){\le}n-rank(A)$$

   我们又知道

$$rank(\beta_1, \beta_2, \beta_3, \cdots, \beta_s) = rank(B)$$

所以：$rank(A) + rank(B) \leq n$.

引理2

设A,B是行数相同的矩阵，(A, B)是由A, B并排组成的矩阵，证明 $rank(A, B) \leq rank(A) + rank(B)$.

即是：分块矩阵的秩大于等于各分块的秩中的最大值，小于等于各分块的秩之和。

引理2. 设A.B为行数相等，并排组成的矩阵. 证明 $rank(A,B) \leq rank(A) + rank(B)$.

证明：对矩阵(A,B)分块有 $\{\vec{\alpha}_1, \vec{\alpha}_2, \cdots, \vec{\alpha}_m, \vec{\beta}_1, \vec{\beta}_2, \cdots, \vec{\beta}_n\}$.

$A = [\vec{\alpha}_1, \vec{\alpha}_2, \cdots, \vec{\alpha}_m]$

$B = [\vec{\beta}_1, \vec{\beta}_2, \cdots, \vec{\beta}_n]$

由于A的一个极大线性无关组A'和B的一个极大线性无关组B'可以表示(A,B).

我们知道(A,B)的一个极大线性无关组的向量个数小于

A', B'的向量个数之和.

故有 $rank(A,B) \leq rank(A') + rank(B') = rank(A) + rank(B)$ □

[证明]

设 A 为 $s \times m$ 矩阵, B 为 $s \times n$ 矩阵, 则 $\max\{\mathrm{r}(A), \mathrm{r}(B)\} \leq \mathrm{r}(A, B) \leq \mathrm{r}(A) + \mathrm{r}(B)$.
(1) 因为 A 和 B 的子式也是分块矩阵 (A, B) 的子式, 所以 $\mathrm{r}(A) \leq \mathrm{r}(A, B), \mathrm{r}(B) \leq \mathrm{r}(A, B)$.
由此可见 $\max\{\mathrm{r}(A), \mathrm{r}(B)\} \leq \mathrm{r}(A, B)$ 成立.
(2) 设 $PA^{\mathrm{T}} = U, QB^{\mathrm{T}} = V$, 其中 P, Q 可逆,
U, V 为行阶梯形矩阵, $\mathrm{r}(U) = \mathrm{r}\left(A^{\mathrm{T}}\right), \mathrm{r}(V) = \mathrm{r}\left(B^{\mathrm{T}}\right)$.
于是

$$\mathrm{r}(\boldsymbol{A}, \boldsymbol{B}) = \mathrm{r}(\boldsymbol{A}, \boldsymbol{B})^{\mathrm{T}} = \mathrm{r}\begin{bmatrix} A^{\mathrm{T}} \\ B^{\mathrm{T}} \end{bmatrix} = \mathrm{r}\left[\begin{pmatrix} \boldsymbol{P} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{Q} \end{pmatrix}\begin{pmatrix} A^{\mathrm{T}} \\ B^{\mathrm{T}} \end{pmatrix}\right] = \mathbf{r}\begin{bmatrix} U \\ V \end{bmatrix} \leq \mathbf{r}\left(A^{\mathrm{T}}\right) + \mathbf{r}\left(B^{\mathrm{T}}\right) = \mathbf{r}(A) + \mathbf{r}(B)$$

本题证明

本题证明：分解 $A_{m\times n}$，$B_{n\times p}$ 矩阵.

$B = [\vec{\beta}_1, \vec{\beta}_2, \cdots, \vec{\beta}_p]$

于是 $AB = [A\vec{\beta}_1, A\vec{\beta}_2, \cdots, A\vec{\beta}_p]$ ---①

设 $A\vec{\beta}_{r_1}, A\vec{\beta}_{r_2}, \cdots, A\vec{\beta}_{r_s}$ 是 ---①的一个极大线性无关组,

那对于余下的 p-s 向量中的任意一个向量 $\vec{\beta}_j$ 而言，都有

$A\vec{\beta}_j = k_{j1}A\vec{\beta}_{r_1} + k_{j2}A\vec{\beta}_{r_2} + \cdots + k_{js}A\vec{\beta}_{r_s}$，即 $\vec{\beta}_j$ 可以被线性表示.(极大线性无关组)

所以对于方程$AX=0$，有 $A\hat{\beta}_j = A\vec{\beta}_j - \sum_{i=1}^{s} k_{ji}A\vec{\beta}_{r_i} = 0$，这样我们可以构

造 $\hat{\beta}_j = \vec{\beta}_j - \sum_{i=1}^{s} k_{ji}\vec{\beta}_{j_i}$ 为$AX=0$的解.

我们有这 p-s 个向量 $\vec{\beta}_1', \vec{\beta}_2', \cdots, \vec{\beta}_{p-s}'$ 构成的矩阵构成$AX=0$的零空间,

我们由引理1知 $rank(\vec{\beta}_1', \vec{\beta}_2', \cdots, \vec{\beta}_{p-s}') \leq n - rank(A)$

由于 $rank(B) = rank(\vec{\beta}_1, \cdots, \vec{\beta}_p) = rank(\vec{\beta}_{r_1}, \vec{\beta}_{r_2}, \cdots, \vec{\beta}_{r_s}, \vec{\beta}_1', \vec{\beta}_2', \cdots, \vec{\beta}_{p-s}')$ ---②

由引理2. 我们知道 $rank(\vec{\beta}_{r_1}, \vec{\beta}_{r_2}, \cdots, \vec{\beta}_{r_s}) = rank(AB)$. $rank(\vec{\beta}_1', \vec{\beta}_2', \cdots, \vec{\beta}_{p-s}') \leq n - rank(A)$

② $\leq rank(\vec{\beta}_{r_1}, \vec{\beta}_{r_2}, \cdots, \vec{\beta}_{r_s}) + rank(\vec{\beta}_1', \cdots, \vec{\beta}_{p-s}') \leq rank(AB) + n - rank(A)$.

即 $rank(B) \leq rank(AB) + n - rank(A)$ □.

记 $B = (\beta_1, \beta_2, \ldots, \beta_p)$，其中 $\beta_i (i = 1, 2, \ldots, p)$ 是 B 的列向量。那么

$$AB = (A\beta_1, A\beta_2, \ldots, A\beta_p)$$

设 $A\beta_{r_1}, A\beta_{r_2}, \ldots, A\beta_{r_s}$ 是 AB 的一个极大线性无关组，在剩下的 $p-s$ 个向量中的任 意一个向量 β_j，我们都有 $A\beta_j = k_{j_1}A\beta_{r_1} + \ldots + k_{j_s}A\beta_{r_s}$，因此

$$\hat{\beta}_j=\beta_j-\sum_{i=1}^{s} k_{j_i}\beta_{r_i}$$

是方程 $AX=0$ 的解，这些 $p-s$ 个解向量 $\hat{\beta}_{j_1},\hat{\beta}_{j_2},\ldots,\hat{\beta}_{j_{p-s}}$ 构成的矩阵，我们有

$$\operatorname{rank}\left(\hat{\beta}_{j_1},\hat{\beta}_{j_2},\ldots,\hat{\beta}_{j_{p-s}}\right)\leq n-\operatorname{rank}(A)$$

而

$$\operatorname{rank}(B)=\operatorname{rank}\left(\beta_1,\beta_2,\ldots,\beta_p\right)=\operatorname{rank}\left(\beta_{r_1},\ldots,\beta_{r_s},\hat{\beta}_{j_1},\hat{\beta}_{j_2},\ldots,\hat{\beta}_{j_{p-s}}\right)$$

因此

$$\operatorname{rank}(B)\leq\operatorname{rank}\left(\beta_{r_1},\ldots,\beta_{r_s}\right)+\operatorname{rank}\left(\hat{\beta}_{j_1},\hat{\beta}_{j_2},\ldots,\hat{\beta}_{j_{p-s}}\right)\leq\operatorname{rank}(AB)+n-\operatorname{rank}(A)$$

故命题成立。

2. 已知二次型 $f\left(x_1,x_2,x_3\right)=x_1^2+5x_2^2+5x_3^2+2x_1x_2-4x_1x_3$
(1) 写出二次型 f 的矩阵表达式.
(2) 用正交变换把二次型 f 化成标准形, 并写出相应的正交矩阵.
(3) 当 $\boldsymbol{x}^T\boldsymbol{x}=2$ 时, $f\left(x_1,x_2,x_3\right)$ 的极大值.

[解答]

(1) $f(x_1,x_2,x_3)=x^TAx=[x_1,x_2,x_3]\begin{bmatrix}1&1&-2\\1&5&0\\-2&0&5\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}$

根据定义求解

(2) 我们认为标准型是不含交叉项的.

于是在 A 已经是对称矩阵的前提下.

利用教材 P396 页定理 2.

可以发现: A 可以正交对角化.

步骤①. 求解特征根 (步骤可参考教材 P396 页例 3).

$$|\lambda E-A|=\begin{vmatrix}1-\lambda&1&-2\\1&5-\lambda&0\\-2&0&5-\lambda\end{vmatrix}=\begin{vmatrix}1-\lambda&1&-2\\1&5-\lambda&0\\0&2(5-\lambda)&5-\lambda\end{vmatrix}=(5-\lambda)\begin{vmatrix}1-\lambda&1&-2\\1&5-\lambda&0\\0&2&1\end{vmatrix}$$

$$=(5-\lambda)\begin{vmatrix}1-\lambda&1+2\times2&1\times2-2\\1&5-\lambda&0\\0&2&1\end{vmatrix}=(5-\lambda)\begin{vmatrix}1-\lambda&5&0\\1&5-\lambda&0\\0&2&1\end{vmatrix}$$

$$\Rightarrow|\lambda E-A|=(5-\lambda)\cdot(-1)^{3+3}\begin{vmatrix}1-\lambda&5\\1&5-\lambda\end{vmatrix}=(5-\lambda)(\lambda^2-6\lambda)=\lambda(5-\lambda)(\lambda-6)$$

于是特征值 $\lambda_1=0$, $\lambda_2=5$, $\lambda_3=6$.

○ 当 $\lambda_1=0$ 时，$\begin{bmatrix}1&1&-2\\1&5&0\\-2&0&5\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix} \Rightarrow \begin{bmatrix}1&1&-2&0\\1&5&0&0\\-2&0&5&0\end{bmatrix} \sim \begin{bmatrix}1&1&-2&0\\0&4&2&0\\0&2&1&0\end{bmatrix} \sim \begin{bmatrix}1&1&-2&0\\0&2&1&0\\0&0&0&0\end{bmatrix}$

$\sim \begin{bmatrix}1&5&0&0\\0&2&1&0\\0&0&0&0\end{bmatrix} \Rightarrow \begin{matrix}x_1+5x_2=0\\2x_2+x_3=0\end{matrix} \Rightarrow \begin{bmatrix}5\\-1\\2\end{bmatrix}=\vec{x}$

$\lambda_1=0 \Rightarrow \vec{x}=\begin{bmatrix}5\\-1\\2\end{bmatrix}$.

○ 当 $\lambda_2=5$ 时，$\begin{bmatrix}-4&1&-2&0\\1&0&0&0\\-2&0&0&0\end{bmatrix} \sim \begin{bmatrix}1&0&0&0\\2&0&0&0\\4&-1&2&0\end{bmatrix} \sim \begin{bmatrix}1&0&0&0\\0&1&-2&0\\0&0&0&0\end{bmatrix} \Rightarrow \begin{matrix}x_1=0\\x_2-2x_3=0\end{matrix} \Rightarrow \vec{x}=\begin{bmatrix}0\\2\\1\end{bmatrix}$

$\lambda_1=5$，$\vec{x}=\begin{bmatrix}0\\2\\1\end{bmatrix}$

○ 当 $\lambda_3=6$ 时，$\begin{bmatrix}-5&1&-2&0\\1&-1&0&0\\-2&0&-1&0\end{bmatrix} \sim \begin{bmatrix}1&-1&0&0\\2&0&1&0\\5&-1&2&0\end{bmatrix} \sim \begin{bmatrix}1&-1&0&0\\0&2&1&0\\0&4&2&0\end{bmatrix} \sim \begin{bmatrix}1&-1&0&0\\0&2&1&0\\0&0&0&0\end{bmatrix} \Rightarrow \vec{x}=\begin{bmatrix}-2\\-2\\4\end{bmatrix}$

取 $\vec{x}=\begin{bmatrix}1\\1\\-2\end{bmatrix}$

经检查，上述特征向量均正交。

现在需单位化即可. $\vec{r}_1=\frac{1}{\sqrt{30}}\begin{bmatrix}5\\-1\\2\end{bmatrix}$ $\vec{r}_2=\frac{1}{\sqrt{5}}\begin{bmatrix}0\\2\\1\end{bmatrix}$ $\vec{r}_3=\frac{1}{\sqrt{5}}\begin{bmatrix}1\\1\\-2\end{bmatrix}$

于是 $P=[\vec{r}_1,\vec{r}_2,\vec{r}_3]=\begin{bmatrix}\frac{5}{\sqrt{30}}&0&\frac{1}{\sqrt{5}}\\-\frac{1}{\sqrt{30}}&\frac{2}{\sqrt{5}}&\frac{1}{\sqrt{5}}\\\frac{2}{\sqrt{30}}&\frac{1}{\sqrt{5}}&-\frac{2}{\sqrt{5}}\end{bmatrix}$

经过正交变换 $\vec{x}=P\vec{y}$，二次型可化为标准型（可参考教材 P_{402} 例4）.

$f(x_1,x_2,x_3)=\vec{x}^TA\vec{x}=(P\vec{y})^TAP\vec{y}=\vec{y}^TP^TAP\vec{y}=\vec{y}^T(P^TAP)\vec{y}=\cancel{\vec{y}^TP\vec{y}}$

$\vec{y}^TD\vec{y}=\cancel{5x_2^2+6x_3^2}\ 5y_2^2+6y_3^2$

(1) f 的矩阵表示为

$$f(x_1,x_2,x_3)=\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}=[x_1,x_2,x_3]\begin{bmatrix}1&1&-2\\1&5&0\\-2&0&5\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}$$

(2) 由矩阵 $\boldsymbol{A}$ 的特征多项式

$$\begin{aligned}|\lambda\boldsymbol{E}-\boldsymbol{A}|&\begin{vmatrix}\lambda-1&-1&2\\-1&\lambda-5&0\\2&0&\lambda-5\end{vmatrix}=\begin{vmatrix}\lambda-1&-1&2\\-1&\lambda-5&0\\0&2(\lambda-5)&\lambda-5\end{vmatrix}\\&=\begin{vmatrix}\lambda-1&-5&2\\-1&\lambda-5&0\\0&0&\lambda-5\end{vmatrix}=(\lambda-5)\left(\lambda^2-6\lambda\right)\end{aligned}$$

得到 $\boldsymbol{A}$ 的特征值是 $0,5,6$.

(1) 当 $\lambda=0$ 时,由 $(0\boldsymbol{E}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{0}$, 即 $\begin{bmatrix}-1&-1&2\\-1&-5&0\\0&2&1\end{bmatrix}\rightarrow\begin{bmatrix}1&5&0\\0&2&1\\0&0&0\end{bmatrix}$

得基础解系 $\boldsymbol{\alpha}_1=[5,-1,2]^T$, 即 $\lambda=0$ 的特征向量.

(2) 当 $\lambda=5$ 时, 由 $(5E-A)x=0$, 即

$$\begin{bmatrix}4&-1&2\\-1&0&0\\2&0&0\end{bmatrix}\rightarrow\begin{bmatrix}1&0&0\\0&-1&2\\0&0&0\end{bmatrix}$$

得基础解系, $\boldsymbol{\alpha}_2=[0,2,1]^T$, 即 $\lambda=5$ 的特征向量.

(3) 当 $\lambda=6$ 时, 由 $(6\boldsymbol{E}-\boldsymbol{A})\boldsymbol{x}=0$, 即

$$\begin{bmatrix}5&-1&2\\-1&1&0\\2&0&1\end{bmatrix}\rightarrow\begin{bmatrix}1&-1&0\\0&2&1\\0&0&0\end{bmatrix}$$

得基础解系 $\boldsymbol{\alpha}_3=[1,1,-2]^T$, 即 $\lambda=6$ 的特征向量.

对于实对称矩阵, 特征值不同, 特征向量已正交, 故只需单位化, 有

$$
\gamma_1=\frac{1}{\sqrt{30}}\begin{bmatrix}5\\-1\\2\end{bmatrix},\gamma_2=\frac{1}{\sqrt{5}}\begin{bmatrix}0\\2\\1\end{bmatrix},\gamma_3=\frac{1}{\sqrt{6}}\begin{bmatrix}1\\1\\-2\end{bmatrix}
$$

那么, 令

$$
\boldsymbol{P}=(\gamma_1,\gamma_2,\gamma_3)=\begin{bmatrix}\frac{5}{\sqrt{30}} & 0 & \frac{1}{\sqrt{6}}\\ -\frac{1}{\sqrt{30}} & \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{6}}\\ \frac{2}{\sqrt{30}} & \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{6}}\end{bmatrix}
$$

经正交变换 $\boldsymbol{x}=\boldsymbol{P}\boldsymbol{y}$, 二次型化为标准形

$$
f\left(x_1,x_2,x_3\right)=\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}=\boldsymbol{y}^T\boldsymbol{A}\boldsymbol{y}=5y_2^2+6y_3^2
$$

(3). 当 $\vec{x}^T\vec{x}=2$ 时, $\vec{x}^T\vec{x}=(P\vec{y}^T)P\vec{y}=\vec{y}^TP^TP\vec{y}=\vec{y}^T\vec{y}=y_1^2+y_2^2+y_3^2=2$.

于是 $\vec{x}^TA\vec{x}=5y_2^2+6y_3^2$

等价于 $\Leftrightarrow x_1+x_2+x_3=2$

$\max(5x_2+6x_3)=\max(5(2-x_1-x_3)+6x_3)$

$=\max(10-5x_1+x_3)$

此时 x_1 min, x_3 max.

$x_1,x_2,x_3\geqslant 0$. 故 $x_3=2$, $x_2=0$, $x_1=0$.

有 max $=12$.

(3) $\boldsymbol{x}^T\boldsymbol{x}=(\boldsymbol{P}\boldsymbol{y})^T(\boldsymbol{P}\boldsymbol{y})=\boldsymbol{y}^T\boldsymbol{P}^T\boldsymbol{P}\boldsymbol{y}=\boldsymbol{y}^T\boldsymbol{y}=y_1^2+y_2^2+y_3^2=2$

$$
\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}=5y_2^2+6y_3^2\leqslant 6\left(y_1^2+y_2^2y_3^2\right)
$$

所以, $f_{\max}=12$.