线性代数 (同济 5 版) 复习提要

黄正华*

2015 年 7 月 27 日

目录

0 全书总结

如果非要给这本书加一个副标题, 我希望是 ——《一个方程组引发的故事》.

0.1 全书概览

我们现在使用的教材是工程数学《线性代数》, 是线性代数学科比较基础的部分. 这一部分的中心是围绕"用高斯消元法求解线性方程组"的问题展开的. 全书中心的例子其实是第三章 P.57 的引例.

由此衍生了三个工具 —— 行列式、矩阵、向量组, 从不同的角度来阐述方程组的求解问题.

要特别注意, 矩阵、向量组、方程组这三个问题的相互转化. 比如, 向量组的问题, 可能需要转化成矩阵的问题来解决; 也可能要借助方程组解的相关理论来解决.

第一章的中心可以认为是克拉默法则, 前面的行列式讨论是为克拉默法则作铺垫的.

高斯消元法的过程, 可以简单地表示为方程组的增广矩阵的初等行变换, 这自然引出了对矩阵的专门讨论. 方程组经过高斯消元法总是会稳定地保留一定数量的方程, 这就对应着秩的问题. 矩阵的细部是向量, 更进一步讨论向量的线性表示、线性相关性才说明了, 为什么矩阵的初等行变换中有一些行会被变为

*Email: huangzh@whu.edu.cn

零, 为什么消元法解方程时有的方程会被消掉. 最大无关组的概念才真正解释了, 为什么消元法解方程组时保留下来的方程个数是稳定不变的.

既然中心的议题是解方程组, 那么关于线性方程组解的理论要非常清楚, 比如"$n-r$"的含义, 有解、无解的充要条件.

0.2 要点 TOP 10

下面的要点列为 TOP 10 是因为其理论重要性、易错等原因.

(I) $|\lambda\boldsymbol{A}|=\lambda^n|\boldsymbol{A}|$. ($\boldsymbol{A}$ 为 n 阶方阵)

(II) 矩阵乘法不满足交换律、消去律.

(III) 矩阵秩的性质 $5\sim 8$.

(IV) 特征值性质: $\lambda_1+\lambda_2+\cdots+\lambda_n=a_{11}+a_{22}+\cdots+a_{nn}$; $\lambda_1\lambda_2\cdots\lambda_n=\big|\boldsymbol{A}\big|$.

(V) 若 λ 是 $\boldsymbol{A}$ 的一个特征值, 则 $\varphi(\lambda)$ 是矩阵多项式 $\varphi(\boldsymbol{A})$ 的特征值.

(VI) 齐次线性方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$ 的基础解系由 $n-r$ 个线性无关的解构成.

(VII) 线性方程组有解、无解的充要条件.

(VIII) 矩阵对角化的充要条件.

(IX) 伴随矩阵的定义, $\boldsymbol{A}\boldsymbol{A}^*=\boldsymbol{A}^*\boldsymbol{A}=|\boldsymbol{A}|\boldsymbol{E}$.

(X) 初等变换不改变矩阵的秩.

0.3 例题 TOP 5

教材上的例子, 如下几个在方法上特别重要:

(I) P.75 第三章例题 13, 带参量的线性方程组解的讨论. 特别重视解法二.

(II) P.123 第四章例题 11, 矩阵对角化. 综合了矩阵对角化的充要条件、"$n-r$"结论.

(III) P.65 第三章例题 3, 用初等变换法解矩阵方程. 题目没有难度, 但是很多人没有接受这个简便的解法.

(IV) P.100 第四章例题 13, 证明矩阵秩的性质 8. $\boldsymbol{A}\boldsymbol{B}=\boldsymbol{O}$, 则视 $\boldsymbol{B}$ 的列为方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$ 的解, 这个观念很重要.

(V) P.120 第五章例题 9, 特征值性质的应用. 或见 P.135 习题 12, 13. 这也是近年考研常见的题型.

1 行列式

在线性代数中最重要的内容就是行列式和矩阵. 虽然表面上看, 行列式和矩阵不过是一种符号或速记, 但从数学史上来看, 优良的数学符号和生动的概念是数学思想产生的动力和钥匙.

1.1 基本要求

(一) 熟悉行列式的计算. 一种是数字型的行列式 (一般是一个 4 阶、5 阶的数字行列式), 另一种是 n 阶行列式. 建议熟悉典型例题: 计算 n 阶行列式

$$D_n = \begin{vmatrix} x & a & \cdots & a \\ a & x & \cdots & a \\ \vdots & \vdots & & \vdots \\ a & a & \cdots & x \end{vmatrix}.$$

及其变形, 比如:

$$D_n = \begin{vmatrix} x_1 & a & \cdots & a \\ a & x_2 & \cdots & a \\ \vdots & \vdots & & \vdots \\ a & a & \cdots & x_n \end{vmatrix}. \qquad D_n = \begin{vmatrix} x_1 & a_2 & \cdots & a_n \\ a_1 & x_2 & \cdots & a_n \\ \vdots & \vdots & & \vdots \\ a_1 & a_2 & \cdots & x_n \end{vmatrix}.$$

计算 n 阶行列式, 有以下常用手法:

- 行累加, 即把行列式的某 $n-1$ 个行, 加到余下的一行. 当行列式的各行的和相同时常使用此技巧.
- 主行消法, 即某行的适当倍数, 加到其余的各行.
- 逐行消法, 即第 i 行乘以 k 加到第 $i+1$ 行, $i = n-1, n-2, \cdots, 1$; 或第 $i+1$ 行乘以 k 加到第 i 行, $i = 1, 2, \cdots, n-1$.
- 逐行相邻互换.

要特别重视**逐行消法**. 当然, 这些方法都是行列式三种基本变换的"高级形式".

(二) 牢记克拉默法则. 克拉默法则告诉我们, 对含有 n 个方程的 n 元线性方程组, 系数行列式 $|A| \neq 0$ 时, 方程组有唯一解. 并且给出了相应的公式.

我们要记住的不是这个公式本身, 因为它并没有实用性. 但是从这里, 我们要建立一个牢固的印象: $|\boldsymbol{A}| = 0$ 时, 齐次线性方程组 $\boldsymbol{Ax} = \boldsymbol{0}$ 有非零解. 因为方程组 $\boldsymbol{Ax} = \boldsymbol{0}$ 是天然有解的: 它至少有一个解, 就是零解, 即 $x_1 = x_2 = \cdots = x_n = 0$. 对齐次线性方程组 $\boldsymbol{Ax} = \boldsymbol{0}$, 我们关心的不是它有没有解, 而是关心它什么时候有非零解. $|\boldsymbol{A}| \neq 0$ 时, 方程组有唯一解, 即零解. 反过来, 我们容易记住: $|\boldsymbol{A}| = 0$ 时, 齐次线性方程组 $\boldsymbol{Ax} = \boldsymbol{0}$ 有非零解.

1.2 常用计算公式

请自己总结.

2 矩阵

矩阵是一个数表, 而行列式本质上是一个数. 这是两者的巨大区别.

2.1 本章重点

(一) 矩阵的运算. 特别是矩阵乘法不满足的几个运算律.

(I) 不满足交换律, 即 $\boldsymbol{AB}=\boldsymbol{BA}$ 一般不成立. (特别地, 若 $\boldsymbol{AB}=\boldsymbol{BA}$, 则称矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ **是可交换的**.) 相应地要注意以下几点:

(1) 矩阵乘法特别地有"左乘"和"右乘"的说法, 比如 $\boldsymbol{B}$ 左乘 $\boldsymbol{A}$ (即 $\boldsymbol{BA}$) 和 $\boldsymbol{B}$ 右乘 $\boldsymbol{A}$ (即 $\boldsymbol{AB}$) 是完全不同的.

(2) 在提取公因子的时候, 要分清楚是从左侧还是右侧提出. 比如 $\boldsymbol{AB}-\boldsymbol{B}$ 可以写为 $(\boldsymbol{A}-\boldsymbol{E})\boldsymbol{B}$, 即 $\boldsymbol{B}$ 只能从右侧提出, 不能写成 $\boldsymbol{B}(\boldsymbol{A}-\boldsymbol{E})$. 而 $\boldsymbol{AB}-\boldsymbol{BC}$, 写成 $(\boldsymbol{A}-\boldsymbol{C})\boldsymbol{B}$ 或 $\boldsymbol{B}(\boldsymbol{A}-\boldsymbol{C})$, 都是错误的.

(3) 由于不满足交换律, 下列公式一般不成立:

- $(\boldsymbol{AB})^k=\boldsymbol{A}^k\boldsymbol{B}^k$;
- $(\boldsymbol{A}+\boldsymbol{B})(\boldsymbol{A}-\boldsymbol{B})=\boldsymbol{A}^2-\boldsymbol{B}^2$;
- 牛顿二项式展开式一般不成立, 比如最简单的公式 $(\boldsymbol{A}+\boldsymbol{B})^2=\boldsymbol{A}^2+2\boldsymbol{AB}+\boldsymbol{B}^2$ 就不成立. 而 $\boldsymbol{A}$ 与 $\lambda\boldsymbol{E}$ 当然是可交换的, 所以牛顿二项式展开式只在下面的情形成立:

$$(\boldsymbol{A}+\lambda\boldsymbol{E})^n=\boldsymbol{A}^n+C_n^1\lambda\boldsymbol{A}^{n-1}+C_n^2\lambda^2\boldsymbol{A}^{n-2}+\cdots+C_n^{n-1}\lambda^{n-1}\boldsymbol{A}+\lambda^n\boldsymbol{E}.$$

(II) 不满足消去律:

(1) 当 $\boldsymbol{AB}=\boldsymbol{O}$ 时, 不能推出 $\boldsymbol{A}=\boldsymbol{O}$ 或 $\boldsymbol{B}=\boldsymbol{O}$.

(2) 当 $\boldsymbol{AB}=\boldsymbol{AC}$, 且 $\boldsymbol{A}\neq\boldsymbol{O}$ 时, 不能得到 $\boldsymbol{B}=\boldsymbol{C}$.

注意, 当 $\boldsymbol{A}$ 可逆时, 消去律是成立的, 即

当 $\boldsymbol{AB}=\boldsymbol{O}$ 时, 且 $|\boldsymbol{A}|\neq 0$ 时, 有 $\boldsymbol{B}=\boldsymbol{O}$;

当 $\boldsymbol{AB}=\boldsymbol{AC}$, 且 $|\boldsymbol{A}|\neq 0$ 时, 有 $\boldsymbol{B}=\boldsymbol{C}$.

矩阵运算不满足交换律, 不满足消去律, 还有一些过去熟知的公式在矩阵理论里并不成立. 前述是线性代数初学者最容易犯的几个错误之一, 为数不少的人会一直犯这个错误.

我们要注意, 虽然矩阵也有所谓的"加法"、"乘法", 但是这和我们熟知的实数加法、乘法是完全不同的. 运算的对象不同, 运算的内容不同, 当然, 运算的规律也不同. 这是两个不同的讨论范围里的不同运算, 相同的只不过是沿用了以前的称谓或记号而已, 我们不要被这一点"相同"而忘记二者本质的不同.

(二) 伴随矩阵.

(1) $\boldsymbol{AA}^*=\boldsymbol{A}^*\boldsymbol{A}=|\boldsymbol{A}|\boldsymbol{E}$. 这个公式要牢记! 其重大意义是由此引入了逆矩阵的讨论. 注意这里的 $\boldsymbol{A}$ 不一定是可逆的.

(2) 若 $|\boldsymbol{A}|\neq 0$, 则 $\boldsymbol{A}^{-1}=\dfrac{1}{|\boldsymbol{A}|}\boldsymbol{A}^*$. 它在理论上给出了求逆矩阵的方法, 但是并不实用, 在第三章将给出一个简单实用的方法(见教材 P.64 例 2).

(三) 逆矩阵

(1) 逆矩阵定义中的条件 "$\boldsymbol{AB}=\boldsymbol{BA}=\boldsymbol{E}$" 是可以弱化的: 设 $\boldsymbol{A}$, $\boldsymbol{B}$ 为方阵, 若 $\boldsymbol{AB}=\boldsymbol{E}$, 则 $\boldsymbol{A}$, $\boldsymbol{B}$ 可逆, 且互为逆矩阵. 更一般地, 对方阵而言, 若 $\boldsymbol{A}_1\boldsymbol{A}_2\cdots\boldsymbol{A}_k=\lambda\boldsymbol{E}$ 且 $\lambda\neq 0$, 则矩阵 $\boldsymbol{A}_1$, $\boldsymbol{A}_2$, $\cdots$, $\boldsymbol{A}_k$ 都是可逆的.

(2) 逆矩阵在运算中实现了除法的功能. 在矩阵中没有除法, 或者说, 我们通过引入逆矩阵, 避免了对除法的讨论.

对于任意的 n 阶方阵 $\boldsymbol{A}$ 都有 $\boldsymbol{AE}=\boldsymbol{EA}=\boldsymbol{A}$. 从乘法的角度来看, n 阶单位矩阵 $\boldsymbol{E}$ 在矩阵中的地位类似于 1 在实数中的地位. 一个实数 a $(a\neq 0)$ 的倒数可以用等式 $aa^{-1}=1$ 来刻划, 相仿地, 我们引入记号 $\boldsymbol{A}^{-1}$ 表示 $\boldsymbol{A}$ 的逆矩阵, 并且满足 $\boldsymbol{AA}^{-1}=\boldsymbol{E}$. 记号 $\boldsymbol{A}^{-1}$ 是特定的, 它不能写成 $\dfrac{1}{\boldsymbol{A}}$.

(3) 矩阵可逆的几个等价说法: 设 $\boldsymbol{A}$ 为 n 阶方阵,
矩阵 $\boldsymbol{A}$ 可逆 $\Longleftrightarrow |\boldsymbol{A}| \neq 0 \Longleftrightarrow \boldsymbol{A}$ 为非奇异矩阵.
其他的等价说法, 在以后会学习到.

2.2 其他重要公式与结论

- $|\lambda \boldsymbol{A}| = \lambda^n |\boldsymbol{A}|$. ($\boldsymbol{A}$ 为 n 阶方阵)
- $(\boldsymbol{A}\boldsymbol{B})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$. ($\boldsymbol{A}$, $\boldsymbol{B}$ 为可逆矩阵)
- $(\boldsymbol{A}\boldsymbol{B})^{\mathrm{T}} = \boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}$.
- $\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \dfrac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$, $(ad-bc \neq 0)$.
- $\begin{pmatrix} \boldsymbol{A}_1 & & & \\ & \boldsymbol{A}_2 & & \\ & & \ddots & \\ & & & \boldsymbol{A}_k \end{pmatrix}^{-1} = \begin{pmatrix} \boldsymbol{A}_1^{-1} & & & \\ & \boldsymbol{A}_1^{-1} & & \\ & & \ddots & \\ & & & \boldsymbol{A}_k^{-1} \end{pmatrix}$.
- $\begin{vmatrix} \boldsymbol{A}_1 & & & \\ & \boldsymbol{A}_2 & & \\ & & \ddots & \\ & & & \boldsymbol{A}_k \end{vmatrix} = |\boldsymbol{A}_1|\,|\boldsymbol{A}_2|\cdots|\boldsymbol{A}_k|$.
- $\begin{pmatrix} \boldsymbol{O} & \boldsymbol{A} \\ \boldsymbol{B} & \boldsymbol{O} \end{pmatrix}^{-1} = \begin{pmatrix} \boldsymbol{O} & \boldsymbol{B}^{-1} \\ \boldsymbol{A}^{-1} & \boldsymbol{O} \end{pmatrix}$. (P.56 习题 27)
- $\begin{vmatrix} \boldsymbol{O} & \boldsymbol{A}_{m\times m} \\ \boldsymbol{B}_{n\times n} & \boldsymbol{O} \end{vmatrix} = (-1)^{mn}\,|\boldsymbol{A}|\,|\boldsymbol{B}|$. (使用 Laplace 展开可得. 或者, 将 $\boldsymbol{A}$ 所在行的最后一行开始, 与 $\boldsymbol{B}$ 所在的 n 行进行相邻互换, 共进行 $m \times n$ 次互换后, 得到 $\begin{vmatrix} \boldsymbol{B}_{n\times n} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{A}_{m\times m} \end{vmatrix}$.)

3 线性方程组

秩是矩阵的一种内在属性. 矩阵的这种内在属性是与生俱来的, 一个矩阵一旦诞生, 它的这种内在属性就确定了. 虽然初等变换可以把它们变得面目全非, 但是它们的这个内在属性是不变的. 等价的矩阵, 看上去各各不同, 但是有一个内在属性是一样的, 那就是它们的秩.

3.1 本章重点

(一) 初等矩阵在矩阵乘法中的功能.

"复制、传递" —— 初等矩阵是怎么由单位阵得来的, 它将把该变换过程传递到它所乘的矩阵; 并遵循"左乘则行变, 右乘则列变"的特点.

具体而言, 设 $\boldsymbol{P}$ 为初等矩阵,

(i) 若 $\boldsymbol{P}$ 左乘矩阵 $\boldsymbol{A}$, 则 $\boldsymbol{PA}$ 的结果是: 把矩阵 $\boldsymbol{A}$ 进行初等行变换, 并且 $\boldsymbol{P}$ 是怎样由单位矩阵 $\boldsymbol{E}$ 通过行变换得来, 矩阵 $\boldsymbol{A}$ 就进行完全相同的行变换.

(ii) 若 $\boldsymbol{P}$ 右乘矩阵 $\boldsymbol{A}$, 则 $\boldsymbol{AP}$ 的结果是: 把矩阵 $\boldsymbol{A}$ 进行初等列变换, 并且 $\boldsymbol{P}$ 是怎样由单位矩阵 $\boldsymbol{E}$ 通过列变换得来, 矩阵 $\boldsymbol{A}$ 就进行完全相同的列变换.

上述过程体现为初等矩阵 $\boldsymbol{P}$ 把自己生成的过程复制、传递到了矩阵 $\boldsymbol{A}$ 上.

设 $\boldsymbol{P}=\begin{pmatrix}1&0&k&0\\&1&0&0\\&&1&0\\&&&1\end{pmatrix}$, $\boldsymbol{A}=\begin{pmatrix}a_1&a_2&a_3&a_4\\b_1&b_2&b_3&b_4\\c_1&c_2&c_3&c_4\\d_1&d_2&d_3&d_4\end{pmatrix}$.

(i) 若计算 $\boldsymbol{PA}$, “左乘则行变”, 意味着把 $\boldsymbol{A}$ 进行初等行变换, $\boldsymbol{P}$ 是由单位矩阵 $\boldsymbol{E}$ 经行变换 r_1+kr_3 得来, 则把 $\boldsymbol{A}$ 就进行相同的行变换, 所以

$$\begin{pmatrix}1&0&k&0\\&1&0&0\\&&1&0\\&&&1\end{pmatrix}\begin{pmatrix}a_1&a_2&a_3&a_4\\b_1&b_2&b_3&b_4\\c_1&c_2&c_3&c_4\\d_1&d_2&d_3&d_4\end{pmatrix}=\begin{pmatrix}a_1+kc_1&a_2+kc_2&a_3+kc_3&a_4+kc_4\\b_1&b_2&b_3&b_4\\c_1&c_2&c_3&c_4\\d_1&d_2&d_3&d_4\end{pmatrix}.$$

(ii) 若计算 $\boldsymbol{AP}$, “右乘则列变”, 此时要视 $\boldsymbol{P}$ 是由单位矩阵 $\boldsymbol{E}$ 通过初等列变换 c_3+kc_1 得来, 并有

$$\begin{pmatrix}a_1&a_2&a_3&a_4\\b_1&b_2&b_3&b_4\\c_1&c_2&c_3&c_4\\d_1&d_2&d_3&d_4\end{pmatrix}\begin{pmatrix}1&0&k&0\\&1&0&0\\&&1&0\\&&&1\end{pmatrix}=\begin{pmatrix}a_1&a_2&a_3+ka_1&a_4\\b_1&b_2&b_3+kb_1&b_4\\c_1&c_2&c_3+kc_1&c_4\\d_1&d_2&d_3+kd_1&d_4\end{pmatrix}.$$

(二) 利用初等变换求 (1) $\boldsymbol{A}^{-1}$; (2) $\boldsymbol{A}^{-1}\boldsymbol{B}$ 或 $\boldsymbol{BA}^{-1}$.

在本章, 矩阵求逆和解矩阵方程, 在方法上已经完全更新.

虽然我们不再通过伴随矩阵求逆矩阵, 但用定义求伴随矩阵的方法仍然要熟练掌握.

(二) 矩阵的秩

(1) 初等变换不改变矩阵的秩. 或者用下面的两种方式表述:

(i) 若 $\boldsymbol{A}\sim\boldsymbol{B}$, 则 $R(\boldsymbol{A})=R(\boldsymbol{B})$. (但 $R(\boldsymbol{A})=R(\boldsymbol{B})$ 不能得 $\boldsymbol{A}\sim\boldsymbol{B}$, 除非两者是同型矩阵.)

(ii) 若 $\boldsymbol{P},\boldsymbol{Q}$ 可逆, 则 $R(\boldsymbol{PAQ})=R(\boldsymbol{A})$.

(2) 矩阵和、差、积的秩.

(i) $R(\boldsymbol{A})-R(\boldsymbol{B})\leqslant R(\boldsymbol{A}\pm\boldsymbol{B})\leqslant R(\boldsymbol{A})+R(\boldsymbol{B})$.

(ii) $R(\boldsymbol{A})+R(\boldsymbol{B})-n\leqslant R(\boldsymbol{AB})\leqslant\min\{R(\boldsymbol{A}),R(\boldsymbol{B})\}$. 其中 $\boldsymbol{A},\boldsymbol{B}$ 分别为 $s\times n$ 和 $n\times m$ 矩阵.

(三) 线性方程组有解判别

(1) 一般的方程 $\boldsymbol{Ax}=\boldsymbol{b}$ 的情形.

对 n 元线性方程组 $\boldsymbol{Ax}=\boldsymbol{b}$, 记 $\boldsymbol{B}=(\boldsymbol{A},\boldsymbol{b})$. 注意到 $R(\boldsymbol{B})$ 比 $R(\boldsymbol{A})$ 只多 0 或 1.

(i) 若 $R(\boldsymbol{B})=R(\boldsymbol{A})+1$, 则说明出现了矛盾方程, 导致方程组无解.

(ii) 若 $R(\boldsymbol{B})=R(\boldsymbol{A})$, 则没有矛盾方程, 方程组有解. 其中,

(a) 当 $R(\boldsymbol{B})=R(\boldsymbol{A})<n$ 时, 说明出现了自由未知量, 导致方程组有无限多解;

(b) 而 $R(\boldsymbol{B})=R(\boldsymbol{A})=n$ 时, 则没有出现自由未知量, 所以方程组有唯一解.

是否出现矛盾方程是方程组有解与否的关键; 是否出现自由未知量又是区分有无限多解和有唯一解的关键. 换成秩的角度去说问题, 就呈现为下面的表达:

n 元线性方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 有解 $\Longleftrightarrow R(\boldsymbol{A})=R(\boldsymbol{B})$. 且 $\begin{cases} R(\boldsymbol{A})=R(\boldsymbol{B})=n, & \text{有唯一解;} \\ R(\boldsymbol{A})=R(\boldsymbol{B})<n, & \text{有无限多解.} \end{cases}$

n 元线性方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 无解 $\Longleftrightarrow R(\boldsymbol{A})\neq R(\boldsymbol{B})$.

(2) 齐次方程组 $\boldsymbol{Ax}=\boldsymbol{0}$ 的情形.

齐次方程组 $\boldsymbol{Ax}=\boldsymbol{0}$ 是天然有解的, 它至少有一个解: 零解. 所以对齐次方程组 $\boldsymbol{Ax}=\boldsymbol{0}$, 我们关心的不是它有没有解, 而是它是否有非零解. 下面的结论要非常的清楚:

n 元齐次方程组 $\boldsymbol{Ax}=\boldsymbol{0}$ 只有零解的充要条件是 $R(\boldsymbol{A})=n$.

n 元齐次方程组 $\boldsymbol{Ax}=\boldsymbol{0}$ 有非零解的充要条件是 $R(\boldsymbol{A})<n$.

注意: (i) n 是未知量的个数, 或者说是矩阵 $\boldsymbol{A}$ 的列数. $R(\boldsymbol{A})<n$ 表明 $\boldsymbol{A}$ 的行阶梯型矩阵中非零行的行数, 小于 n, 说明出现了自由未知量, 导致方程组的不唯一, 所以有非零解.

(ii) 这里的矩阵 $\boldsymbol{A}$ 不一定是方阵. 这个结论较第一章 P.25 的定理 5 就更一般化了, 而且是充要条件.

线性代数的基本内容的中心问题是解线性方程组, 关于线性方程组的解的判别在理论上是极端重要的, 必须熟练掌握. 特别是本章的例题 12, 该题型给出了两种解法. 解法二是我们要特别推荐的: 克拉默法则我们总是记得的, 系数行列式 $|\boldsymbol{A}|\neq 0$ 就得到了使方程组有唯一解的 λ 的值, 然后将其他的 λ 值代入方程组, 直接解方程组即可. 其优点是避免对含参数矩阵的初等变换, 缺点是仅适用于系数矩阵为方阵的情形.

(四) 新添矩阵可逆的等价说法.

n 阶矩阵 $\boldsymbol{A}$ 可逆, 新添下列等价说法:

(1) $\boldsymbol{A}$ 是满秩矩阵; 或 $R(\boldsymbol{A})=n$. (不可逆矩阵又称为降秩矩阵.)

(2) $\boldsymbol{A}$ 的标准形是 $\boldsymbol{E}$; 或 $\boldsymbol{A}\sim\boldsymbol{E}$.

(3) $\boldsymbol{A}$ 可以表达为有限个初等矩阵的乘积.

(4) 齐次线性方程组 $\boldsymbol{Ax}=\boldsymbol{0}$ 只有零解.

(5) 非齐次线性方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 有唯一解.

注意, 第一章的克拉默法则只告诉了我们矩阵 $\boldsymbol{A}$ 可逆是方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 有唯一解的充分条件.

3.2 要避免的错误

求矩阵秩时, 行变换和列变换可以随意进行, 行变、列变、两者交叉进行, 都可以(因为"初等变换不改变矩阵的秩"). 而解方程和求逆矩阵时只能进行一种变换:

(1) 对增广矩阵 $\boldsymbol{B}=(\boldsymbol{A},\boldsymbol{b})$ 进行初等变换求解方程时, 只有行变换, 不能有列变换; (因该过程本质上是消元法, 当然只能方程与方程之间进行运算, 在对应的增广矩阵中就体现为初等行变换.)

(2) 用矩阵初等变换 $(\boldsymbol{A},\boldsymbol{E})\sim(\boldsymbol{E},\boldsymbol{A}^{-1})$ 求逆矩阵时, 只有行变换, 不能有列变换. 其他的情形类似.

4 向量组的线性相关性

最大无关组是原向量组的"全权代表". 最大无关组的概念, 才真正阐明了秩的涵义.

4.1 本章要点

(一) 向量的线性表示.

有关向量的线性表示, 下面的说法是等价的:

向量 $\boldsymbol{b}$ 能由向量组 $\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_m$ 线性表示.

$\Longleftrightarrow$ 线性方程组 $x_1\boldsymbol{a}_1+x_2\boldsymbol{a}_2+\cdots+x_m\boldsymbol{a}_m=\boldsymbol{b}$ 有解.

$\Longleftrightarrow$ $R(\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_m)=R(\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_m,\boldsymbol{b})$.

上述结论的朴素理解: $R(\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_m)=R(\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_m,\boldsymbol{b})$, 意味着往向量组 $\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_m$ 中添加向量 $\boldsymbol{b}$, 并没有使得向量组的秩增加, 其根本原因在于向量 $\boldsymbol{b}$ 能由 $\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_m$ 线性表示.

进而, $R(\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_m)=R(\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_m,\boldsymbol{b}_1,\boldsymbol{b}_2,\cdots,\boldsymbol{b}_s)$, 也可理解为往向量组 $\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_m$ 中添加向量 $\boldsymbol{b}_1,\boldsymbol{b}_2,\cdots,\boldsymbol{b}_s$, 并没有使得向量组的秩增加, 所以, 向量组 B: $\boldsymbol{b}_1,\boldsymbol{b}_2,\cdots,\boldsymbol{b}_s$ 能由向量组 A: $\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_m$ 线性表示的充分必要条件是

$$R(\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_m)=R(\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_m,\boldsymbol{b}_1,\boldsymbol{b}_2,\cdots,\boldsymbol{b}_s).$$

(二) 线性相关与线性无关.

对于线性相关, 下面的说法是等价的:

向量组 A: $\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_m$ $(m\geqslant 2)$ 线性相关.

$\Longleftrightarrow$ 向量组 A 中至少存在一个向量是其余 $m-1$ 个向量的线性组合.

$\Longleftrightarrow$ 线性方程组 $x_1\boldsymbol{a}_1+x_2\boldsymbol{a}_2+\cdots x_m\boldsymbol{a}_m=\boldsymbol{0}$ 有非零解.

$\Longleftrightarrow$ $\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_m$ 的秩小于向量的个数 m, 即 $R(\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_m)<m$.

对于线性无关, 下面的说法是等价的:

向量组 A: $\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_m$ $(m\geqslant 2)$ 线性无关.

$\Longleftrightarrow$ 线性方程组 $x_1\boldsymbol{a}_1+x_2\boldsymbol{a}_2+\cdots x_m\boldsymbol{a}_m=\boldsymbol{0}$ 只有零解.

$\Longleftrightarrow$ $R(\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_m)=m$.

从上述说法要得到的理解是:

(1) 注意向量组、方程组、矩阵问题的相互转换;

(2) 得到一个朴素的认识: $R(\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_m)<m$ 的根本原因在于, 列向量 $\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_m$ 中有多余的向量, 或说存在某向量可以被其他的向量线性表示, 当然整个向量组是线性相关的.

(三) 矩阵等价与向量组等价的区别与联系.

设有 n 维向量组 A: $\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_m$, n 维向量组 B: $\boldsymbol{b}_1,\boldsymbol{b}_2,\cdots,\boldsymbol{b}_m$. 矩阵 $\boldsymbol{A}=\big(\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_m\big)$, 矩阵 $\boldsymbol{B}=\big(\boldsymbol{b}_1,\boldsymbol{b}_2,\cdots,\boldsymbol{b}_m\big)$. 则

(1) 向量组等价, 可得矩阵等价; (注意这里所设的两向量组中向量的个数相同, 否则两矩阵的列数不同, 会导致矩阵不是同型矩阵, 就不能得到矩阵等价.)

(2) 矩阵等价, 不能得到向量组等价.

例如, 设 $\boldsymbol{A}=\big(\boldsymbol{a}_1,\boldsymbol{a}_2\big)=\begin{pmatrix}1&0\\0&1\\0&0\\0&0\end{pmatrix}$, $\boldsymbol{B}=\big(\boldsymbol{b}_1,\boldsymbol{b}_2\big)=\begin{pmatrix}0&0\\0&0\\1&0\\0&1\end{pmatrix}$, 由 $R(\boldsymbol{A})=R(\boldsymbol{B})=2$, 知 $\boldsymbol{A}\sim\boldsymbol{B}$.

但向量组 $\boldsymbol{a}_1,\boldsymbol{a}_2$ 与向量组 $\boldsymbol{b}_1,\boldsymbol{b}_2$ 不是等价的, 因为这里 $R(\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{b}_1,\boldsymbol{b}_2)=4$, 不满足两向量组等价的充要

条件 $R(\boldsymbol{a}_1,\boldsymbol{a}_2)=R(\boldsymbol{b}_1,\boldsymbol{b}_2)=R(\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{b}_1,\boldsymbol{b}_2)$. 见教材 P.84 定理 2 的推论, 两向量组等价的充要条件是 $R(\boldsymbol{A})=R(\boldsymbol{B})=R(\boldsymbol{A},\boldsymbol{B})$, 而不是 $R(\boldsymbol{A})=R(\boldsymbol{B})$.

(四) 研究最大无关组的意义.

最大无关组和原向量组是等价的, 是原向量组的简约, 更是原向量组的“全权代表”.

最大无关组从理论上弄清了, 用消元法解线性方程组时, 为什么最后会剩余稳定数量的方程, 事实上那些剩下的方程就是原方程组的“最大无关组”, 和原方程组是等价的, 是同解的.

线性相关、线性表示的概念也可以解释用消元法解线性方程组的相关问题: 方程组是“线性相关”的, 说明有多余的方程; 能被其他的方程“线性表示”的方程就是多余的. (“多余”是相对的, 方程的去、留不是绝对的, 因最大无关组一般不唯一.)

最大无关组也使线性方程组在解的表示上, 得到了简洁、完备的表达.

从更广泛的含义上看, 最大无关组还充当了坐标系的功能.

(五) “$n-r$”的含义.

定理 7 是说: 对 n 元齐次线性方程组 $\boldsymbol{Ax}=\boldsymbol{0}$, 设 $R(\boldsymbol{A})=r$, 则方程组 $\boldsymbol{Ax}=\boldsymbol{0}$ 的基础解系包含 $n-r$ 个向量.

r 是 $\boldsymbol{A}$ 的秩, 也是 $\boldsymbol{A}$ 的行阶梯型矩阵的非零行的行数, 是非自由未知量的个数. (非自由未知量一般取自非零行的第一个非零元所对应的未知量, 一个非零行只能确定一个非自由未知量.)

n 是未知量的总数, 所以 $n-r$ 是自由未知量的个数. 有多少个自由未知量, 基础解系里就对应有多少个向量.

4.2 重新理解矩阵秩的性质

矩阵秩的性质在第三章列出了 8 条. 我们对其中的几条重要性质重新加以证明和理解.

例 1 (**性质** 5) 证明: $\max\{R(\boldsymbol{A}),R(\boldsymbol{B})\}\leqslant R(\boldsymbol{A},\boldsymbol{B})\leqslant R(\boldsymbol{A})+R(\boldsymbol{B})$.

证明 因为 $\boldsymbol{A}$ 的列均可由 $(\boldsymbol{A},\boldsymbol{B})$ 的列线性表出, 所以

$$R(\boldsymbol{A})\leqslant R(\boldsymbol{A},\boldsymbol{B}), \tag{1}$$

同理 $R(\boldsymbol{B})\leqslant R(\boldsymbol{A},\boldsymbol{B})$. 所以

$$\max\{R(\boldsymbol{A}),R(\boldsymbol{B})\}\leqslant R(\boldsymbol{A},\boldsymbol{B}).$$

设 $\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_r$ 为 $\boldsymbol{A}$ 的列向量的极大线性无关组, $\boldsymbol{b}_1,\boldsymbol{b}_2,\cdots,\boldsymbol{b}_s$ 为 $\boldsymbol{B}$ 的列向量的极大线性无关组. 则 $(\boldsymbol{A},\boldsymbol{B})$ 的列向量均可由向量组 $\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_r,\boldsymbol{b}_1,\boldsymbol{b}_2,\cdots,\boldsymbol{b}_s$ 线性表出, 所以

$$R(\boldsymbol{A},\boldsymbol{B})\leqslant R(\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_r,\boldsymbol{b}_1,\boldsymbol{b}_2,\cdots,\boldsymbol{b}_s).$$

而向量组 $\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_r,\boldsymbol{b}_1,\boldsymbol{b}_2,\cdots,\boldsymbol{b}_s$ 的秩不可能超过其向量的个数 $r+s$, 即 $R(\boldsymbol{A})+R(\boldsymbol{B})$, 所以

$$R(\boldsymbol{A},\boldsymbol{B})\leqslant R(\boldsymbol{A})+R(\boldsymbol{B}). \tag{2}$$

得证 $\max\{R(\boldsymbol{A}),R(\boldsymbol{B})\}\leqslant R(\boldsymbol{A},\boldsymbol{B})\leqslant R(\boldsymbol{A})+R(\boldsymbol{B})$.

注 1 对 (1) 式的朴素理解是, 在矩阵 $\boldsymbol{A}$ 的右侧添加新的列, 只会有可能使秩在原来的基础上得到增加; 当 $\boldsymbol{B}$ 的列向量能被 $\boldsymbol{A}$ 的列向量线性表出时, 等号成立.

对 (2) 式的朴素理解是, 对矩阵 $(\boldsymbol{A},\boldsymbol{B})$, 有可能 $\boldsymbol{A}$ 的列向量与 $\boldsymbol{B}$ 的列向量出现线性相关, 合并为 $(\boldsymbol{A},\boldsymbol{B})$ 的秩一般会比 $R(\boldsymbol{A})+R(\boldsymbol{B})$ 要小. 当$\boldsymbol{A}$ 和 $\boldsymbol{B}$ 两者列向量的极大线性无关组线性无关时, (2) 式的等号成立. 更极端的情形是 $\boldsymbol{A}$ 的列向量组与 $\boldsymbol{B}$ 的列向量组线性无关.

矩阵秩的性质 5 其实还可以写成

$$\max\{R(\boldsymbol{A}),R(\boldsymbol{B})\}\leqslant R\begin{pmatrix}\boldsymbol{A}\\ \boldsymbol{B}\end{pmatrix}\leqslant R(\boldsymbol{A})+R(\boldsymbol{B}).$$

上式第一个不等号也是说明, 给一个矩阵添加行, 有可能使得矩阵的秩增加.

例 2 (**性质** 6) 证明: $R(\boldsymbol{A}+\boldsymbol{B})\leqslant R(\boldsymbol{A})+R(\boldsymbol{B})$.

证明 因为 $\boldsymbol{A}+\boldsymbol{B}$ 的列均可由 $(\boldsymbol{A},\boldsymbol{B})$ 的列线性表出, 所以 $R(\boldsymbol{A}+\boldsymbol{B})\leqslant R(\boldsymbol{A},\boldsymbol{B})\leqslant R(\boldsymbol{A})+R(\boldsymbol{B})$. 得证

$$R(\boldsymbol{A}+\boldsymbol{B})\leqslant R(\boldsymbol{A})+R(\boldsymbol{B}). \tag{3}$$

注 2 注意 (2) 式、(3) 式的右侧都是 $R(\boldsymbol{A})+R(\boldsymbol{B})$. 就是说把矩阵 $\boldsymbol{A}$ 和$\boldsymbol{B}$ 合并、相加, 只可能使秩得以减少.

例 3 (**性质** 7) 证明: $R(\boldsymbol{AB})\leqslant\min\{R(\boldsymbol{A}),R(\boldsymbol{B})\}$.

证明 矩阵 $\boldsymbol{AB}$ 的列向量是矩阵 $\boldsymbol{A}$ 的列向量的线性组合, 事实上, 设

$$\boldsymbol{AB}=(\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_m)\begin{pmatrix}b_{11}&b_{12}&\cdots&b_{1s}\\ b_{21}&b_{22}&\cdots&b_{2s}\\ \vdots&\vdots&&\vdots\\ b_{m1}&b_{m2}&\cdots&b_{ms}\end{pmatrix},$$

知矩阵 $\boldsymbol{AB}$ 的第 1 列为 $b_{11}\boldsymbol{a}_1+b_{21}\boldsymbol{a}_2+\cdots+b_{m1}\boldsymbol{a}_m$, …, 第 s 列为 $b_{1s}\boldsymbol{a}_1+b_{2s}\boldsymbol{a}_2+\cdots+b_{ms}\boldsymbol{a}_m$. 矩阵 $\boldsymbol{AB}$ 的列向量可以被矩阵 $\boldsymbol{A}$ 的列向量线性表示, 由 P.85 定理 3, 知

$$R(\boldsymbol{AB})\leqslant R(\boldsymbol{A}).$$

类似地, 矩阵 $\boldsymbol{AB}$ 的行向量是矩阵 $\boldsymbol{B}$ 的行向量的线性组合, 有 $R(\boldsymbol{AB})\leqslant R(\boldsymbol{B})$. 得证 $R(\boldsymbol{AB})\leqslant\min\{R(\boldsymbol{A}),R(\boldsymbol{B})\}$.

从这个性质及 P.85 定理 3 得到的共同理解是: 对一个向量组进行线性组合可能会使向量组的秩减小.

5 相似矩阵及二次型

矩阵乘以向量, 在功能上相当于把一个向量变换为另一个向量. 一个矩阵的特征向量是这样一种特定的向量, 它经过这种变换后方向不变(或正好反向), 只发生长度上的伸缩. 特征值则反映了特征向量的伸缩倍数(及方向).

5.1 重点释疑

(一) 怎么理解特征值、特征向量.

矩阵 $\boldsymbol{A}$ 左乘向量 $\boldsymbol{x}$, 其结果是一个同维数的向量, 比如 $\boldsymbol{A}=\begin{pmatrix}-2&1&1\\0&2&0\\-4&1&3\end{pmatrix}$, 取 $\boldsymbol{x}=\begin{pmatrix}1\\2\\3\end{pmatrix}$, 有

$$\boldsymbol{Ax}=\begin{pmatrix}-2&1&1\\0&2&0\\-4&1&3\end{pmatrix}\begin{pmatrix}1\\2\\3\end{pmatrix}=\begin{pmatrix}3\\4\\7\end{pmatrix},$$

可见矩阵 $\boldsymbol{A}$ 左乘向量 $\boldsymbol{x}$, 相当于把向量 $\boldsymbol{x}$ 作了一个变换, 把 $\boldsymbol{x}$ 转到了一个新的位置, 而且长度也发生了变化. 这里 $(1,2,3)^{\mathrm{T}}$ 与变换之后的结果 $(3,4,7)^{\mathrm{T}}$ 看不出有什么关联, 但是, 矩阵 $\boldsymbol{A}$ 左乘某些特定的向量 $\boldsymbol{x}$, 会出现比较特别的现象, 就是乘积的结果相当于把向量 $\boldsymbol{x}$ 在原方向伸缩或反方向伸缩, 即

$$\boldsymbol{Ax}=\lambda\boldsymbol{x},$$

具有这种特点的向量, 就称为矩阵 $\boldsymbol{A}$ 的特征向量, 数 λ 反映了伸缩的倍数及方向, 称为与 $\boldsymbol{x}$ 对应的特征值.

比如 $\boldsymbol{x}_1=(1,0,1)^{\mathrm{T}}$, $\boldsymbol{x}_2=(0,1,-1)^{\mathrm{T}}$ 就是矩阵 $\boldsymbol{A}$ 的特征向量:

$$\boldsymbol{Ax_1}=\begin{pmatrix}-2&1&1\\0&2&0\\-4&1&3\end{pmatrix}\begin{pmatrix}1\\0\\1\end{pmatrix}=\begin{pmatrix}-1\\0\\-1\end{pmatrix}=-\boldsymbol{x}_1,$$

$$\boldsymbol{Ax_2}=\begin{pmatrix}-2&1&1\\0&2&0\\-4&1&3\end{pmatrix}\begin{pmatrix}0\\1\\-1\end{pmatrix}=\begin{pmatrix}0\\2\\-2\end{pmatrix}=2\boldsymbol{x}_2.$$

一个 n 阶矩阵 $\boldsymbol{A}$ 一旦产生, 那么 n 维空间[1] 的某处就存在着一些向量, 它们与矩阵 $\boldsymbol{A}$ 有着一种天然的内在联系: $\boldsymbol{A}$ 乘以这些向量相当于只是把这些向量在原方向或反方向上伸长或缩短.

特征值、特征向量的专业词汇分别是 eigenvalue, eigenvector. eigen 是德文词汇, 意思是自己的, 特有的. eigen 一词很恰当地反映了矩阵和其特征向量的天然联系和隶属性. 特征值在一些教材里称为“本征值”, 这个翻译比较贴近 eigenvalue 的本意.

矩阵 $\boldsymbol{A}$ 左乘零向量总是等于零向量的, 所以, 讨论特征向量时, 是把零向量排除在外的. 请牢记: 零向量不是特征向量; 特征向量是非零向量.

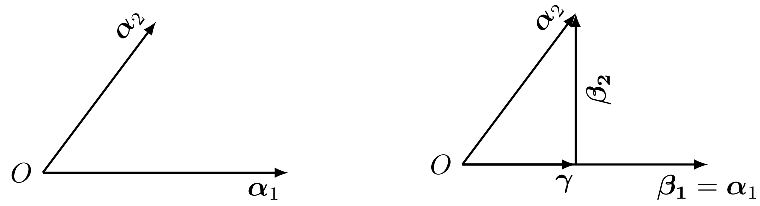
(二) 施密特正交化方法怎么理解.

以三维空间为例. 设有不共面的三个向量 $\boldsymbol{\alpha}_1$, $\boldsymbol{\alpha}_2$, $\boldsymbol{\alpha}_3$, 由这三个向量, 我们构造一组两两正交的向量 $\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$, $\boldsymbol{\beta}_3$.

设有两个向量 $\boldsymbol{\alpha}_1$, $\boldsymbol{\alpha}_2$ 如图 1(a).

先直接取 $\boldsymbol{\beta}_1=\boldsymbol{\alpha}_1$, 如图 1(b).

[1]准确地讲, 应该是 n 维复向量空间. 因为方阵在实数域上不一定总有特征值. 比如 $\boldsymbol{A}=\begin{pmatrix}0&-1\\1&0\end{pmatrix}$, $|\boldsymbol{A}-\lambda\boldsymbol{E}|=\lambda^2+1$, 在实数域上无根, 故 $\boldsymbol{A}$ 在实数域上无特征值. 教材也讲明了: n 阶矩阵 $\boldsymbol{A}$ 在复数范围内有 n 个特征值.



(a) (b)

图 1:

记 $\boldsymbol{\alpha}_2$ 在 $\boldsymbol{\alpha}_1$ 上的投影向量为 $\boldsymbol{\gamma}$, 则可令 $\boldsymbol{\beta}_2 \triangleq \boldsymbol{\alpha}_2 - \boldsymbol{\gamma}$, 使得 $\boldsymbol{\beta}_2 \perp \boldsymbol{\beta}_1$. 下面给出 $\boldsymbol{\gamma}$ 的表达式.

$\boldsymbol{\alpha}_2$ 在 $\boldsymbol{\alpha}_1$ 上的投影为:

$$\mathrm{Prj}_{\boldsymbol{\alpha}_1} \boldsymbol{\alpha}_2 = \|\boldsymbol{\alpha}_2\| \cdot \cos(\widehat{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2}) = \|\boldsymbol{\alpha}_2\| \cdot \frac{[\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2]}{\|\boldsymbol{\alpha}_1\|\|\boldsymbol{\alpha}_2\|} = \frac{[\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2]}{\|\boldsymbol{\alpha}_1\|},$$

则 $\boldsymbol{\alpha}_2$ 在 $\boldsymbol{\alpha}_1$ 上的投影向量为:

$$\boldsymbol{\gamma} = \mathrm{Prj}_{\boldsymbol{\alpha}_1} \boldsymbol{\alpha}_2 \cdot \frac{\boldsymbol{\alpha}_1}{\|\boldsymbol{\alpha}_1\|} = \frac{[\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2]}{\|\boldsymbol{\alpha}_1\|} \cdot \frac{\boldsymbol{\alpha}_1}{\|\boldsymbol{\alpha}_1\|} = \frac{[\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2]}{[\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_1]}\boldsymbol{\alpha}_1. \tag{4}$$

所以

$$\boldsymbol{\beta}_2 = \boldsymbol{\alpha}_2 - \boldsymbol{\gamma} = \boldsymbol{\alpha}_2 - \frac{[\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2]}{[\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_1]}\boldsymbol{\alpha}_1 = \boldsymbol{\alpha}_2 - \frac{[\boldsymbol{\beta}_1, \boldsymbol{\alpha}_2]}{[\boldsymbol{\beta}_1, \boldsymbol{\beta}_1]}\boldsymbol{\beta}_1.$$

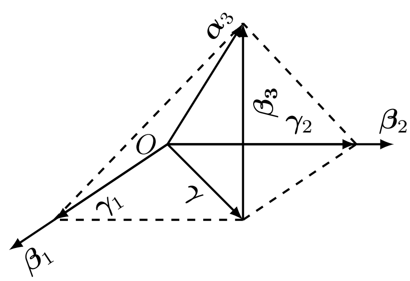


图 2:

记 $\boldsymbol{\alpha}_3$ 在 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2$ 所在平面的投影向量为 $\boldsymbol{\gamma}$, 若令 $\boldsymbol{\beta}_3 \triangleq \boldsymbol{\alpha}_3 - \boldsymbol{\gamma}$, 则 $\boldsymbol{\beta}_3$ 垂直于 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2$ 所在的平面, 从而 $\boldsymbol{\beta}_3 \perp \boldsymbol{\beta}_1$, $\boldsymbol{\beta}_3 \perp \boldsymbol{\beta}_2$, 如图 2.

记 $\boldsymbol{\alpha}_3$ 在 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2$ 上的投影向量分别为 $\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2$, 则 $\boldsymbol{\gamma} = \boldsymbol{\gamma}_1 + \boldsymbol{\gamma}_2$. 与 (4) 式同理有

$$\boldsymbol{\gamma}_1 = \frac{[\boldsymbol{\beta}_1, \boldsymbol{\alpha}_3]}{[\boldsymbol{\beta}_1, \boldsymbol{\beta}_1]}\boldsymbol{\beta}_1, \qquad \boldsymbol{\gamma}_2 = \frac{[\boldsymbol{\beta}_2, \boldsymbol{\alpha}_3]}{[\boldsymbol{\beta}_2, \boldsymbol{\beta}_2]}\boldsymbol{\beta}_2$$

所以

$$\boldsymbol{\beta}_3 = \boldsymbol{\alpha}_3 - \boldsymbol{\gamma} = \boldsymbol{\alpha}_3 - \boldsymbol{\gamma}_1 - \boldsymbol{\gamma}_2 = \boldsymbol{\alpha}_3 - \frac{[\boldsymbol{\beta}_1, \boldsymbol{\alpha}_3]}{[\boldsymbol{\beta}_1, \boldsymbol{\beta}_1]}\boldsymbol{\beta}_1 - \frac{[\boldsymbol{\beta}_2, \boldsymbol{\alpha}_3]}{[\boldsymbol{\beta}_2, \boldsymbol{\beta}_2]}\boldsymbol{\beta}_2.$$

对一般的 n 维空间, 其施密特正交化公式可以类似地理解和记忆.

(三) 为什么要强调正交变换.

本章的知识是为二次型化标准形服务的. 线性代数主要讨论线性函数, 二次型按理不在讨论之列, 但二次型化标准形可以转化为矩阵的问题, 是矩阵的一个应用.

把二次型化为标准形, 可以方便地研究其性质. 例如中心是原点的二次曲线 $x^2-xy+y^2=4$, 让曲线绕原点旋转适当角度, 令

$$\begin{pmatrix}\widetilde{x}\\ \widetilde{y}\end{pmatrix}=\begin{pmatrix}\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix},$$

可以使它化为标准形

$$\frac{\widetilde{x}^2}{8}+\frac{\widetilde{y}^2}{\frac{8}{3}}=1,$$

从这个标准形, 我们容易识别曲线的类别, 研究曲线的性质. 为使曲线在旋转过程中保持形状不变, 要求使用正交变换.

不使用正交变换, 也可以把二次型 x^2-xy+y^2 标准化. 比如, 由 $x^2-xy+y^2=(x-\frac{1}{2}y)^2+\frac{3}{4}y^2$, 令 $\widetilde{x}=(x-\frac{1}{2}y)$, $\widetilde{y}=\frac{\sqrt{3}}{2}y$, 得 $\widetilde{x}^2+\widetilde{y}^2=4$. 很显然, 这种非正交的变换, 改变了曲线的形状.

(四) 例题 12 的说明.

教材 P.125 例题 12 给出了对称矩阵正交对角化的示例. 这是本章基本而重要的题型, 强调几点要注意的事项. 为表述方便, 把例题的主要内容罗列如下:

$$\boldsymbol{A}=\begin{pmatrix}0 & -1 & 1\\ -1 & 0 & 1\\ 1 & 1 & 0\end{pmatrix},$$

特征值为 $\lambda_1=-2$, $\lambda_2=\lambda_3=1$.

对应 $\lambda_1=-2$ 的特征向量取为 $\boldsymbol{\xi}_1=(-1,-1,1)^{\mathrm{T}}$, 单位化得 $\boldsymbol{p}_1=\frac{1}{\sqrt{3}}(-1,-1,1)^{\mathrm{T}}$.

对应 $\lambda_2=\lambda_3=1$ 的线性无关的特征向量取为 $\boldsymbol{\xi}_2=(-1,1,0)^{\mathrm{T}}$, $\boldsymbol{\xi}_3=(1,0,1)^{\mathrm{T}}$. 正交化得 $\boldsymbol{\eta}_2=(-1,1,0)^{\mathrm{T}}$, $\boldsymbol{\eta}_3=\frac{1}{2}(1,1,2)^{\mathrm{T}}$. 再单位化得 $\boldsymbol{p}_2=\frac{1}{\sqrt{2}}(-1,1,0)^{\mathrm{T}}$, $\boldsymbol{p}_3=\frac{1}{\sqrt{6}}(1,1,2)^{\mathrm{T}}$.

取正交矩阵 $\boldsymbol{P}=(\boldsymbol{p}_1,\boldsymbol{p}_2,\boldsymbol{p}_3)=\begin{pmatrix}-\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}}\\ -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}}\\ \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}}\end{pmatrix}$, 得

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{\Lambda}=\begin{pmatrix}-2 & & \\ & 1 & \\ & & 1\end{pmatrix}.$$

(1) 要看清楚题目的要求. 如果没有要求 $\boldsymbol{P}$ 为正交阵, 则作一般的对角化即可, 即取

$$\boldsymbol{P}=(\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\boldsymbol{\xi}_3)=\begin{pmatrix}-1 & -1 & 1\\ -1 & 1 & 0\\ 1 & 0 & 1\end{pmatrix},$$

则有

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{\Lambda} = \begin{pmatrix} -2 & & \\ & 1 & \\ & & 1 \end{pmatrix}.$$

(2) 正交化过程只出现在重根特征值对应的特征向量. 因为对称矩阵 $\boldsymbol{A}$ 的属于不同特征值的特征向量是正交的, 所以只需要将重根特征值对应的特征向量进行施密特正交化, 而无需将 $\boldsymbol{A}$ 所有的特征向量放在一起正交化.

(3) $\boldsymbol{P}$ 中 $\boldsymbol{p}_1, \boldsymbol{p}_2, \boldsymbol{p}_3$ 的排列顺序要和 $\boldsymbol{\Lambda}$ 中对角元素 $\lambda_1, \lambda_2, \lambda_3$ 的排列顺序一致. 比如若取

$$\boldsymbol{P} = (\boldsymbol{p}_2, \boldsymbol{p}_1, \boldsymbol{p}_3) = \begin{pmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \\ 0 & \frac{1}{\sqrt{3}} & \frac{2}{\sqrt{6}} \end{pmatrix},$$

则由 $\boldsymbol{A}(\boldsymbol{p}_2, \boldsymbol{p}_1, \boldsymbol{p}_3) = (\lambda_2\boldsymbol{p}_2, \lambda_1\boldsymbol{p}_1, \lambda_3\boldsymbol{p}_3) = (\boldsymbol{p}_2, \boldsymbol{p}_1, \boldsymbol{p}_3)\,\mathrm{diag}(\lambda_2, \lambda_1, \lambda_3)$, 得

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{\Lambda} = \begin{pmatrix} 1 & & \\ & -2 & \\ & & 1 \end{pmatrix}.$$

(4) 正交化过程其实是可以避免的: 对重根求对应方程组的基础解系时, 可以直接凑出一组正交的基础解系. 对应 $\lambda_2 = \lambda_3 = 1$, 求解 $(\boldsymbol{A} - \boldsymbol{E})\boldsymbol{x} = \boldsymbol{0}$ 时, 得同解方程 $x_1 + x_2 - x_3 = 0$(见教材 P.125), 下面构造一组正交的基础解系: 若取 $\boldsymbol{\xi}_2 = (1, 0, 1)^{\mathrm{T}}$, 要满足正交, 则应有 $\boldsymbol{\xi}_3 = (1, \square, -1)^{\mathrm{T}}$, 又要满足方程, 所以 $\boldsymbol{\xi}_3 = (1, -2, -1)^{\mathrm{T}}$, 从而得到一组正交的基础解系:

$$\boldsymbol{\xi}_2 = (1, 0, 1)^{\mathrm{T}}, \qquad \boldsymbol{\xi}_3 = (1, -2, -1)^{\mathrm{T}}. \tag{5}$$

类似地, 还可以取

$$\boldsymbol{\xi}_2 = (1, -1, 0)^{\mathrm{T}}, \qquad \boldsymbol{\xi}_3 = (1, 1, 2)^{\mathrm{T}}. \tag{6}$$

或者 $\boldsymbol{\xi}_2 = (0, 1, 1)^{\mathrm{T}}$, $\boldsymbol{\xi}_3 = (-2, 1, -1)^{\mathrm{T}}$, 等等.

(5) 满足条件的正交阵不是唯一的(即便我们忽略 $\boldsymbol{P}$ 中 $\boldsymbol{p}_1, \boldsymbol{p}_2, \boldsymbol{p}_3$ 的排列顺序). 把 (5) 式中的向量单位化得

$$\boldsymbol{p}_2 = \frac{1}{\sqrt{2}}(1, 0, 1)^{\mathrm{T}}, \qquad \boldsymbol{p}_3 = \frac{1}{\sqrt{6}}(1, -2, -1)^{\mathrm{T}},$$

从而得满足条件的正交阵

$$\boldsymbol{P} = (\boldsymbol{p}_1, \boldsymbol{p}_2, \boldsymbol{p}_3) = \begin{pmatrix} -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{3}} & 0 & -\frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \end{pmatrix}.$$

把 (5) 式中的向量单位化得满足条件的正交阵

$$\boldsymbol{P} = (\boldsymbol{p}_1, \boldsymbol{p}_2, \boldsymbol{p}_3) = \begin{pmatrix} -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \end{pmatrix}.$$

5.2 重要结论

(一) 特征值、特征向量的性质.

(1) 设 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 是 n 阶矩阵 $\boldsymbol{A}$ 的 n 个特征值, 则

(a) $\lambda_1 + \lambda_2 + \cdots + \lambda_n = a_{11} + a_{22} + \cdots + a_{nn}$;

(b) $\lambda_1 \lambda_2 \cdots \lambda_n = \left|\boldsymbol{A}\right|$.

(2) 设 λ_0 是 $\boldsymbol{A}$ 的一个特征值, $\boldsymbol{x}$ 是 $\boldsymbol{A}$ 的对应于 λ_0 的特征向量, 则

(a) 若 $\boldsymbol{A}$ 可逆, 则 $\lambda_0 \neq 0$, 且 $\dfrac{1}{\lambda_0}$ 是 $\boldsymbol{A}^{-1}$ 的特征值; $\dfrac{|\boldsymbol{A}|}{\lambda_0}$ 是 $\boldsymbol{A}^*$ 的特征值.

(b) $k\lambda_0$ 为 $k\boldsymbol{A}$ 特征值; λ_0^m 是 $\boldsymbol{A}^m$ 特征值.

(c) $\varphi(\lambda_0)$ 是矩阵多项式 $\varphi(\boldsymbol{A})$ 的特征值, 其中

$$
\begin{aligned}
\varphi(\lambda_0) &= a_m \lambda_0^m + \cdots + a_1 \lambda_0 + a_0, \\
\varphi(\boldsymbol{A}) &= a_m \boldsymbol{A}^m + \cdots + a_1 \boldsymbol{A} + a_0 \boldsymbol{E}.
\end{aligned}
$$

而且 $\boldsymbol{x}$ 仍然是矩阵 $\boldsymbol{A}^{-1}$, $\boldsymbol{A}^*$, $k\boldsymbol{A}$, $\boldsymbol{A}^m$, $\varphi(\boldsymbol{A})$ 的分别对应于特征值 $\dfrac{1}{\lambda_0}$, $\dfrac{|\boldsymbol{A}|}{\lambda_0}$, $k\lambda_0$, λ_0^m, $\varphi(\lambda_0)$ 的特征向量.

(3) 特征向量之间的关系:

(a) 矩阵 $\boldsymbol{A}$ 的属于不同特征值的特征向量是线性无关的. (教材 P.120 定理 2.)

(b) 设 $\lambda_1, \lambda_2, \cdots, \lambda_m$ 是矩阵 $\boldsymbol{A}$ 的 m 个互异特征值, 对应于 λ_i $(i = 1, 2, \cdots, m)$ 的线性无关的特征向量有 r_i 个, 则由所有这些特征向量(共 $r_1 + r_2 + \cdots + r_m$ 个)构成的向量组是线性无关的.

(c) 对称矩阵 $\boldsymbol{A}$ 的属于不同特征值的特征向量, 是两两正交的. (教材 P.124 定理 6.)

(4) 特征值所对应的特征向量的个数:

(a) 每个特征值都对应着至少一个特征向量.

(b) k 重特征值对应的线性无关的特征向量的个数不超过 k.

(c) 若 $\boldsymbol{A}$ 为对称阵, 则 $\boldsymbol{A}$ 的每个特征值对应的线性无关特征向量的个数恰好等于该特征值的重数.(教材 P.124 定理 7 的推论.)

这几个结论可用来初步排除计算中的错误, 比如单根却对应着多个线性无关的特征向量, 或者 $(\boldsymbol{A} - \lambda_0 \boldsymbol{E})\boldsymbol{x} = \boldsymbol{0}$ 的基础解系中向量的个数超过了 λ_0 的重数, 等等.

(5) 方阵 $\boldsymbol{A}$ 可逆 $\Longleftrightarrow$ 0 不是 $\boldsymbol{A}$ 的特征值.

(二) 正交矩阵的性质.

$$
\begin{aligned}
&\text{方阵 } \boldsymbol{A} \text{ 为正交矩阵} \\
\Longleftrightarrow\ & \boldsymbol{A}^{\mathrm{T}} \boldsymbol{A} = \boldsymbol{E} \iff \boldsymbol{A}\boldsymbol{A}^{\mathrm{T}} = \boldsymbol{E} \\
\Longleftrightarrow\ & \boldsymbol{A} \text{ 可逆, 且 } \boldsymbol{A}^{-1} = \boldsymbol{A}^{\mathrm{T}} \\
\Longleftrightarrow\ & \boldsymbol{A} \text{ 的行(列)向量组两两正交, 且都是单位向量} \\
\Longleftrightarrow\ & \boldsymbol{A} \text{ 的行(列)向量组是一组规范正交基.}
\end{aligned}
$$

若 $\boldsymbol{A}$ 为正交矩阵, 则 $|\boldsymbol{A}| = \pm 1$; 其特征值 λ 满足 $|\lambda| = 1$. 正交变换具有保持向量的内积、长度、夹角不变的特性.

(三) 矩阵对角化.

矩阵 $\boldsymbol{A}$ 与对角阵相似的充要条件是: $\boldsymbol{A}$ 有 n 个线性无关的特征向量.

矩阵 $\boldsymbol{A}$ 与对角阵相似的充要条件是: $\boldsymbol{A}$ 的每个特征值对应的线性无关特征向量的个数等于该特征值的重数.

n 阶矩阵 $\boldsymbol{A}$ 有 n 个互不相同的特征值, 则 $\boldsymbol{A}$ 可以对角化. 此条件是充分的, 但不是必要的.

矩阵对角化这个问题非常重要. 教材上的例题和习题已经很好了(P.123 例题 11, P.135 习题 15, 16, 20, 21, 22).

(四) 对称矩阵正定的充要条件.

对称矩阵 $\boldsymbol{A}$ 为正定的 $\Longleftrightarrow$ 二次型 $f(\boldsymbol{x}) = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 为正定的 $\Longleftrightarrow$ $\boldsymbol{A}$ 的特征值全为正 $\Longleftrightarrow$ $\boldsymbol{A}$ 的各阶主子式都为正.

(五) 等价、相似、合同、正交相似.

设 $\boldsymbol{A}$, $\boldsymbol{B}$ 均为 $m \times n$ 矩阵, $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 等价 $\Longleftrightarrow$ 存在 m 阶可逆阵 $\boldsymbol{P}$ 和 n 阶可逆阵 $\boldsymbol{Q}$, 使 $\boldsymbol{PAQ} = \boldsymbol{B}$.

设 $\boldsymbol{A}$, $\boldsymbol{B}$ 均为 n 阶方阵,

$\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相似 $\Longleftrightarrow$ 存在可逆阵 $\boldsymbol{P}$, 使 $\boldsymbol{P}^{-1}\boldsymbol{AP} = \boldsymbol{B}$.

$\boldsymbol{A}$ 与 $\boldsymbol{B}$ 合同 $\Longleftrightarrow$ 存在可逆阵 $\boldsymbol{P}$, 使 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{AP} = \boldsymbol{B}$.

$\boldsymbol{A}$ 与 $\boldsymbol{B}$ 正交相似 $\Longleftrightarrow$ 存在可逆阵 $\boldsymbol{P}$, 使 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{AP} = \boldsymbol{P}^{-1}\boldsymbol{AP} = \boldsymbol{B}$.

等价、相似、合同、正交相似的区别和联系:

- 等价的矩阵不必是方阵, 后面三个都是方阵之间的关系.
- 相似、合同、正交相似都是等价的一种; 正交相似关系最强, 等价关系最弱.
- 相似与合同没有什么关系, 仅当 $\boldsymbol{P}$ 为正交阵时, 有 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{AP} = \boldsymbol{P}^{-1}\boldsymbol{AP}$, 这时相似与合同是一致的.