线性方程组有解判别准则的另一种解释:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \cdots \\ b_n \end{pmatrix}$$

$$\begin{array}{cccc} \beta_1 & \beta_2 & & \beta_n \end{array} \qquad b$$

$A=(\beta_1 \beta_2 \cdots \beta_n)$

$B=(\beta_1 \beta_2 \cdots \beta_n\, b)$

线性方程组有解

$\Updownarrow$

$r(A)=r(B)$ $\Longleftrightarrow$ $r\{\beta_1,\beta_2,\cdots,\beta_n\} = r\{\beta_1,\beta_2,\cdots,\beta_n,b\}$

$\Updownarrow$

b 可由 $\beta_1,\beta_2,\cdots,\beta_n$ 线性表示.

第 3 章　线性方程组解的结构

第 1 章的 1.2.4 节讨论了 n 元 n 个方程构成的线性方程组的解，我们知道由克莱姆法则 (定理 1.2.10) 可知线性方程组当系数行列式非零时有唯一解的结论，并给出了解的表达式. 但该法则具有很大的局限性：首先，若未知量的个数 n 与方程组中方程个数 m 不同时，该法则就不能适用；其次，n 元 n 个方程构成的方程组的系数行列式非零时，当 n 较大时，用行列式求解的计算量非常大，此时该法则很不实用. 我们在这一章将用求矩阵的秩时用过的初等变换法讨论由 n 元 m 个方程构成的更一般的方程组是否有解、解的结构和求解方法.

3.1　高斯消元法与矩阵的行变换

一般的线性方程组表示为

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ \cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m. \end{cases} \tag{3.1}$$

由第 2 章矩阵和向量的知识，线性方程组 (3.1) 也可表示成

$$x_1\boldsymbol{\alpha}_1 + x_2\boldsymbol{\alpha}_2 + \cdots + x_n\boldsymbol{\alpha}_n = \boldsymbol{b}, \tag{3.2}$$

其中向量

$$\boldsymbol{\alpha}_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}, \quad \boldsymbol{\alpha}_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix}, \quad \cdots, \quad \boldsymbol{\alpha}_n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}, \quad \boldsymbol{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}, \tag{3.3}$$

求解方程组 (3.1) 即为求解式 (3.2)，也就是求 $\boldsymbol{b}$ 表示为 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n$ 的线性组合.

线性方程组 (3.1) 还可表示成矩阵的形式

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \tag{3.4}$$

其中

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \quad \boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad \boldsymbol{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

称 $\boldsymbol{A}$ 为方程组 (3.1) 的**系数矩阵**, $\boldsymbol{x}$ 为**未知向量**, $\boldsymbol{b}$ 为**右端向量**. 使方程组 (3.4) 成立的已知向量称为该方程组的**解向量**. 将系数矩阵与右端向量合在一起构成的矩阵 $\boldsymbol{B}=(\boldsymbol{A}\quad \boldsymbol{b})$, 称为该方程组的**增广矩阵**，也可用增广矩阵来表示线性方程组.

现考虑方程组 (3.1) 的解. 若用一个由 n 个数组成的有序数组 $(s_1,s_2,\cdots,s_n)$ 代替变量 $x_1,x_2,\cdots,x_n$ 后，方程组 (3.1) 的所有方程都成立，则称 $x_1=s_1,x_2=s_2,\cdots,x_n=s_n$ 为方程组 (3.1) 的一个**解**. 该解写成向量形式 $(s_1,s_2,\cdots,s_n)$ 或 $(s_1,s_2,\cdots,s_n)^{\mathrm{T}}$, 就称为方程组 (3.1) 的**解向量**. 方程组 (3.1) 的所有解或所有解向量所成的集合称为方程组 (3.1) 的**解集**. 为了求出方程组的解集，我们要想办法让方程组的解集保持不变，而让方程组越来越简单，直至能容易地得出它们的解集. 为此，我们引入同解方程组的概念.

定义 3.1.1 (同解方程组)　具有相同解集的两个方程组称为**同解方程组**.

我们不仅要让方程组保持解集不变，还要让方程组向简单的方向进行变化. 我们知道：交换方程组中的两个方程；某个方程乘以一个非零的数；一个方程减去另一个方程，这些方程组的变换依然保持方程组的解集. 显然方程组的这些变换等价于对增广矩阵 $\boldsymbol{B}$ 施行相应的行初等变换. 当通过一系列的行初等变换将原增广矩阵 $\boldsymbol{B}$ 变换成行简化梯形矩阵时，方程组的解也就得到了, 这就是解线性方程组的**高斯 (Gauss) 消元法**.

例 3.1.1　解方程组

$$\begin{cases} 2x_1+x_2+3x_3=-5,\\ 3x_1+x_2+2x_3=-1,\\ 4x_1+3x_2+8x_3=-14. \end{cases}$$

解　此方程组的系数矩阵、未知向量、右端向量分别为

$$\boldsymbol{A}=\begin{pmatrix} 2&1&3\\ 3&1&2\\ 4&3&8 \end{pmatrix},\quad \boldsymbol{x}=\begin{pmatrix} x_1\\ x_2\\ x_3 \end{pmatrix},\quad \boldsymbol{b}=\begin{pmatrix} -5\\ -1\\ -14 \end{pmatrix}.$$

对其增广矩阵作行初等变换：

$$\boldsymbol{B}=\begin{pmatrix} 2&1&3&-5\\ 3&1&2&-1\\ 4&3&8&-14 \end{pmatrix} \xrightarrow{r_2+(-1)r_1} \begin{pmatrix} 2&1&3&-5\\ 1&0&-1&4\\ 4&3&8&-14 \end{pmatrix}$$

$$\xrightarrow{r_3+(-2)r_1} \begin{pmatrix} 2&1&3&-5\\ 1&0&-1&4\\ 0&1&2&-4 \end{pmatrix} \xrightarrow{r_1\leftrightarrow r_2} \begin{pmatrix} 1&0&-1&4\\ 2&1&3&-5\\ 0&1&2&-4 \end{pmatrix}$$

$$\xrightarrow{r_2+(-2)r_1} \begin{pmatrix} 1&0&-1&4\\ 0&1&5&-13\\ 0&1&2&-4 \end{pmatrix} \xrightarrow{r_3+(-1)r_2} \begin{pmatrix} 1&0&-1&4\\ 0&1&5&-13\\ 0&0&-3&9 \end{pmatrix}$$

$$\xrightarrow{-\frac{1}{3}r_3}\begin{pmatrix}1&0&-1&4\\0&1&5&-13\\0&0&1&-3\end{pmatrix}\xrightarrow{r_1+r_3}\begin{pmatrix}1&0&0&1\\0&1&5&-13\\0&0&1&-3\end{pmatrix}$$

$$\xrightarrow{r_2+(-5)r_3}\begin{pmatrix}1&0&0&1\\0&1&0&2\\0&0&1&-3\end{pmatrix},$$

最后得到方程组

$$\begin{cases}x_1=1,\\x_2=2,\\x_3=-3.\end{cases}$$

即得方程组的解为 $x_1=1,x_2=2,x_3=-3$. □

在用行初等变换将增广矩阵简化为行简化梯形矩阵时，既可像上述例子的行初等变换那样，逐列简化，也可先化为行梯形矩阵，再化为行简化梯形矩阵，见下例.

例 3.1.2　解方程组

$$\begin{cases}x_1+x_2-x_3+x_4=2,\\x_1-x_2+x_3+x_4=2,\\x_1-x_2-x_3-x_4=0.\end{cases}$$

解　对此方程组的增广矩阵作行初等变换:

$$\boldsymbol{B}=\begin{pmatrix}1&1&-1&1&2\\1&-1&1&1&2\\1&-1&-1&-1&0\end{pmatrix}\xrightarrow[r_3+(-1)r_1]{r_2+(-1)r_1}\begin{pmatrix}1&1&-1&1&2\\0&-2&2&0&0\\0&-2&0&-2&-2\end{pmatrix}$$

$$\xrightarrow{r_3+(-1)r_2}\begin{pmatrix}1&1&-1&1&2\\0&-2&2&0&0\\0&0&-2&-2&-2\end{pmatrix}\xrightarrow[(-\frac{1}{2})r_3]{(-\frac{1}{2})r_2}\begin{pmatrix}1&1&-1&1&2\\0&1&-1&0&0\\0&0&1&1&1\end{pmatrix}$$

$$\xrightarrow{r_1+(-1)r_2}\begin{pmatrix}1&0&0&1&2\\0&1&-1&0&0\\0&0&1&1&1\end{pmatrix}\xrightarrow{r_2+r_3}\begin{pmatrix}1&0&0&1&2\\0&1&0&1&1\\0&0&1&1&1\end{pmatrix},$$

最后得到方程组

$$\begin{cases}x_1+x_4=2,\\x_2+x_4=1,\\x_3+x_4=1,\end{cases}$$

即

$$\begin{cases}x_1=2-x_4,\\x_2=1-x_4,\\x_3=1-x_4.\end{cases}$$

令 $x_4=t$ 为任意实数，则方程组有无穷多组解，该方程组的解为 $x_1=2-t,x_2=1-t,x_3=1-t,x_4=t(t\in\mathbf{R})$. □

例 3.1.3 解方程组

$$\begin{cases}2x_1+5x_2-8x_3=8,\\4x_1+3x_2-9x_3=9,\\2x_1+3x_2-5x_3=7,\\x_1+8x_2-7x_3=10.\end{cases}$$

解 对此方程组的增广矩阵作行初等变换：

$$\boldsymbol{B}=\begin{pmatrix}2&5&-8&8\\4&3&-9&9\\2&3&-5&7\\1&8&-7&10\end{pmatrix}\xrightarrow{r_1+(-1)r_4}\begin{pmatrix}1&-3&-1&-2\\4&3&-9&9\\2&3&-5&7\\1&8&-7&10\end{pmatrix}$$

$$\xrightarrow{r_2+(-2)r_3}\begin{pmatrix}1&-3&-1&-2\\0&-3&1&-5\\2&3&-5&7\\1&8&-7&10\end{pmatrix}\xrightarrow[r_4+(-1)r_1]{r_3+(-2)r_1}\begin{pmatrix}1&-3&-1&-2\\0&-3&1&-5\\0&9&-3&11\\0&11&-6&12\end{pmatrix}$$

$$\xrightarrow[r_4+4r_2]{r_3+3r_2}\begin{pmatrix}1&-3&-1&-2\\0&-3&1&-5\\0&0&0&-4\\0&-1&-2&-8\end{pmatrix}\xrightarrow{r_2+(-3)r_4}\begin{pmatrix}1&-3&-1&-2\\0&0&7&19\\0&0&0&-4\\0&-1&-2&-8\end{pmatrix}$$

$$\xrightarrow[r_2\leftrightarrow r_3]{r_3\leftrightarrow r_4}\begin{pmatrix}1&-3&-1&-2\\0&-1&-2&-8\\0&0&7&19\\0&0&0&-4\end{pmatrix},$$

最后得到方程组为

$$\begin{cases}x_1-3x_2-x_3=-2,\\-x_2-2x_3=-8,\\7x_3=19,\\0=-4.\end{cases}$$

此为矛盾方程组，故该方程组无解. □

3.2* 高斯消元法的矩阵表示

解线性方程组的高斯消元法也可以只对增广矩阵的行初等变换过程进行到行梯形矩阵，然后直接对相应方程组用代入方法求得方程组的解，见下述例子.

例 3.2.1　解下列方程组

$$\begin{cases} 2x_1+x_2+3x_3=3, \\ 2x_1+3x_2+2x_3=-2, \\ 4x_1+6x_2+8x_3=0. \end{cases}$$

解　对此方程组的增广矩阵作行初等变换:

$$\boldsymbol{B}=\begin{pmatrix} 2 & 1 & 3 & 3 \\ 2 & 3 & 2 & -2 \\ 4 & 6 & 8 & 0 \end{pmatrix} \xrightarrow[r_3+(-2)r_1]{r_2+(-1)r_1} \begin{pmatrix} 2 & 1 & 3 & 3 \\ 0 & 2 & -1 & -5 \\ 0 & 4 & 2 & -6 \end{pmatrix} \xrightarrow{r_3+(-2)r_2} \begin{pmatrix} 2 & 1 & 3 & 3 \\ 0 & 2 & -1 & -5 \\ 0 & 0 & 4 & 4 \end{pmatrix},$$

最后得到方程组为

$$\begin{cases} 2x_1+x_2+3x_3=3, \\ 2x_2-x_3=-5, \\ 4x_3=4. \end{cases}$$

用代入方法容易解得此方程组的解为 $x_1=1, x_2=-2, x_3=1$. □

线性方程组也可以用另外一种方法求解: 将系数矩阵分解为一个下三角矩阵和一个上三角矩阵的乘积, 然后分两步求得方程组的解.

例 3.2.2　解下列方程组

$$\begin{cases} 2x_1+x_2+3x_3=3, \\ 2x_1+3x_2+2x_3=-2, \\ 4x_1+6x_2+8x_3=0. \end{cases}$$

解　考虑系数矩阵

$$\begin{aligned} \boldsymbol{A}&=\begin{pmatrix} 2 & 1 & 3 \\ 2 & 3 & 2 \\ 4 & 6 & 8 \end{pmatrix} \\ &=\begin{pmatrix} 1 & & \\ * & 1 & \\ * & * & 1 \end{pmatrix}\times\begin{pmatrix} * & * & * \\ & * & * \\ & & * \end{pmatrix}=\begin{pmatrix} 1 & & \\ * & 1 & \\ * & * & 1 \end{pmatrix}\times\begin{pmatrix} 2 & 1 & 3 \\ & * & * \\ & & * \end{pmatrix} \\ &=\begin{pmatrix} 1 & & \\ 1 & 1 & \\ 2 & * & 1 \end{pmatrix}\times\begin{pmatrix} 2 & 1 & 3 \\ & * & * \\ & & * \end{pmatrix}=\begin{pmatrix} 1 & & \\ 1 & 1 & \\ 2 & * & 1 \end{pmatrix}\times\begin{pmatrix} 2 & 1 & 3 \\ & 2 & -1 \\ & & * \end{pmatrix} \\ &=\begin{pmatrix} 1 & & \\ 1 & 1 & \\ 2 & 2 & 1 \end{pmatrix}\times\begin{pmatrix} 2 & 1 & 3 \\ & 2 & -1 \\ & & * \end{pmatrix}=\begin{pmatrix} 1 & & \\ 1 & 1 & \\ 2 & 2 & 1 \end{pmatrix}\times\begin{pmatrix} 2 & 1 & 3 \\ & 2 & -1 \\ & & 4 \end{pmatrix}. \end{aligned}$$

所以有

$$\boldsymbol{A}=\begin{pmatrix} 2 & 1 & 3 \\ 2 & 3 & 2 \\ 4 & 6 & 8 \end{pmatrix}=\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 2 & 2 & 1 \end{pmatrix}\times\begin{pmatrix} 2 & 1 & 3 \\ 0 & 2 & -1 \\ 0 & 0 & 4 \end{pmatrix}=\boldsymbol{LU}.$$

下面分两步用代入法求解该方程组，第一步求解方程组：$\boldsymbol{Ly}=\boldsymbol{b}$，即

$$\begin{cases} y_1 & = 3, \\ y_1 + y_2 & = -2, \\ 2y_1 + 2y_2 + y_3 & = 0, \end{cases}$$

易得解为：$y_1=3, y_2=-5, y_3=4$.

第二步求解方程组：$\boldsymbol{Ux}=\boldsymbol{y}$，即

$$\begin{cases} 2x_1 + x_2 + 3x_3 = 3, \\ \qquad 2x_2 - x_3 = -5, \\ \qquad\qquad 4x_3 = 4, \end{cases}$$

易得解为：$x_1=1, x_2=-2, x_3=1$. 此即为原方程组的解. □

上述例子说明了对于方程组 $\boldsymbol{Ax}=\boldsymbol{b}$, 若有 $\boldsymbol{A}=\boldsymbol{LU}$, 其中 $\boldsymbol{L}$ 和 $\boldsymbol{U}$ 分别是下三角矩阵和上三角矩阵，则令 $\boldsymbol{Ux}=\boldsymbol{y}$，原方程组的求解可通过先解方程组

$$\boldsymbol{Ly}=\boldsymbol{b},$$

再解方程组

$$\boldsymbol{Ux}=\boldsymbol{y},$$

从而得到方程组的解.

上述两种解方程组的方法，看似两种完全不同的方法，但当我们对于第一种解法过程中的行初等变换用初等矩阵来表示时，这两种方法其实就是一种方法. 具体说明如下.

设方程组为 $\boldsymbol{Ax}=\boldsymbol{b}$, 其中

$$\boldsymbol{A}=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}, \quad \boldsymbol{b}=\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix},$$

且 $\boldsymbol{A}$ 可逆.

对增广矩阵 $\boldsymbol{B}=(\boldsymbol{A}\ \ \boldsymbol{b})$ 进行行初等变换化为行梯形矩阵的过程，可用左乘等价初等矩阵表示. 为说明方便起见，假设简化过程中不需要进行任何行交换 (保证每次的主对角元素非零). 下面描述左乘初等矩阵化行梯形矩阵的过程.

第一步，将第一列主对角线下面的元素消为零. 左乘初等矩阵如下：

$$\boldsymbol{B}_2=\boldsymbol{E}\left(n,1\left(-\frac{a_{n1}}{a_{11}}\right)\right)\boldsymbol{E}\left(n-1,1\left(-\frac{a_{n-1,1}}{a_{11}}\right)\right)\cdots\boldsymbol{E}\left(2,1\left(-\frac{a_{21}}{a_{11}}\right)\right)\boldsymbol{B}.$$

令 $\boldsymbol{P}_1=\boldsymbol{E}\left(n,1\left(-\frac{a_{n1}}{a_{11}}\right)\right)\boldsymbol{E}\left(n-1,1\left(-\frac{a_{n-1,1}}{a_{11}}\right)\right)\cdots\boldsymbol{E}\left(2,1\left(-\frac{a_{21}}{a_{11}}\right)\right)$, 则有

$$\boldsymbol{P}_1=\begin{pmatrix} 1 & 0 & \cdots & 0 \\ -\dfrac{a_{21}}{a_{11}} & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ -\dfrac{a_{n1}}{a_{11}} & 0 & \cdots & 1 \end{pmatrix}=\begin{pmatrix} 1 & 0 & \cdots & 0 \\ * & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ * & 0 & \cdots & 1 \end{pmatrix}.$$

第二步，将第二列主对角元素下的元素消为零. 左乘初等矩阵如下

$$\boldsymbol{B}_3=\boldsymbol{E}\left(n,2\left(-\frac{a'_{n2}}{a'_{22}}\right)\right)\boldsymbol{E}\left(n-1,2\left(-\frac{a'_{n-1,2}}{a'_{22}}\right)\right)\cdots\boldsymbol{E}\left(3,2\left(-\frac{a'_{32}}{a'_{22}}\right)\right)\boldsymbol{B}_2.$$

令 $\boldsymbol{P}_2=\boldsymbol{E}\left(n,2\left(-\frac{a'_{n2}}{a'_{22}}\right)\right)\boldsymbol{E}\left(n-1,2\left(-\frac{a'_{n-1,2}}{a'_{22}}\right)\right)\cdots\boldsymbol{E}\left(3,2\left(-\frac{a'_{32}}{a'_{22}}\right)\right)$, 则有

$$\boldsymbol{P}_2=\begin{pmatrix}1&0&0&\cdots&0\\0&1&0&\cdots&0\\0&*&1&\cdots&0\\\vdots&\vdots&\vdots&&\vdots\\0&*&0&\cdots&1\end{pmatrix}.$$

依次执行左乘初等矩阵将各列主对角元素下面的元素消去为零. 最后一步对矩阵 $\boldsymbol{B}_{n-1}$ 消去第 $n-1$ 列主对角元素下面的元素，即左乘初等矩阵如下

$$\boldsymbol{B}_n=\boldsymbol{E}\left(n,n-1\left(-\frac{a''_{n,n-1}}{a''_{n-1,n-1}}\right)\right)\boldsymbol{B}_{n-1}=\boldsymbol{P}_{n-1}\boldsymbol{B}_{n-1}=\begin{pmatrix}1&\cdots&0&0\\\vdots&&\vdots&\vdots\\0&\cdots&1&0\\0&\cdots&*&1\end{pmatrix}\boldsymbol{B}_{n-1}.$$

将所有步骤合起来，就有

$$\boldsymbol{P}_{n-1}\boldsymbol{P}_{n-2}\cdots\boldsymbol{P}_2\boldsymbol{P}_1\boldsymbol{B}=\boldsymbol{P}(\boldsymbol{A}\ \ \boldsymbol{b})=(\boldsymbol{U}\ \ \boldsymbol{b}')=\left(\begin{array}{cccc|c}*&*&\cdots&*&*\\0&*&\cdots&*&*\\\vdots&\vdots&&\vdots&\vdots\\0&0&\cdots&*&*\end{array}\right),$$

于是 $\boldsymbol{PA}=\boldsymbol{U}$, 故

$$\boldsymbol{A}=\boldsymbol{P}^{-1}\boldsymbol{U}=(\boldsymbol{P}_1^{-1}\boldsymbol{P}_2^{-1}\cdots\boldsymbol{P}_{n-1}^{-1})\boldsymbol{U}=\boldsymbol{LU},$$

其中

$$\boldsymbol{L}=\begin{pmatrix}1&0&\cdots&0\\*&1&\cdots&0\\\vdots&\vdots&&\vdots\\*&0&\cdots&1\end{pmatrix}\begin{pmatrix}1&0&\cdots&0\\0&1&\cdots&0\\\vdots&\vdots&&\vdots\\0&*&\cdots&1\end{pmatrix}\cdots\begin{pmatrix}1&\cdots&0&0\\\vdots&&\vdots&\vdots\\0&\cdots&1&0\\0&\cdots&*&1\end{pmatrix}=\begin{pmatrix}1&0&\cdots&0\\*&1&\cdots&0\\\vdots&\vdots&&\vdots\\*&*&\cdots&1\end{pmatrix},$$

$$\boldsymbol{A}=\boldsymbol{LU}=\begin{pmatrix}1&0&\cdots&0\\*&1&\cdots&0\\\vdots&\vdots&&\vdots\\*&*&\cdots&1\end{pmatrix}\times\begin{pmatrix}*&*&\cdots&*\\0&*&\cdots&*\\\vdots&\vdots&&\vdots\\0&0&\cdots&*\end{pmatrix}.$$

从而有 $\boldsymbol{Ax}=\boldsymbol{b}$ 即为 $\boldsymbol{Ux}=\boldsymbol{b}'=\boldsymbol{L}^{-1}\boldsymbol{b}$.

设 $\boldsymbol{Ux}=\boldsymbol{y}$, 则 $\boldsymbol{y}=\boldsymbol{L}^{-1}\boldsymbol{b}$, 即 $\boldsymbol{Ly}=\boldsymbol{b}$. 此即为先求解 $\boldsymbol{Ly}=\boldsymbol{b}$ 再求解 $\boldsymbol{Ux}=\boldsymbol{y}$ 的方法.

3.3 线性方程组的可解性

本节讨论线性方程组是否有解，是否有无穷多解的判定条件，归结为下面的定理.

定理1.

定理 3.3.1 线性方程组 (3.1) 有解的充要条件是系数矩阵的秩等于增广矩阵的秩. 且当 $\mathrm{r}(\boldsymbol{A})=\mathrm{r}(\boldsymbol{B})=n$ 时，方程组有唯一解；而当 $\mathrm{r}(\boldsymbol{A})=\mathrm{r}(\boldsymbol{B})<n$ 时，方程组有无穷多组解.

证明 线性方程组 (3.1) 的系数矩阵 $\boldsymbol{A}$ 和增广矩阵 $\boldsymbol{B}$ 为

$$\boldsymbol{A}=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix},\quad \boldsymbol{B}=(\boldsymbol{A}\ \boldsymbol{b})=\left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right).$$

经过适当的系数矩阵部分的列交换, 并进行一系列高斯消元法的行初等变换，最后可以化为行梯形矩阵： 必要时进行列交换

$$\left(\begin{array}{ccccccc|c} c_{11} & c_{12} & \cdots & c_{1r} & c_{1,r+1} & \cdots & c_{1n} & d_1 \\ 0 & c_{22} & \cdots & c_{2r} & c_{2,r+1} & \cdots & c_{2n} & d_2 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & c_{rr} & c_{r,r+1} & \cdots & c_{rn} & d_r \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & d_{r+1} \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \end{array}\right)=(\boldsymbol{C}\ \boldsymbol{d})=\widetilde{\boldsymbol{C}},$$

其中 $c_{11},c_{22},\cdots,c_{rr}$ 均不等于 0.

为便于描述，我们假设整个简化过程未进行任何列交换. 由于 $\boldsymbol{A}$ 经过初等变换化为 $\boldsymbol{C}$，$\boldsymbol{B}$ 经过初等变换化为 $\widetilde{\boldsymbol{C}}$，故有 $\mathrm{r}(\boldsymbol{A})=\mathrm{r}(\boldsymbol{C})=r$，$\mathrm{r}(\boldsymbol{B})=\mathrm{r}(\widetilde{\boldsymbol{C}})$. 当 $d_{r+1}\neq 0$ 时，$\mathrm{r}(\boldsymbol{B})=\mathrm{r}(\widetilde{\boldsymbol{C}})=r+1$；当 $d_{r+1}=0$ 时，$\mathrm{r}(\boldsymbol{B})=\mathrm{r}(\widetilde{\boldsymbol{C}})=r$.

当 $d_{r+1}\neq 0$ 时，即 $\mathrm{r}(\boldsymbol{B})>\mathrm{r}(\boldsymbol{A})$，原方程组经过高斯消元过程，化为同解方程组

$$\begin{cases} c_{11}x_1+c_{12}x_2+\cdots+c_{1r}x_r+\cdots+c_{1n}x_n=d_1, \\ \qquad\quad c_{22}x_2+\cdots+c_{2r}x_r+\cdots+c_{2n}x_n=d_2, \\ \qquad\qquad\qquad\qquad\qquad\qquad\cdots\cdots \quad (\#) \\ \qquad\qquad\qquad\qquad c_{rr}x_r+\cdots+c_{rn}x_n=d_r, \\ \qquad\qquad\qquad\qquad\qquad\qquad 0=d_{r+1}\neq 0. \end{cases}$$

最后一个方程显然矛盾，故原方程组无解.

当 $d_{r+1}=0$ 时，即 $\mathrm{r}(\boldsymbol{B})=\mathrm{r}(\boldsymbol{A})$，则原方程组化为

$$\begin{cases} c_{11}x_1+c_{12}x_2+\cdots+c_{1r}x_r+\cdots+c_{1n}x_n=d_1, \\ \qquad c_{22}x_2+\cdots+c_{2r}x_r+\cdots+c_{2n}x_n=d_2, \\ \qquad\qquad \cdots\cdots \\ \qquad\qquad c_{rr}x_r+\cdots+c_{rn}x_n=d_r. \end{cases} \tag{3.5}$$

令 $x_{r+1}=0,\cdots,x_n=0$, 解得

$$\begin{cases} x_r=d_r/c_{rr}, \\ x_{r-1}=(d_{r-1}-c_{r-1,r}x_r)/c_{r-1,r-1}, \\ \cdots\cdots \\ x_1=(d_1-c_{12}x_2-c_{13}x_3-\cdots-c_{1r}x_r)/c_{11}, \end{cases}$$

则 $(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$ 显然是方程组 (3.5) 的解向量，也是原方程组 (3.1) 的一个解.

若 $\mathrm{r}(\boldsymbol{A})=\mathrm{r}(\boldsymbol{B})=r$，当 $r=n$ 时，方程组 (3.5) 再化为如下形式：

$$\begin{cases} x_n=d_n/c_{nn}, \\ x_{n-1}=(d_{n-1}-c_{nn}x_n)/c_{n-1,n-1}, \\ \cdots\cdots \\ x_2=(d_2-c_{23}x_3-\cdots-c_{2n}x_n)/c_{22}, \\ x_1=(d_1-c_{12}x_2-\cdots-c_{1n}x_n)/c_{11}, \end{cases}$$

则易知方程组 (3.5) 的解唯一确定，从而同解方程组 (3.1) 有唯一解.

若 $\mathrm{r}(\boldsymbol{A})=\mathrm{r}(\boldsymbol{B})=r<n$ 时，易知对任一 $t\in\mathbf{R}$, 令 $x_{r+1}=0,\cdots,x_{n-1}=0,x_n=t$，解得

$$\begin{cases} x_r=d_r/c_{rr}-c_{rn}t/c_{rr}, \\ x_{r-1}=(d_{r-1}-c_{r-1,r}x_r)/c_{r-1,r-1}-c_{r-1,n}t/c_{r-1,r-1}, \\ \cdots\cdots \\ x_1=(d_1-c_{12}x_2-c_{13}x_3-\cdots-c_{1r}x_r)/c_{11}-c_{1n}t/c_{11}, \end{cases}$$

则 $(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$ 满足方程组 (3.5), 于是方程组 (3.5) 有无穷多解，从而原方程组 (3.1) 有无穷多解. □

若将方程组 (3.1) 的系数矩阵 $\boldsymbol{A}$ 按列分块为 $\boldsymbol{A}=(\boldsymbol{\alpha}_1\ \cdots\ \boldsymbol{\alpha}_n)$，则方程组 (3.1) 与方程组 (3.2)、方程组 (3.4) 等价. 下面我们从向量相关性的角度出发来看上述定理的意义.

现在考虑方程组 (3.2)，即

$$x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+\cdots+x_n\boldsymbol{\alpha}_n=\boldsymbol{b}$$

的有解性.

方程组 (3.2) 有解表示 $\boldsymbol{b}$ 可由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 线性表示，从而两个向量组 $\{\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n\}$ 和 $\{\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n,\boldsymbol{b}\}$ 等价，这与两向量组有相同的秩

$$\mathrm{r}(\boldsymbol{A})=\mathrm{r}\{\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n\}=\mathrm{r}\{\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n,\boldsymbol{b}\}=\mathrm{r}(\boldsymbol{B})$$

是一回事.

另外，考虑 $\mathrm{r}(\boldsymbol{A})=\mathrm{r}(\boldsymbol{B})$ 时 $\boldsymbol{b}$ 由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 线性表示的形式.

当 $\mathrm{r}(\boldsymbol{A})=\mathrm{r}(\boldsymbol{B})=n$ 时，向量组 $\{\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n\}$ 线性无关，故表示形式唯一，即方程组 (3.2) 有唯一解.

当 $\mathrm{r}(\boldsymbol{A})=\mathrm{r}(\boldsymbol{B})<n$ 时，向量组 $\{\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n\}$ 线性相关，故表示形式不唯一，有无穷多种表示形式，即方程组 (3.2) 有无穷多解.

对于前面解方程组的例 3.1.1、例 3.1.2、例 3.1.3，用系数矩阵和增广矩阵秩的关系可以很容易判别解的存在性和唯一性. 下面我们再来看一个方程组有解性判别的例子.

例 3.3.1　解方程组

$$\begin{cases}3x_1+\lambda x_2+x_3=4,\\ x_1+2x_2-4x_3=\mu,\\ x_1-x_2+9x_3=19.\end{cases}$$

解

$$\boldsymbol{B}=\begin{pmatrix}3&\lambda&1&4\\1&2&-4&\mu\\1&-1&9&19\end{pmatrix}\xrightarrow{r_1\leftrightarrow r_3}\begin{pmatrix}1&-1&9&19\\1&2&-4&\mu\\3&\lambda&1&4\end{pmatrix}$$

$$\xrightarrow[r_3+(-3)r_1]{r_2+(-1)r_1}\begin{pmatrix}1&-1&9&19\\0&3&-13&\mu-19\\0&\lambda+3&-26&-53\end{pmatrix}\xrightarrow{r_3+(-2)r_2}\begin{pmatrix}1&-1&9&19\\0&3&-13&\mu-19\\0&\lambda-3&0&-2\mu-15\end{pmatrix}=\boldsymbol{B}_1.$$

当 $\lambda\neq 3$ 时，$\mathrm{r}(\boldsymbol{A})=\mathrm{r}(\boldsymbol{B})=3$，方程组有唯一解. 进一步化简 $\boldsymbol{B}$,

$$\boldsymbol{B}_1\xrightarrow[(\frac{1}{\lambda-3})r_3]{\substack{r_1+\frac{9}{13}r_2\\(-\frac{1}{13})r_2}}\begin{pmatrix}1&\dfrac{14}{13}&0&19+\dfrac{9}{13}(\mu-19)\\0&-\dfrac{3}{13}&1&-\dfrac{1}{13}(\mu-19)\\0&1&0&-\dfrac{2\mu+15}{\lambda-3}\end{pmatrix}$$

$$\xrightarrow[r_2\leftrightarrow r_3]{\substack{r_1-\frac{14}{13}r_3\\r_2+\frac{3}{13}r_3}}\begin{pmatrix}1&0&0&\dfrac{9\mu+76}{13}+\dfrac{14}{13}\times\dfrac{2\mu+15}{\lambda-3}\\0&1&0&-\dfrac{2\mu+15}{\lambda-3}\\0&0&1&-\dfrac{\mu-19}{13}-\dfrac{3}{13}\times\dfrac{2\mu+15}{\lambda-3}\end{pmatrix},$$

得方程组的唯一解为

$$x_1=\frac{9\mu+76}{13}+\frac{14}{13}\times\frac{2\mu+15}{\lambda-3},x_2=-\frac{2\mu+15}{\lambda-3},x_3=-\frac{\mu-19}{13}-\frac{3}{13}\times\frac{2\mu+15}{\lambda-3}.$$

当 $\lambda=3,\mu=-\dfrac{15}{2}$ 时，$\mathrm{r}(\boldsymbol{A})=\mathrm{r}(\boldsymbol{B})=2$，方程组有无穷多解. 进一步化简 $\boldsymbol{B}$,

$$\boldsymbol{B}_1=\begin{pmatrix}1&-1&9&19\\0&3&-13&-\dfrac{53}{2}\\0&0&0&0\end{pmatrix}\xrightarrow[\frac{1}{3}r_2]{r_1+\frac{1}{3}r_2}\begin{pmatrix}1&0&\dfrac{14}{3}&\dfrac{61}{6}\\0&1&-\dfrac{13}{3}&-\dfrac{53}{6}\\0&0&0&0\end{pmatrix},$$

得方程组的解为：$x_1=\dfrac{61}{6}-14t, x_2=-\dfrac{53}{6}+13t, x_3=3t, t\in\mathbf{R}$.

当 $\lambda=3,\mu\neq-\dfrac{15}{2}$ 时，$2=\mathrm{r}(\boldsymbol{A})<\mathrm{r}(\boldsymbol{B})=3$，方程组无解. □

3.4　线性方程组解的性质与结构

本节我们讨论各种线性方程组的解集的形式和特点. 我们先考虑较简单的右端向量为 $\boldsymbol{\theta}$ 的所谓齐次线性方程组的解集，再讨论右端非零的方程组的解集.

3.4.1　齐次方程组解的结构

考虑方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{\theta}$，它的解有如下的特点.

定理2. **定理 3.4.1**　*若 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$ 是方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{\theta}$ 的解，则其线性组合 $k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2$ 也是该方程组的解.*

证明　直接验证：$\boldsymbol{A}(k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2)=k_1\boldsymbol{A}\boldsymbol{\alpha}_1+k_2\boldsymbol{A}\boldsymbol{\alpha}_2=k_1\boldsymbol{\theta}+k_2\boldsymbol{\theta}=\boldsymbol{\theta}$. □

定义1. **定义 3.4.1** (齐次线性方程组、基础解系)　右端为零的线性方程组称为**齐次线性方程组**；能线性表示出齐次方程组所有解的极大无关向量组称为该齐次线性方程组的**基础解系**.

定义 3.4.2 (方程组的特解、通解)　方程组的某一个解称为方程组的**特解**；方程组所有的解的集合称为方程组的**通解**.

对于齐次线性方程组，$\boldsymbol{x}=\boldsymbol{\theta}$ 显然是它的解，但齐次线性方程组是否有非零解？如果有，又如何求出所有可能的非零解？是我们特别要关心的问题. 下面我们就讨论齐次线性方程组有非零解的条件及所有非零解的表示.

推论1. **定理 3.4.2**　*设 $\boldsymbol{A}\in\mathbf{R}^{m\times n}$，若 $\mathrm{r}(\boldsymbol{A})=n$，则 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{\theta}$ 只有零解；若 $\mathrm{r}(\boldsymbol{A})<n$，则 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{\theta}$ 有非零解.*

证明　由定理3.3.1 知，当 $\mathrm{r}(\boldsymbol{A})=n$ 时，$\boldsymbol{A}\boldsymbol{x}=\boldsymbol{\theta}$ 只有唯一解 $\boldsymbol{x}=\boldsymbol{\theta}$，即零解. 当 $\mathrm{r}(\boldsymbol{A})<n$ 时，$\boldsymbol{A}\boldsymbol{x}=\boldsymbol{\theta}$ 有无穷多解，故除零解外还有非零解. □

例 3.4.1　判别下列齐次方程组

$$\begin{cases}x_1+x_2=0,\\2x_1+x_2+x_3=0,\\x_1+x_3=0,\\-x_1+2x_2+x_3=0\end{cases}$$

是否有非零解.

解　对系数矩阵作行初等变换

$$\begin{pmatrix}1&1&0\\2&1&1\\1&0&1\\-1&2&1\end{pmatrix}\xrightarrow[r_4+r_1]{\substack{r_2+(-2)r_1\\r_3+(-1)r_1}}\begin{pmatrix}1&1&0\\0&-1&1\\0&-1&1\\0&3&1\end{pmatrix}\xrightarrow{\substack{r_3+(-1)r_2\\r_4+3r_2}}\begin{pmatrix}1&1&0\\0&-1&1\\0&0&0\\0&0&4\end{pmatrix}\xrightarrow{r_3\leftrightarrow r_4}\begin{pmatrix}1&1&0\\0&-1&1\\0&0&4\\0&0&0\end{pmatrix},$$

其秩为 3，所以方程组只有零解 (列向量线性无关). □

推论 3.4.3 方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{\theta}, \boldsymbol{A}\in\mathbf{R}^{n\times n}$ 有非零解的充要条件是 $|\boldsymbol{A}|=0$.

证明 由定理 3.4.2，

$$\boldsymbol{A}\boldsymbol{x}=\boldsymbol{\theta}\text{有非零解} \iff \mathrm{r}(\boldsymbol{A})<n \iff |\boldsymbol{A}|=0. \qquad \square$$

推论 3.4.4 若 $\boldsymbol{A}\in\mathbf{R}^{m\times n}$，且 $m<n$，则 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{\theta}$ 有非零解.

证明 由定理 3.4.2 直接可得方程组有非零解. □

例 3.4.2 讨论含参的 3 元方程组

$$\begin{cases} x_1+x_2+\lambda x_3=0, \\ 2x_1+x_2+x_3=0, \\ x_1+x_3=0 \end{cases}$$

有非零解的条件.

解 计算系数行列式

$$\begin{vmatrix} 1 & 1 & \lambda \\ 2 & 1 & 1 \\ 1 & 0 & 1 \end{vmatrix} \overset{c_3+(-1)c_1}{=} \begin{vmatrix} 1 & 1 & \lambda-1 \\ 2 & 1 & -1 \\ 1 & 0 & 0 \end{vmatrix} = \begin{vmatrix} 1 & \lambda-1 \\ 1 & -1 \end{vmatrix} = -\lambda.$$

易知当 $\lambda=0$ 时方程组有非零解，当 $\lambda\neq0$ 时，方程组只有零解. □

当一个齐次线性方程组有非零解时，我们如何去求得方程组的包括零解和非零解的所有解，即方程组的通解. 先看下面的例子.

例 3.4.3 求解下列齐次方程组

$$\begin{cases} x_1-x_2+x_4=0, \\ x_1-x_2+x_3-x_4=0, \\ x_1-x_2-x_3+3x_4=0, \\ x_1-x_2+2x_3-3x_4=0. \end{cases}$$

解 将系数矩阵 $\boldsymbol{A}$ 化为行简化梯形矩阵

$$\boldsymbol{A}=\begin{pmatrix} 1 & -1 & 0 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 3 \\ 1 & -1 & 2 & -3 \end{pmatrix} \xrightarrow[r_4+(-1)r_1]{\substack{r_2+(-1)r_1 \\ r_3+(-1)r_1}} \begin{pmatrix} 1 & -1 & 0 & 1 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & -1 & 2 \\ 0 & 0 & 2 & -4 \end{pmatrix}$$

$$\xrightarrow{\substack{r_3+r_2 \\ r_4+(-2)r_2}} \begin{pmatrix} 1 & -1 & 0 & 1 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

于是我们得到简化的同解方程组

$$\begin{cases} x_1-x_2+x_4=0, \\ x_3-2x_4=0. \end{cases}$$

再改写上述方程组为 (将行简化梯形矩阵每行最左边的 1 所在列以外的列对应的变量移到方程组右边)

$$\begin{cases} x_1 = x_2 - x_4, \\ x_3 = 2x_4. \end{cases}$$

分别令 x_2, x_4 为 $x_2 = s, x_4 = t, s, t \in \mathbf{R}$，则可得方程组的解为

$$\begin{cases} x_1 = s - t, \\ x_2 = s, \\ x_3 = 2t, \\ x_4 = t, \end{cases} \qquad s, t \in \mathbf{R}.$$

□

上述例子中，简化的同解方程组中有些变量需要移到方程组的右边，并设为自由参数，而有些变量则保留在左边. 这些保留在左边的变量正好是行简化梯形矩阵每一行最左边的 1 对应的变量，称为**非自由变量**；那些移到右边的变量则是行简化梯形矩阵最左边的 1 以外的列对应的变量，可以任意取值，称为**自由变量**.

下面我们通过一个定理给出一般齐次线性方程组的通解形式.

定理3. 设 $Ax=0$ 的系数矩阵 A 经过行初等变换化为如下行简化梯形:

$$\begin{pmatrix} 1 & 0 & \cdots & 0 & c_{1,r+1} & \cdots & c_{1n} \\ 0 & 1 & \cdots & 0 & c_{2,r+1} & \cdots & c_{2n} \\ \vdots & & & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 & c_{r,r+1} & \cdots & c_{rn} \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix}, \text{则 } \alpha_1 = \begin{pmatrix} -c_{1,r+1} \\ -c_{2,r+1} \\ \vdots \\ -c_{r,r+1} \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad \alpha_2 = \begin{pmatrix} -c_{1,r+2} \\ -c_{2,r+2} \\ \vdots \\ -c_{r,r+2} \\ 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} \quad \cdots \quad \alpha_{n-r} = \begin{pmatrix} -c_{1n} \\ -c_{2n} \\ \vdots \\ -c_{rn} \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

是 $Ax=0$ 的一个基础解系. 其通解为 $k_1\alpha_1+k_2\alpha_2+\cdots+k_{n-r}\alpha_{n-r}$.
其中 $k_1,k_2,\cdots,k_{n-r}$ 为任意实数.

证: 该行简化梯形矩阵对应的线性方程组为:

$$\begin{cases} x_1 \qquad + c_{1,r+1}x_{r+1} + \cdots + c_{1n}x_n = 0 \\ \quad x_2 \quad + c_{2,r+1}x_{r+1} + \cdots + c_{2n}x_n = 0 \\ \qquad \cdots\cdots \\ \qquad x_r + c_{r,r+1}x_{r+1} + \cdots + c_{rn}x_n = 0 \end{cases} \qquad (\#\#)$$

分别令 $(x_{r+1},\cdots,x_n)^T = (1,0,\cdots,0)^T, (0,1,\cdots,0)^T, (0,0,\cdots,1)^T$.

易见 $\alpha_1,\alpha_2,\cdots\alpha_{n-r}$ 均是方程组(##)的解.

方程组(##)中的未知量 $x_{r+1},\cdots,x_n$ 称为自由未知量.

由于 $(\alpha_1,\alpha_2,\cdots\alpha_{n-r})$ 后 $n-r$ 行构成 E_{n-r}，故 $\alpha_1,\alpha_2,\cdots,\alpha_{n-r}$ 线性无关.

设 $\alpha=(s_1,s_2,\cdots s_r,s_{r+1},\cdots,s_n)^T$ 是方程组(##)的任意解.

由于 $\beta = s_{r+1}\alpha_1+s_{r+2}\alpha_2+\cdots+s_n\alpha_{n-r} = s_{r+1}\begin{pmatrix}\cdots\end{pmatrix} + s_{r+2}\begin{pmatrix}\cdots\end{pmatrix} + \cdots + s_n\begin{pmatrix}\cdots\end{pmatrix} = \begin{pmatrix} * \\ * \\ \vdots \\ * \\ s_{r+1} \\ s_{r+2} \\ \vdots \\ s_n \end{pmatrix}$

是方程组的一个解.

且其后 $n-r$ 个分量(自由变量)取值与 α 后 $n-r$ 个变量分别相同

故必有 $\alpha=\beta$. 即 α 可以由 $\alpha_1,\alpha_2,\cdots,\alpha_{n-r}$ 线性表示.

进一步，将自由变量移到方程组的右边，得如下方程组

$$\begin{cases} x_1 = -d_{12}x_2 - \cdots - d_{1i_2+1}x_{i_2+1} - \cdots - d_{1i_r+1}x_{i_r+1} - \cdots - d_{1n}x_n, \\ x_{i_2} = -d_{2i_2+1}x_{i_2+1} - \cdots - d_{2i_r+1}x_{i_r+1} - \cdots - d_{2n}x_n, \\ x_{i_3} = -d_{3i_3+1}x_{i_3+1} - \cdots - d_{3i_r+1}x_{i_r+1} - \cdots - d_{3n}x_n, \\ \cdots\cdots \\ x_{i_r} = -d_{ri_r+1}x_{i_r+1} - \cdots - d_{rn}x_n. \end{cases} \tag{3.7}$$

对 $n-r$ 个自由变量 $x_2,\cdots,x_{i_2-1},x_{i_2+1},\cdots,x_{i_r-1},x_{i_r+1},\cdots,x_n$ 分别取 $n-r$ 组数据 $(1,0,\cdots,0),(0,1,0,\cdots,0),\cdots,(0,\cdots,0,1)$ 代入方程组 (3.7)，则可得 $n-r$ 组非自由变量 $x_1,x_{i_2},\cdots,x_{i_r}$ 的值，从而构成 $n-r$ 组方程组的解，其解向量为：$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{n-r}$. 下面证明它们就是方程组的一个基础解系.

由

$$k_1\boldsymbol{\alpha}_1+\cdots+k_{n-r}\boldsymbol{\alpha}_{n-r}=\boldsymbol{\theta}$$

可知对应于 $n-r$ 个自由变量的分量分别是 $k_1,k_2,\cdots,k_{n-r}$，由上式这些分量都应该为零，故 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_{n-r}$ 线性无关.

再看方程组的任意一个解向量：$\boldsymbol{\beta}=(s_1,s_2,\cdots,s_n)^{\mathrm{T}}$，它也是方程组 (3.7) 的解. 令

$$\boldsymbol{\gamma}=k_1\boldsymbol{\alpha}_1+\cdots+k_{n-r}\boldsymbol{\alpha}_{n-r},$$

其中 $k_1,k_2,\cdots,k_{n-r}$ 分别取自由变量 $x_2,\cdots,x_{i_2-1},x_{i_2+1},\cdots,x_{i_r-1},x_{i_r+1},\cdots,x_n$ 对应的分量 $s_2,\cdots,s_{i_2-1},s_{i_2+1},\cdots,s_{i_r-1},s_{i_r+1},\cdots,s_n$，则 $\boldsymbol{\gamma}$ 是方程组的解向量，并且其自由变量的值分别为解向量 $\boldsymbol{\beta}$ 的值. 由方程组 (3.7) 可知方程组的解由自由变量的值唯一确定，故

$$\boldsymbol{\beta}=\boldsymbol{\gamma}=k_1\boldsymbol{\alpha}_1+\cdots+k_{n-r}\boldsymbol{\alpha}_{n-r},$$

即任意解向量都可表示为 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_{n-r}$ 的线性组合. 从而它们构成了方程组的一个基础解系. 方程组的通解形式是显然的. □

注 1　若系数矩阵简化后的行简化梯形矩阵为

$$\begin{pmatrix} 1 & 0 & \cdots & 0 & d_{1,r+1} & \cdots & d_{1n} \\ 0 & 1 & \cdots & 0 & d_{2,r+1} & \cdots & d_{2n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 & d_{r,r+1} & \cdots & d_{rn} \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix},$$

则

$$\boldsymbol{\alpha}_1=\begin{pmatrix}-d_{1,r+1}\\ \vdots\\ -d_{r,r+1}\\ 1\\ 0\\ \vdots\\ 0\end{pmatrix},\quad \boldsymbol{\alpha}_2=\begin{pmatrix}-d_{1,r+2}\\ \vdots\\ -d_{r,r+2}\\ 0\\ 1\\ \vdots\\ 0\end{pmatrix},\quad \cdots,\quad \boldsymbol{\alpha}_{n-r}=\begin{pmatrix}-d_{1n}\\ \vdots\\ -d_{rn}\\ 0\\ \vdots\\ 0\\ 1\end{pmatrix},$$

是齐次方程组的一个基础解系.

注 2　若齐次方程组有非零解，则该方程组的基础解系并不唯一.

事实上，若齐次方程组 $\boldsymbol{Ax}=\boldsymbol{\theta}$ 有基础解系 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_{n-r}$，则 $2\boldsymbol{\alpha}_1,\cdots,2\boldsymbol{\alpha}_{n-r}$ 和 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2+\cdots+\boldsymbol{\alpha}_{n-r}$ 均为原齐次方程组的基础解系.

由上述定理可知，若齐次方程组的系数矩阵 $\boldsymbol{A}$ 的秩为 r，则其基础解系的向量个数为 $n-r$. 故有如下推论.

推论2、**推论 3.4.6**　若 $\boldsymbol{A}\in\mathbf{R}^{m\times n}$，则 $\mathrm{r}(\boldsymbol{A})+\mathrm{r}(N(\boldsymbol{A}))=n$，其中 $N(\boldsymbol{A})$ 表示 $\boldsymbol{Ax}=\boldsymbol{\theta}$ 的基础解系为列构成的矩阵.

r(A)+基础解系中所含向量个数=未知量个数

例 3.4.4　求齐次方程组

$$\begin{cases}x_1+2x_2+x_3=0,\\ x_1+x_2=0,\\ 2x_1+x_2-x_3=0\end{cases}$$

的基础解系及通解.

解　对系数矩阵作初等行变换

$$\begin{pmatrix}1&2&1\\1&1&0\\2&1&-1\end{pmatrix}\xrightarrow[r_3+(-2)r_1]{r_2+(-1)r_1}\begin{pmatrix}1&2&1\\0&-1&-1\\0&-3&-3\end{pmatrix}$$

$$\xrightarrow[r_3+(-3)r_2]{r_1+2r_2}\begin{pmatrix}1&0&-1\\0&-1&-1\\0&0&0\end{pmatrix}\xrightarrow{(-1)r_2}\begin{pmatrix}1&0&-1\\0&1&1\\0&0&0\end{pmatrix},$$

得一个基础解系为

$$\boldsymbol{\alpha}=\begin{pmatrix}1\\-1\\1\end{pmatrix}.$$

于是所求通解为 $\boldsymbol{x}=k\boldsymbol{\alpha}$，$k\in\mathbf{R}$.　□

例 3.4.5 求齐次方程组

$$\begin{cases} x_1+2x_2-x_3+x_4=0, \\ 2x_1+2x_2+x_3=0, \\ x_1+4x_2-4x_3+3x_4=0, \\ 4x_1+6x_2-x_3+2x_4=0, \\ 2x_1-2x_2+7x_3-4x_4=0 \end{cases}$$

的基础解系.

解 对系数矩阵作初等行变换

$$\begin{pmatrix} 1 & 2 & -1 & 1 \\ 2 & 2 & 1 & 0 \\ 1 & 4 & -4 & 3 \\ 4 & 6 & -1 & 2 \\ 2 & -2 & 7 & -4 \end{pmatrix} \xrightarrow[\substack{r_4+(-4)r_1 \\ r_5+(-2)r_1}]{\substack{r_2+(-2)r_1 \\ r_3+(-1)r_1}} \begin{pmatrix} 1 & 2 & -1 & 1 \\ 0 & -2 & 3 & -2 \\ 0 & 2 & -3 & 2 \\ 0 & -2 & 3 & -2 \\ 0 & -6 & 9 & -6 \end{pmatrix}$$

$$\xrightarrow[\substack{r_4+(-1)r_2 \\ r_5+(-3)r_2 \\ (-\frac{1}{2})r_2}]{\substack{r_1+r_2 \\ r_3+r_2}} \begin{pmatrix} 1 & 0 & 2 & -1 \\ 0 & 1 & -3/2 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

容易求得一个基础解系为

$$\boldsymbol{\alpha}_1=\begin{pmatrix} -2 \\ 3/2 \\ 1 \\ 0 \end{pmatrix}, \quad \boldsymbol{\alpha}_2=\begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \end{pmatrix}.$$

□

例 3.4.6 求齐次方程组

$$\begin{cases} 2x_1-4x_2+2x_3+7x_4=0, \\ 3x_1-6x_2+4x_3+4x_4=0, \\ 4x_1-8x_2+4x_3+15x_4=0 \end{cases}$$

的基础解系.

解 对系数矩阵作初等行变换

$$\begin{pmatrix} 2 & -4 & 2 & 7 \\ 3 & -6 & 4 & 4 \\ 4 & -8 & 4 & 15 \end{pmatrix} \xrightarrow{\substack{(\frac{1}{2})r_1 \\ r_2+(-3)r_1 \\ r_3+(-4)r_1}} \begin{pmatrix} 1 & -2 & 1 & 3.5 \\ 0 & 0 & 1 & -6.5 \\ 0 & 0 & 0 & 1 \end{pmatrix} \xrightarrow{\substack{r_1+(-1)r_2 \\ r_1+(-10)r_3 \\ r_2+6.5r_3}} \begin{pmatrix} 1 & -2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

取自由变量 x_2 为 1，解得基础解系为 $\boldsymbol{\alpha}=(2,1,0,0)^{\mathrm{T}}$. □

[例7] 证明：若 $A \in R^{n\times n}$，则

$$r(A^*) = \begin{cases} n, & r(A)=n \\ 1, & r(A)=n-1 \\ 0, & r(A)\le n-2 \end{cases}$$

当 $r(A) < n-2$ 时，$A^*=0$，故 $r(A^*)=0$.

当 $r(A)=n$ 时，$|A^*|=|A|^{n-1}\neq 0$ $\therefore r(A^*)=n$.

当 $r(A)=n-1$ 时，记 $A^*=(\xi_1,\xi_2,\cdots,\xi_n)$.

一方面，由 A^* 定义知，$A^*\neq 0$. 故 $r(A^*)\geq 1$

另一方面，由 $AA^*=A(\xi_1,\xi_2,\cdots,\xi_n)=|A|E=0$ 可知 $A\xi_i=0$.

$\therefore \xi_1,\xi_2,\cdots\xi_n$ 均是 $Ax=0$ 的解.

由于 $r(A)=n-1$ $\therefore Ax=0$ 的基础解系含一个向量.

$\therefore r(A^*)=r\{\xi_1,\xi_2,\cdots,\xi_n\}\le 1$

$\therefore r(A^*)=1$

[例8] 若 $A_{m\times l}$, $B_{l\times n}$ 满足 $AB=0$，则 $r(A)\le l-r(B)$.

设 $B=(\xi_1,\xi_2,\cdots,\xi_n)$.

由于 $AB=0$，故 $\xi_1,\xi_2,\cdots,\xi_n$ 均是 $Ax=0$ 的解.

$\therefore r(B)=r\{\xi_1,\xi_2,\cdots,\xi_n\}\le l-r(A)$

亦即 $r(A)+r(B)\le l$ $r(A)\le l-r(B)$

$r(A)=l-r(B)$ 的另一种证法：$AB=0 \Rightarrow (AB)^T=B^TA^T=0. \Rightarrow B^Tx=0 \cdots$

例 9. 已知向量 $\alpha_1,\alpha_2,\alpha_3\in R^n$ 线性无关，α_1,α_2 是 $A^Tx=0$ 的基础解系，α_2,α_3 是 $B^Tx=0$ 的基础解系，其中 $A\in R^{n\times m}$, $B\in R^{n\times p}$，求 $r(A\ B)$.

解. 设 $A^Tx=0$ 的通解为 $k_1\alpha_1+k_2\alpha_2$，$B^Tx=0$ 的通解为 $\ell_1\alpha_2+\ell_2\alpha_3$，其中 k_1,k_2,ℓ_1,ℓ_2 为任意实数。则

$$\begin{pmatrix} A^T \\ B^T \end{pmatrix} x = 0$$

的解应是满足 $k_1\alpha_1+k_2\alpha_2=\ell_1\alpha_2+\ell_2\alpha_3$ 的所有线性组合，即

$$k_1\alpha_1+(k_2-\ell_1)\alpha_2-\ell_2\alpha_3=0.$$

由题设知，$k_1=0$, $\ell_2=0$ 且 $k_2=\ell_1$.

这表明该方程组的通解应是 $k\alpha_2$，k 是任意实数。于是，

$$r(A\ B)=r\begin{pmatrix} A^T \\ B^T \end{pmatrix}=n-1.$$

3.4.2 非齐次方程组解的结构

非齐次线性方程组解的结构

$$Ax=b,\ b\neq 0.$$

定义 2. 对于非齐次方程组 $Ax=b$，齐次方程组 $Ax=0$ 称为其**导出组**。

非齐次方程组 $Ax=b$ 的解与其导出组 $Ax=0$ 的解有何关系?

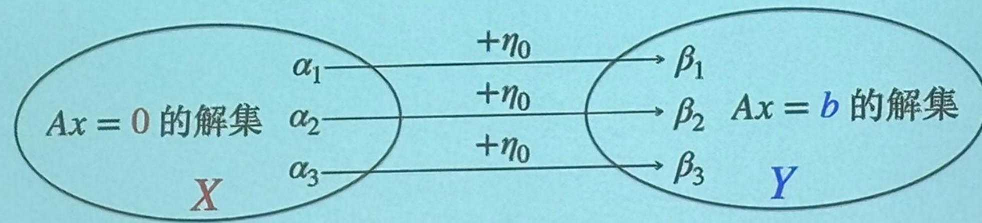
定理 4. 设 η_0 是 $Ax=b$ 的一个特解。则对 $Ax=0$ 的任一解 α，$\eta_0+\alpha$ 是 $Ax=b$ 的解；对 $Ax=b$ 的任一解 η，存在 $Ax=0$ 的解 α' 使得 $\eta=\eta_0+\alpha'$.

证明. 因为 $A(\eta_0+\alpha)=A\eta_0+A\alpha=b+0=b$，故 $\eta_0+\alpha$ 是 $Ax=b$ 的解。

另一方面，令 $\alpha'=\eta-\eta_0$，则有

$$A\alpha'=A(\eta-\eta_0)=A\eta-A\eta_0=b-b=0,$$

即 α' 是 $Ax=0$ 的解，且 $\eta=\eta_0+(\eta-\eta_0)=\eta_0+\alpha'$.



如何求非齐次方程组 $Ax=b$ 的通解?

1. 用行初等变换把 $Ax=b$ 的增广阵 $(A\ b)$ 化成如下行梯形:

$$\begin{pmatrix} 1 & 0 & \cdots & 0 & c_{1,r+1} & \cdots & c_{1n} & d_1 \\ 0 & 1 & \cdots & 0 & c_{2,r+1} & \cdots & c_{2n} & d_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 & c_{r,r+1} & \cdots & c_{rn} & d_r \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}$$

2. 求出 $Ax=b$ 的一个特解及导出组 $Ax=0$ 的一个基础解系:

$$\eta_0=\begin{pmatrix} d_1 \\ d_2 \\ \cdots \\ d_r \\ 0 \\ 0 \\ \cdots \\ 0 \end{pmatrix},\quad \alpha_1=\begin{pmatrix} -c_{1,r+1} \\ -c_{2,r+1} \\ \cdots \\ -c_{r,r+1} \\ 1 \\ 0 \\ \cdots \\ 0 \end{pmatrix},\quad \alpha_2=\begin{pmatrix} -c_{1,r+2} \\ -c_{2,r+2} \\ \cdots \\ -c_{r,r+2} \\ 0 \\ 1 \\ \cdots \\ 0 \end{pmatrix},\ \ldots,\ \alpha_{n-r}=\begin{pmatrix} -c_{1n} \\ -c_{2n} \\ \cdots \\ -c_{rn} \\ 0 \\ 0 \\ \cdots \\ 1 \end{pmatrix}.$$

3. 写出 $Ax=b$ 的通解:

$$\eta_0+k_1\alpha_1+k_2\alpha_2+\cdots+k_{n-r}\alpha_{n-r},$$

其中 $k_1,k_2,\ldots,k_{n-r}$ 是任意的实数。

于是所求通解为: $\boldsymbol{x}=\boldsymbol{\eta}+k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2$. □

例 3.4.11　将向量 $\boldsymbol{\beta}$ 表示为向量组 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3\}$ 的线性组合，其中

$$\boldsymbol{\beta}=\begin{pmatrix}2\\0\\-1\end{pmatrix},\ \boldsymbol{\alpha}_1=\begin{pmatrix}\lambda\\2\\1\end{pmatrix},\ \boldsymbol{\alpha}_2=\begin{pmatrix}-2\\3\\1\end{pmatrix},\ \boldsymbol{\alpha}_3=\begin{pmatrix}1\\\mu\\2\end{pmatrix}.$$

解　此即求解

$$\boldsymbol{\beta}=x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+x_3\boldsymbol{\alpha}_3,$$

或等价地求解方程组

$$\begin{cases}\lambda x_1-2x_2+x_3=2,\\2x_1+3x_2+\mu x_3=0,\\x_1+x_2+2x_3=-1.\end{cases}$$

对方程组的增广矩阵进行行初等变换

$$\boldsymbol{B}=\begin{pmatrix}\lambda&-2&1&2\\2&3&\mu&0\\1&1&2&-1\end{pmatrix}\xrightarrow{r_1\leftrightarrow r_3}\begin{pmatrix}1&1&2&-1\\2&3&\mu&0\\\lambda&-2&1&2\end{pmatrix}$$

$$\xrightarrow[r_3+(-\lambda)r_1]{r_2+(-2)r_1}\begin{pmatrix}1&1&2&-1\\0&1&\mu-4&2\\0&-2-\lambda&1-2\lambda&2+\lambda\end{pmatrix}$$

$$\xrightarrow[r_3+(\lambda+2)r_2]{r_1+(-1)r_2}\begin{pmatrix}1&0&6-\mu&-3\\0&1&\mu-4&2\\0&0&(\mu-6)(\lambda+2)+5&3(\lambda+2)\end{pmatrix}=\boldsymbol{B}_1$$

当 $(\mu-6)(\lambda+2)+5=0$ 时，显然有 $\lambda\neq-2$，故有 $2=\mathrm{r}(\boldsymbol{A})<\mathrm{r}(\boldsymbol{B})=3$，所以此时无解.

当 $(\mu-6)(\lambda+2)+5\neq0$ 时，或者 $\lambda=-2$ 或者 $\lambda\neq-2,\mu\neq6-\dfrac{5}{\lambda+2}$，均有

$$\boldsymbol{B}_1\xrightarrow{\left(\frac{1}{(\lambda+2)(\mu-6)+5}\right)r_3}\begin{pmatrix}1&0&6-\mu&-3\\0&1&\mu-4&2\\0&0&1&\dfrac{3(\lambda+2)}{(\lambda+2)(\mu-6)+5}\end{pmatrix}$$

$$\xrightarrow[r_2+(4-\mu)r_3]{r_1+(\mu-6)r_3}\begin{pmatrix}1&0&0&-\dfrac{15}{(\lambda+2)(\mu-6)+5}\\0&1&0&-1-\dfrac{3(2\lambda-1)}{(\lambda+2)(\mu-6)+5}\\0&0&1&\dfrac{3(\lambda+2)}{(\lambda+2)(\mu-6)+5}\end{pmatrix}.$$

即方程组的唯一解为

$$\begin{cases}x_1=-\dfrac{15}{(\lambda+2)(\mu-6)+5},\\x_2=-1-\dfrac{3(2\lambda-1)}{(\lambda+2)(\mu-6)+5},\\x_3=\dfrac{3(\lambda+2)}{(\lambda+2)(\mu-6)+5}.\end{cases}$$

故当且仅当 $\lambda = -2$ 或 $\lambda \neq -2, \mu \neq 6 - \dfrac{5}{\lambda+2}$ 时，$\boldsymbol{\beta}$ 可由 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3$ 线性表示，且线性组合

$$\boldsymbol{\beta} = -\frac{15}{(\lambda+2)(\mu-6)+5}\boldsymbol{\alpha}_1 - \left(1 + \frac{3(2\lambda-1)}{(\lambda+2)(\mu-6)+5}\right)\boldsymbol{\alpha}_2 + \frac{3(\lambda+2)}{(\lambda+2)(\mu-6)+5}\boldsymbol{\alpha}_3$$

是唯一的. □

3.5* 线性最小二乘法

最小二乘法最初是由高斯在解决法国政府进行的私人土地边界的测量数据优化的过程中提出的，在统计数据处理、曲线拟合等许多方面有重要的应用.

如有一组测量的数据

$$(t_1, b_1), (t_2, b_2), \cdots, (t_m, b_m),$$

要找出最吻合的三次多项式 $p(t) = a_3t^3 + a_2t^2 + a_1t + a_0$，也就是我们要找这样的多项式系数 a_3, a_2, a_1, a_0，使得

$$\sum_{i=1}^{m}(p(t_i) - b_i)^2 = \boldsymbol{z}^{\mathrm{T}}\boldsymbol{z}$$

最小，其中

$$\boldsymbol{z} = \boldsymbol{A}\boldsymbol{x} - \boldsymbol{b},\ \ \boldsymbol{A} = \begin{pmatrix} t_1^3 & t_1^2 & t_1 & 1 \\ t_2^3 & t_2^2 & t_2 & 1 \\ \vdots & \vdots & \vdots & \vdots \\ t_m^3 & t_m^2 & t_m & 1 \end{pmatrix}, \boldsymbol{x} = \begin{pmatrix} a_3 \\ a_2 \\ a_1 \\ a_0 \end{pmatrix}, \boldsymbol{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

本质上，求解矛盾线性方程组 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ 也可以通过求 $(\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b})^{\mathrm{T}}(\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b})$ 的最小值点 $\boldsymbol{x}$ 来实现.

若有线性方程组

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b},\ \ \boldsymbol{A} \in \mathbf{R}^{m\times n}, \boldsymbol{x} \in \mathbf{R}^n, \boldsymbol{b} \in \mathbf{R}^m, \tag{3.8}$$

称使得 $(\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b})^{\mathrm{T}}(\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b})$ 最小的 $\boldsymbol{x}$ 为线性方程组 (3.8) 的**最小二乘解**. 求解最小二乘解的方法称为**最小二乘法**.

最小二乘法有多种方法，下面我们介绍**法方程法**. 称线性方程组

$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}, \tag{3.9}$$

为线性方程组 (3.8) 的**法方程**. 法方程法就是通过求法方程的解得到原方程组的最小二乘解的方法，该方法以下列定理为基础.

定理 3.5.1 *线性方程组 (3.8) 的法方程 (3.9) 总有解. 当 $\boldsymbol{A}$ 列满秩时，法方程有唯一解，否则有无穷多组解.*

证明 易知

$$\mathrm{r}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}) \leqslant \mathrm{r}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\ \ \boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}) = \mathrm{r}(\boldsymbol{A}^{\mathrm{T}}(\boldsymbol{A}\ \ \boldsymbol{b})) \leqslant \mathrm{r}(\boldsymbol{A}^{\mathrm{T}}) = \mathrm{r}(\boldsymbol{A}).$$

又由 $\boldsymbol{Ax}=\boldsymbol{\theta}$ 可得 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{Ax}=\boldsymbol{\theta}$，而 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{Ax}=\boldsymbol{\theta}$ 可得 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{Ax}=(\boldsymbol{Ax})^{\mathrm{T}}(\boldsymbol{Ax})=0$，从而 $\boldsymbol{Ax}=\boldsymbol{\theta}$. 即 $\boldsymbol{Ax}=\boldsymbol{\theta}$ 与 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{Ax}=\boldsymbol{\theta}$ 同解，故有 $\mathrm{r}(N(\boldsymbol{A}))=\mathrm{r}(N(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})$，从而也有 $\mathrm{r}(\boldsymbol{A})=\mathrm{r}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})$. 最后可得

$$\mathrm{r}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})\leqslant\mathrm{r}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\ \ \boldsymbol{A}^{\mathrm{T}}\boldsymbol{b})\leqslant\mathrm{r}(\boldsymbol{A})=\mathrm{r}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}),$$

故 $\mathrm{r}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})=\mathrm{r}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\ \ \boldsymbol{A}^{\mathrm{T}}\boldsymbol{b})$，即法方程 (3.9) 有解.

进一步，由 $\mathrm{r}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})=\mathrm{r}(\boldsymbol{A})$ 可知 $\boldsymbol{A}$ 列满秩时，$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 是满秩的方阵，故法方程有唯一解，否则有无穷多组解. □

定理 3.5.2　*$\boldsymbol{x}$ 是线性方程组 (3.8) 的最小二乘解的充要条件是 $\boldsymbol{x}$ 是法方程 (3.9) 的解.*

证明 设 $\boldsymbol{z}=\boldsymbol{Ax}-\boldsymbol{b},\boldsymbol{y}=\boldsymbol{x}+t\boldsymbol{d}$，其中 $\boldsymbol{x},\boldsymbol{y},\boldsymbol{d}$ 为向量，t 为参数. 则有

$$f(t)=(\boldsymbol{Ay}-\boldsymbol{b})^{\mathrm{T}}(\boldsymbol{Ay}-\boldsymbol{b})=(\boldsymbol{z}+t\boldsymbol{Ad})^{\mathrm{T}}(\boldsymbol{z}+t\boldsymbol{Ad})=\boldsymbol{z}^{\mathrm{T}}\boldsymbol{z}+2\boldsymbol{d}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{z}t+\boldsymbol{d}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{Ad}t^2,$$

对 t 求导数得

$$f'(t)=2\boldsymbol{d}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{z}+2\boldsymbol{d}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{Ad}t=2\boldsymbol{d}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}(\boldsymbol{z}+t\boldsymbol{Ad}).$$

若 $\boldsymbol{x}$ 是方程组 (3.8) 的最小二乘解，则对任意的向量 $\boldsymbol{d}$，$f(t)$ 的最小值点在 $t=0$ 处，即有 $f'(0)=2\boldsymbol{d}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{z}=0$. 因为 $\boldsymbol{d}$ 是任意向量，可取 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n$，则有 $\boldsymbol{EA}^{\mathrm{T}}\boldsymbol{z}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{z}=\boldsymbol{A}^{\mathrm{T}}(\boldsymbol{Ax}-\boldsymbol{b})=\boldsymbol{\theta}$，即 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{Ax}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}$.

若有 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{Ax}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}$，则 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{z}=\boldsymbol{\theta}$，故有

$$f(t)=\boldsymbol{z}^{\mathrm{T}}\boldsymbol{z}+\boldsymbol{z}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{Ad}t^2\geqslant\boldsymbol{z}^{\mathrm{T}}\boldsymbol{z}.$$

即对任意的 $\boldsymbol{d}$，$t=0$ 为最小值点，即 $\boldsymbol{x}$ 为 $(\boldsymbol{Ay}-\boldsymbol{b})^{\mathrm{T}}(\boldsymbol{Ay}-\boldsymbol{b})$ 的最小值点，故 $\boldsymbol{x}$ 是 (3.8) 的最小二乘解. □

当线性方程组有解时，方程组的最小二乘解就是方程组的解. 当最小二乘解有多个时，我们还需要考虑如何选择最简单的解，本节简单介绍只有一个最小二乘解的情况，即方程组系数矩阵为列满秩的情况. 下面我们来看一个求线性最小二乘解的例子.

例 3.5.1　求方程组

$$\begin{cases}x_1+2x_2+x_3=6,\\2x_1-x_2+x_3=6,\\2x_1-x_2-x_3=12,\\x_1-3x_2-2x_3=-6\end{cases}$$

的最小二乘解.

解　将方程组写成矩阵形式:

$$\boldsymbol{Ax}=\boldsymbol{b},$$

其中

$$\boldsymbol{A}=\begin{pmatrix}1&2&1\\2&-1&1\\2&-1&-1\\1&-3&-2\end{pmatrix},\quad \boldsymbol{b}=\begin{pmatrix}6\\6\\12\\-6\end{pmatrix}.$$

其法方程为：$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}$. 求解法方程

$$(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\ \boldsymbol{A}^{\mathrm{T}}\boldsymbol{b})=\begin{pmatrix}10&-5&-1&36\\-5&15&8&12\\-1&8&7&12\end{pmatrix}\longrightarrow\begin{pmatrix}1&0&0&5\\0&1&0&3\\0&0&1&-1\end{pmatrix}$$

故方程组的最小二乘解为：$\boldsymbol{x}=(5,3,-1)^{\mathrm{T}}$. □

习　题　三

1. 解下列方程组：

(1) $\begin{cases}3x+2y-z=6,\\x+y=3,\\2x+y+3z=7;\end{cases}$　　(2) $\begin{cases}4x+3y+3z=19,\\-3x+4y+7z=26,\\6x-y-4z=-8;\end{cases}$

(3) $\begin{cases}x+2z=6,\\-3x+5y=-11,\\3y+4z=5;\end{cases}$　　(4) $\begin{cases}x_1+2x_2+x_3-x_4=4,\\2x_1+3x_2-x_3+x_4=-1,\\3x_1+3x_2+7x_3+2x_4=6,\\4x_1+5x_2+x_3+3x_4=-1;\end{cases}$

(5) $\begin{cases}3x_1+7x_2+2x_4=2,\\7x_1+15x_2-3x_3+5x_4=7,\\3x_1+9x_2-2x_3=2,\\2x_1+5x_2+2x_3+x_4=-1;\end{cases}$　　(6) $\begin{cases}2x_1+x_2=0,\\x_1+2x_2+x_3=0,\\x_2+2x_3+x_4=5,\\x_3+2x_4+x_5=2,\\x_4+2x_5=-5;\end{cases}$

(7) $\begin{cases}x_1+2x_2=3,\\2x_1+5x_2=5,\\2x_3+x_4+x_5=2,\\x_3+2x_4+x_5=-1,\\x_3+x_4+2x_5=-1.\end{cases}$

2. 设 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_s\in\mathbf{R}^n$ 是非齐次方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$, $\boldsymbol{b}\neq\boldsymbol{\theta}$ 的解，证明 $\boldsymbol{\beta}=k_1\boldsymbol{\alpha}_1+\cdots+k_s\boldsymbol{\alpha}_s$ 是齐次方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{\theta}$ 的解的充要条件是 $k_1+\cdots+k_s=0$, 其中 $\boldsymbol{A}\in\mathbf{R}^{m\times n}$.

3. 证明 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{\theta}$ 与 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=\boldsymbol{\theta}$ 是同解方程组, 其中 $\boldsymbol{A}\in\mathbf{R}^{m\times n}$.

若 $A^TAx=0$
则 $x^TA^TAx=0$
∴ $(Ax)^TAx=0$
∴ $Ax=0$
即 $A^TAx=0$ 与 $Ax=0$ 同解.

4. 判断下列方程组是否有解，若有解，求出方程组的解:

(1) $\begin{cases}2x+y-2z=1,\\3x+y+5z=9,\\x+y+3z=5;\end{cases}$　　(2) $\begin{cases}2x+y-2z=2,\\3x+3y+5z=14,\\x-y-9z=10;\end{cases}$

(3) $\begin{cases}2x-y+2z=-4,\\3x-2y+5z=-5,\\x-y+3z=-1;\end{cases}$　　(4) $\begin{cases}3x+y-4z=1,\\7x-2y+9z=12,\\5x-5y+7z=2,\\3x+3y-5z=4;\end{cases}$

(5) $\begin{cases} x+2y-3z=-6, \\ 7x+3y+6z=4, \\ -5x+3y+8z=10, \\ x+y+6z=4; \end{cases}$　(6) $\begin{cases} x+2y-3z=3, \\ 3x+2y-z=5, \\ -2x+3y-8z=1, \\ 2x+y=3; \end{cases}$

(7) $\begin{cases} x_1-2x_2+x_3+7x_4+2x_5=13, \\ 3x_1+7x_2-x_3+2x_4+x_5=20, \\ 2x_1-17x_2+6x_3+33x_4+9x_5=39; \end{cases}$

(8) $\begin{cases} 2x_1-3x_2+x_3+x_4-x_5=-5, \\ -x_1+3x_2+4x_3-2x_4+2x_5=1, \\ 4x_1-9x_2-10x_3+5x_4-2x_5=-4. \end{cases}$

5. 求下列带参方程组的解 (何时有解, 何时无解, 何时有无穷多解):

(1) $\begin{cases} 2x+y=4, \\ 3x+\lambda y+2z=13, \\ 2x-2y+3z=7; \end{cases}$

(2) $\begin{cases} x_1+2x_2+x_3-x_4=1, \\ x_1+3x_2+2x_3+3x_4=1, \\ 3x_1+2x_2+x_3-11x_4=a, \\ 2x_1+3x_2+3x_3+bx_4=-6. \end{cases}$

6. 证明线性方程组

$$\begin{cases} x_1-2x_2+x_3=b_1, \\ x_2-2x_3+x_4=b_2, \\ \quad\cdots\cdots \\ x_{n-2}-2x_{n-1}+x_n=b_{n-2}, \\ x_1+x_{n-1}-2x_n=b_{n-1}, \\ -2x_1+x_2+x_n=b_n \end{cases}$$

有解的充要条件为 $\sum_{i=1}^{n} b_i=0$.

7. 求下列齐次方程组的基础解系和通解:

(1) $\begin{cases} 2x+y+7z=0, \\ 3x+3y+5z=0, \\ 7x+2y+30z=0; \end{cases}$　(2) $\begin{cases} x_1+2x_2-x_3+x_4=0, \\ 2x_1+x_2+4x_3-x_4=0, \\ x_1-2x_2+7x_3-3x_4=0; \end{cases}$

(3) $\begin{cases} 3x-y+5z=0, \\ 2x+y+5z=0, \\ x+3y+5z=0, \\ x+2y+4z=0; \end{cases}$　(4) $\begin{cases} 2x_1-4x_2+x_3+2x_4=0, \\ -x_1+2x_2+x_3+5x_4=0, \\ 3x_1-6x_2+3x_3+9x_4=0, \\ 2x_1-4x_2+2x_3+6x_4=0. \end{cases}$

8. 设 $\begin{vmatrix} a_{01} & \cdots & a_{0n} & b_0 \\ a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{n1} & \cdots & a_{nn} & b_n \end{vmatrix} \neq 0$，证明 $\begin{cases} a_{01}x_1 + \cdots + a_{0n}x_n = b_0, \\ a_{11}x_1 + \cdots + a_{1n}x_n = b_1, \\ \quad\cdots\cdots \\ a_{n1}x_1 + \cdots + a_{nn}x_n = b_n \end{cases}$ 无解.

9. 证明：方程组 $\begin{pmatrix} \boldsymbol{A} \\ \boldsymbol{b}^{\mathrm{T}} \end{pmatrix} \boldsymbol{x} = \boldsymbol{\theta}$ 与方程组 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{\theta}$ 是同解方程组的充要条件是 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{y} = \boldsymbol{b}$ 有解, 其中 $\boldsymbol{A} \in \mathbf{R}^{m\times n}$.

10. 已知 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{\theta}$ 的一个基础解系是 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_r$，求 $(\boldsymbol{A}\ \ \boldsymbol{A})\begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{pmatrix} = \boldsymbol{\theta}$ 的一个基础解系, 其中 $\boldsymbol{A} \in \mathbf{R}^{m\times n}$.

11. 已知齐次方程组 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{\theta}$ 的系数矩阵 $\boldsymbol{A} \in \mathbf{R}^{n\times n}$ 有 $A_{11} \neq 0$ 且 $|\boldsymbol{A}| = 0$，证明 $(A_{11}, \cdots, A_{1n})^{\mathrm{T}}$ 是 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{\theta}$ 的一个基础解系，其中 $A_{11}, \cdots, A_{1n}$ 是行列式 $|\boldsymbol{A}|$ 第一行元素的代数余子式.

12. 设 $\boldsymbol{A} \in \mathbf{R}^{r\times n}, \boldsymbol{B} \in \mathbf{R}^{(n-r)\times n}$ 都是行满秩矩阵, 且有关系 $\boldsymbol{A}\boldsymbol{B}^{\mathrm{T}} = \boldsymbol{O}$, 证明 $\boldsymbol{B}^{\mathrm{T}}$ 的列构成 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{\theta}$ 的一个基础解系, $\boldsymbol{A}^{\mathrm{T}}$ 的列构成 $\boldsymbol{B}\boldsymbol{y} = \boldsymbol{\theta}$ 的一个基础解系.

13. 证明：若 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{\theta}$ 的解一定是 $\boldsymbol{B}\boldsymbol{x} = \boldsymbol{\theta}$ 的解，则 $\boldsymbol{B}$ 的行向量都能表示成 $\boldsymbol{A}$ 的行向量的线性组合，其中 $\boldsymbol{A} \in \mathbf{R}^{m\times n}, \boldsymbol{B} \in \mathbf{R}^{p\times n}$.

14. 证明：若 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3 \in \mathbf{R}^n$ 是 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{\theta}$ 的一个基础解系，则 $\boldsymbol{\beta}_1 = 2\boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_2 + \boldsymbol{\alpha}_3, \boldsymbol{\beta}_2 = \boldsymbol{\alpha}_1 + 2\boldsymbol{\alpha}_2 + \boldsymbol{\alpha}_3, \boldsymbol{\beta}_3 = \boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_2 + 2\boldsymbol{\alpha}_3$ 也是 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{\theta}$ 的一个基础解系, 其中 $\boldsymbol{A} \in \mathbf{R}^{m\times n}$.

15. 设线性方程组

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = 0, \\ a_{21}x_1 + \cdots + a_{2n}x_n = 0, \\ \quad\cdots\cdots \\ a_{n-1,1}x_1 + \cdots + a_{n-1,n}x_n = 0 \end{cases}$$

11. 由 $A_{11} \neq 0$ 可得 $r(A) \geq n-1$.
由 $|A| = 0$ 可得 $r(A) \leq n-1$.
因此，$r(A) = n-1$.
从而 $Ax = 0$ 的基础解系包含一个非 0 向量。
由于 $AA^* = |A|E = O$,
故 $(A_{11}\ A_{12}\ \cdots\ A_{1n})^T$ 是 $Ax = 0$ 的一个非 0 解，
从而是 $Ax = 0$ 的一个基础解系。

的系数矩阵

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & & \vdots \\ a_{n-1,1} & \cdots & a_{n-1,n} \end{pmatrix}$$

为行满秩，M_i 为 $\boldsymbol{A}$ 中划去第 i 列后剩下的 $(n-1)\times(n-1)$ 阶矩阵的行列式，证明 $(M_1, -M_2, \cdots, (-1)^{n-1}M_n)^{\mathrm{T}}$ 为方程组的一个基础解系.

16. 设 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{\theta}$ 的一个基础解系为 $\boldsymbol{\alpha}_1 = (0,0,1,1)^{\mathrm{T}}, \boldsymbol{\alpha}_2 = (1,2,0,-1)^{\mathrm{T}}$, $\boldsymbol{B}\boldsymbol{x} = \boldsymbol{\theta}$ 的一个基础解系为 $\boldsymbol{\beta}_1 = (1,-1,1,1)^{\mathrm{T}}, \boldsymbol{\beta}_2 = (2,1,1,0)^{\mathrm{T}}$, 其中 $\boldsymbol{A} \in \mathbf{R}^{m\times 4}$, $\boldsymbol{B} \in \mathbf{R}^{p\times 4}$, 求

$$\begin{pmatrix} \boldsymbol{A} \\ \boldsymbol{B} \end{pmatrix} \boldsymbol{x} = \boldsymbol{\theta}$$

的一个基础解系.

17. 设方程组

$$\begin{cases} x_1 - x_2 + x_3 = 0, \\ -2x_1 + x_2 + x_4 = 0 \end{cases}$$

与方程组

$$\begin{cases} 2x_1 - x_2 + ax_3 + bx_4 = 0, \\ x_1 - ax_2 + bx_3 - x_4 = 0 \end{cases}$$

有相同的解集，求 a, b 的值.

18. 求下列非齐次方程组的通解：

(1) $\begin{cases} 3x + 2y + 2z = -3, \\ 2x + 3y + 8z = 3, \\ 4x + 3y + 4z = -3; \end{cases}$ (2) $\begin{cases} 2x_1 + x_2 + 7x_3 + 3x_4 = 5, \\ 5x_1 + 16x_2 + 31x_3 - 6x_4 = 26, \\ x_1 - x_2 + 2x_3 + 3x_4 = 1. \end{cases}$

19. 将向量 $\boldsymbol{\beta}$ 表示为向量组 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3\}$ 的线性组合，其中

$$\boldsymbol{\beta} = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}, \quad \boldsymbol{\alpha}_1 = \begin{pmatrix} -1 \\ 1 \\ -5 \end{pmatrix}, \quad \boldsymbol{\alpha}_2 = \begin{pmatrix} \lambda \\ -1 \\ 5 \end{pmatrix}, \quad \boldsymbol{\alpha}_3 = \begin{pmatrix} 2 \\ \lambda \\ 4 \end{pmatrix}.$$

20. 设 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_r \in \mathbf{R}^n$ 是 $\boldsymbol{Ax} = \boldsymbol{\theta}$ 的一个基础解系, 其中 $\boldsymbol{A} \in \mathbf{R}^{m\times n}$，$\boldsymbol{\eta} \in \mathbf{R}^n$ 是 $\boldsymbol{Ax} = \boldsymbol{b}, \boldsymbol{b} \neq \boldsymbol{\theta}$ 的一个特解，又设 $\boldsymbol{\beta}_0 = \boldsymbol{\eta}, \boldsymbol{\beta}_1 = \boldsymbol{\eta} + \boldsymbol{\alpha}_1, \cdots, \boldsymbol{\beta}_r = \boldsymbol{\eta} + \boldsymbol{\alpha}_r$，证明：

(1) $\boldsymbol{\beta}_0, \cdots, \boldsymbol{\beta}_r$ 线性无关;

(2) $\boldsymbol{Ax} = \boldsymbol{b}$ 的任意解 $\boldsymbol{\beta}$ 都能表示成 $\boldsymbol{\beta}_0, \cdots, \boldsymbol{\beta}_r$ 的线性组合，即 $\boldsymbol{\beta} = k_0\boldsymbol{\beta}_0 + k_1\boldsymbol{\beta}_1 + \cdots + k_r\boldsymbol{\beta}_r$，其中 $k_0 + k_1 + \cdots + k_r = 1$.

21*. 求下列方程组的最小二乘解

$$\begin{cases} 3x_1 + x_2 - 2x_3 = -5, \\ x_1 + 3x_2 - 3x_3 = -5, \\ 2x_1 - 7x_3 = -3, \\ x_1 - x_2 + 3x_3 = 11. \end{cases}$$

三大关系：等价矩阵
相似矩阵

第 4 章　矩阵的特征值与特征向量

矩阵的特征值与特征向量可以用来计算在 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ 意义下的最简矩阵，而 $\boldsymbol{A}$ 与 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ 表示在不同坐标系下的同一个线性变换.

4.1　相 似 矩 阵

矩阵分解经常能解决很多问题，我们要考虑一种特殊的矩阵分解：$\boldsymbol{A}=\boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^{-1}$，其中 $\boldsymbol{\Lambda}=\operatorname{diag}(\lambda_1,\cdots,\lambda_n)$. 这种分解非常有用，问题是对一般方阵，这种分解是否存在？在存在这种分解的情况下，又怎样去求出其中的 $\boldsymbol{P}$ 和 $\boldsymbol{\Lambda}$？下面就要讨论这些问题.

定义1、**定义 4.1.1** (相似矩阵)　对于同阶方阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$, 如果存在可逆矩阵 $\boldsymbol{P}$, 使得

$$\boldsymbol{B}=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P},$$

则称 $\boldsymbol{A}$ **相似于** $\boldsymbol{B}$, 记为 $\boldsymbol{A}\sim\boldsymbol{B}$. 称 $\boldsymbol{B}$ 为 $\boldsymbol{A}$ 的**相似矩阵**, 而称 $\boldsymbol{P}$ 为 $\boldsymbol{A}$ 到 $\boldsymbol{B}$ 的**相似变换矩阵**.

容易证明，相似矩阵满足等价关系的三个性质：

(1) **自反性**：对任意矩阵 $\boldsymbol{A}$，都成立 $\boldsymbol{A}\sim\boldsymbol{A}$;　　$B=P^{-1}AP$

(2) **对称性**：若 $\boldsymbol{A}\sim\boldsymbol{B}$，则 $\boldsymbol{B}\sim\boldsymbol{A}$;　　$B^{-1}=P^{-1}A^{-1}(P^{-1})^{-1}$

(3) **传递性**：若 $\boldsymbol{A}\sim\boldsymbol{B},\boldsymbol{B}\sim\boldsymbol{C}$, 则 $\boldsymbol{A}\sim\boldsymbol{C}$.　　$\Rightarrow B^{-1}=P^{-1}A^{-1}P$

此外, 相似矩阵之间也有许多共同的性质：

性质 1　若 $\boldsymbol{A}\sim\boldsymbol{B}$, 则 $|\boldsymbol{A}|=|\boldsymbol{B}|$, 从而 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 可逆性相同.

证明　由 $\boldsymbol{B}=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ 得 $|\boldsymbol{B}|=|\boldsymbol{P}^{-1}||\boldsymbol{A}||\boldsymbol{P}|=|\boldsymbol{P}|^{-1}|\boldsymbol{A}||\boldsymbol{P}|=|\boldsymbol{A}|$. □

性质 2　若 $\boldsymbol{A}\sim\boldsymbol{B}$, 且 $\boldsymbol{A}$ 或 $\boldsymbol{B}$ 可逆, 则 $\boldsymbol{A}^{-1}\sim\boldsymbol{B}^{-1}$.

证明　若 $\boldsymbol{A}\sim\boldsymbol{B}$, 则由性质 1 知 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 同为不可逆或同为可逆. 在 $\boldsymbol{A}$ 或 $\boldsymbol{B}$ 为可逆的情形, 对等式 $\boldsymbol{B}=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ 两边取逆矩阵得 $\boldsymbol{B}^{-1}=\boldsymbol{P}^{-1}\boldsymbol{A}^{-1}\boldsymbol{P}$, 故 $\boldsymbol{A}^{-1}\sim\boldsymbol{B}^{-1}$. □

性质 3　若 $\boldsymbol{A}\sim\boldsymbol{B}$, 则 $\boldsymbol{A}^n\sim\boldsymbol{B}^n,k\boldsymbol{A}\sim k\boldsymbol{B}$，其中 n 为自然数，k 为任意实数.

证明　因为 $\boldsymbol{A}\sim\boldsymbol{B}$, 所以存在可逆矩阵 $\boldsymbol{P}$ 使 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{B}$, 于是对任何正整数 n, 有

$$\begin{aligned}\boldsymbol{B}^n&=\underbrace{(\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P})(\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P})\cdots(\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P})}_{n\text{ 个括弧}}\\&=\boldsymbol{P}^{-1}\boldsymbol{A}(\boldsymbol{P}\boldsymbol{P}^{-1})\boldsymbol{A}(\boldsymbol{P}\boldsymbol{P}^{-1})\boldsymbol{A}\cdots(\boldsymbol{P}\boldsymbol{P}^{-1})\boldsymbol{A}\boldsymbol{P}=\boldsymbol{P}^{-1}\boldsymbol{A}^n\boldsymbol{P},\end{aligned}$$

即 $\boldsymbol{A}^n\sim\boldsymbol{B}^n$. 另外易知 $k\boldsymbol{B}=\boldsymbol{P}^{-1}(k\boldsymbol{A})\boldsymbol{P}$. □

性质 3 在求矩阵的正整数幂时非常有用, 且看下面的例子.

$A=P^{-1}BP$

一个矩阵是否与某个对角阵相似？

求 $A^n=P^{-1}B^nP=P^{-1}\begin{pmatrix} a_1^n & & & \\ & a_2^n & & \\ & & a_3^n\cdots & \\ & & & a_n^n \end{pmatrix}P$

例 4.1.1　若 $\boldsymbol{P}^{-1}\boldsymbol{AP}$ 为对角矩阵, 其中

$$\boldsymbol{P}=\begin{pmatrix} 4 & 1 \\ -1 & -1 \end{pmatrix},\quad \boldsymbol{A}=\begin{pmatrix} 3 & 4 \\ -1 & -2 \end{pmatrix},$$

试求 $\boldsymbol{A}^n$(n 为正整数).

解　容易求得 $\boldsymbol{P}^{-1}=\dfrac{1}{3}\begin{pmatrix} 1 & 1 \\ -1 & -4 \end{pmatrix}$, 于是

$$\boldsymbol{B}=\boldsymbol{P}^{-1}\boldsymbol{AP}=\frac{1}{3}\begin{pmatrix} 1 & 1 \\ -1 & -4 \end{pmatrix}\begin{pmatrix} 3 & 4 \\ -1 & -2 \end{pmatrix}\begin{pmatrix} 4 & 1 \\ -1 & -1 \end{pmatrix}=\begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}.$$

由 $\boldsymbol{P}^{-1}\boldsymbol{AP}=\boldsymbol{B}$ 得 $\boldsymbol{A}=\boldsymbol{PBP}^{-1}$, 于是

$$\begin{aligned}\boldsymbol{A}^n=(\boldsymbol{PBP}^{-1})^n=\boldsymbol{PB}^n\boldsymbol{P}^{-1}&=\begin{pmatrix} 4 & 1 \\ -1 & -1 \end{pmatrix}\begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}^n\frac{1}{3}\begin{pmatrix} 1 & 1 \\ -1 & -4 \end{pmatrix}\\ &=\frac{1}{3}\begin{pmatrix} 4 & 1 \\ -1 & -1 \end{pmatrix}\begin{pmatrix} 2^n & 0 \\ 0 & (-1)^n \end{pmatrix}\begin{pmatrix} 1 & 1 \\ -1 & -4 \end{pmatrix}\\ &=\frac{1}{3}\begin{pmatrix} 2^{n+2}+(-1)^{n+1} & 4(2^n+(-1)^{n+1}) \\ -2^n+(-1)^n & -2^n+4(-1)^n \end{pmatrix}.\end{aligned}$$

$P_n(x)$ □

性质 4　若 $\boldsymbol{A}\sim\boldsymbol{B}$, 则 $f(\boldsymbol{A})\sim f(\boldsymbol{B})$, 其中 $f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$ 为任意多项式.

$f(A)=a_nA^n+a_{n-1}A^{n-1}+\cdots+a_1A+a_0E$.

证明　由 $\boldsymbol{B}=\boldsymbol{P}^{-1}\boldsymbol{AP}$ 可得

$$\begin{aligned}f(\boldsymbol{B})&=a_n\boldsymbol{B}^n+\cdots+a_1\boldsymbol{B}+a_0\boldsymbol{E}\\ &=a_n\underbrace{(\boldsymbol{P}^{-1}\boldsymbol{AP})(\boldsymbol{P}^{-1}\boldsymbol{AP})\cdots(\boldsymbol{P}^{-1}\boldsymbol{AP})}_{n\text{ 个括弧}}+\cdots+a_1\boldsymbol{P}^{-1}\boldsymbol{AP}+a_0\boldsymbol{E}\\ &=a_n\boldsymbol{P}^{-1}\boldsymbol{A}^n\boldsymbol{P}+\cdots+a_1\boldsymbol{P}^{-1}\boldsymbol{AP}+a_0\boldsymbol{P}^{-1}\boldsymbol{EP}\\ &=\boldsymbol{P}^{-1}(a_n\boldsymbol{A}^n+\cdots+a_1\boldsymbol{A}+a_0\boldsymbol{E})\boldsymbol{P}=\boldsymbol{P}^{-1}f(\boldsymbol{A})\boldsymbol{P},\end{aligned}$$

即 $f(\boldsymbol{A})\sim f(\boldsymbol{B})$. □

相似矩阵有其几何意义，在第六章将看到相似矩阵是同一个线性变换在不同坐标系下的表示 (定理 6.4.6).　矩阵相似的必要条件?

4.2　特征值与特征向量

这一节我们扩大到复数范围进行讨论.

对于一个 n 阶方阵 $\boldsymbol{A}$，如果有 $\boldsymbol{P}=(\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\cdots,\boldsymbol{\xi}_n),\boldsymbol{\Lambda}=\mathrm{diag}(\lambda_1,\cdots,\lambda_n)$ 使得 $\boldsymbol{A}=\boldsymbol{P\Lambda P}^{-1}$，则 $\boldsymbol{AP}=\boldsymbol{P\Lambda}$，所以有

$$\boldsymbol{A\xi}_i=\lambda_i\boldsymbol{\xi}_i,\ i=1,2,\cdots,n.$$

下面就从 $\boldsymbol{A\xi}_i=\lambda_i\boldsymbol{\xi}_i$ 入手讨论上面提到的问题.

Q_1: $A\sim\mathrm{diag}(\lambda_1,\cdots,\lambda_n)$ 的条件是什么?

Q_2: $A\sim\mathrm{diag}(\lambda_1,\cdots,\lambda_n)$ 时,

$A^n=P^{-1}\mathrm{diag}(\cdots)^nP$.

如何求 P?

由 $P^{-1}AP=\begin{pmatrix}\lambda_1 & & \\ & \lambda_2\ddots & \\ & & \lambda_n\end{pmatrix}$

$PP^{-1}AP=P\begin{pmatrix}\lambda_1 & & \\ & \lambda_2\ddots & \\ & & \lambda_n\end{pmatrix}$　$P=(\xi_1,\xi_2,\cdots,\xi_n)$

即 $A(\xi_1,\xi_2\cdots\xi_n)=(\xi_1,\xi_2\cdots\xi_n)\begin{pmatrix}\lambda_1 & & \\ & \lambda_2\ddots & \\ & & \lambda_n\end{pmatrix}$

$\Rightarrow(A\xi_1,A\xi_2,\cdots,A\xi_n)=(\lambda_1\xi_1,\lambda_2\xi_2,\cdots,\lambda_n\xi_n)$

$A\xi_i=\lambda_i\xi_i\quad(i=1,2,\cdots,n)$

若存在可逆阵 P，有 $P^{-1}AP = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$

则 ξ_i 是 A 的特征向量，且 $\{\xi_1, \xi_2, \cdots, \xi_n\}$ 线性无关

则 A 必有 n 个线性无关的特征向量.

如何求 ξ_i?　　　　是关于 λ 的 n 次方程

$A\xi = \lambda\xi \iff (\lambda E - A)\xi = 0 \iff (\lambda E - A)x = 0$ 有非零解 $\iff |\lambda E - A| = 0$

即 λ 为矩阵 A 的特征值 $\iff |\lambda E - A| = 0$

ξ 为 A 的属于 λ 的特征向量 $\iff \xi$ 是 $(\lambda E - A)x = 0$ 的非零解.

定义2. **定义 4.2.1** (特征值、特征向量)　设 $\boldsymbol{A}$ 是实数域 $\mathbf{R}$ 或复数域 $\mathbf{C}$ 上的一个方阵，$\lambda \in \mathbf{C}$，若存在非零向量 $\boldsymbol{\xi}$ 使得 $\boldsymbol{A\xi} = \lambda\boldsymbol{\xi}$，则称 λ 为矩阵 $\boldsymbol{A}$ 的**特征值**，$\boldsymbol{\xi}$ 称为 $\boldsymbol{A}$ 的属于特征值 λ 的**特征向量**. ξ 决定 λ.

对于特征向量而言，有如下性质：

定理 4.2.1　设方阵 $\boldsymbol{A}$ 有特征值 λ，$\boldsymbol{\xi}_1, \boldsymbol{\xi}_2$ 为属于 λ 的特征向量，则它们的任意不等于零向量的线性组合 $\boldsymbol{\eta} = k_1\boldsymbol{\xi}_1 + k_2\boldsymbol{\xi}_2 (k_1, k_2 \in \mathbf{R})$ 仍是属于 λ 的特征向量.

证明　直接验证如下

$$\boldsymbol{A\eta} = \boldsymbol{A}(k_1\boldsymbol{\xi}_1 + k_2\boldsymbol{\xi}_2) = k_1\boldsymbol{A\xi}_1 + k_2\boldsymbol{A\xi}_2 = k_1\lambda\boldsymbol{\xi}_1 + k_2\lambda\boldsymbol{\xi}_2 = \lambda(k_1\boldsymbol{\xi}_1 + k_2\boldsymbol{\xi}_2) = \lambda\boldsymbol{\eta}.$$

□

由定理 4.2.1 可知若 $\boldsymbol{\xi}$ 是属于 λ 的一个特征向量，则 $k \neq 0$ 时 $k\boldsymbol{\xi}$ 也是属于 λ 的特征向量. 于是有：如果 λ 是 $\boldsymbol{A}$ 的特征值，则属于 λ 的特征向量有无穷多个.

现在我们讨论如何求出 $\boldsymbol{A}$ 的全部特征值和全部特征向量. 因为方程 $\boldsymbol{A\xi} = \lambda\boldsymbol{\xi}$ 等价于齐次线性方程组 $(\lambda\boldsymbol{E} - \boldsymbol{A})\boldsymbol{\xi} = \boldsymbol{\theta}$，故求特征值、特征向量即是求 $(\lambda\boldsymbol{E} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{\theta}$ 非零解的问题. 由于此齐次线性方程组有非零解的充要条件是系数行列式为零，即 $|\lambda\boldsymbol{E} - \boldsymbol{A}| = 0$，由此可求得特征值 λ，进一步可求得 $(\lambda\boldsymbol{E} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{\theta}$ 的非零解，此即属于特征值 λ 的特征向量.

定义3. **定义 4.2.2** (特征多项式、特征方程、特征矩阵)　$|\lambda\boldsymbol{E} - \boldsymbol{A}|$ 称为 $\boldsymbol{A}$ 的**特征多项式**；方程 $|\lambda\boldsymbol{E} - \boldsymbol{A}| = 0$ 称为 $\boldsymbol{A}$ 的**特征方程**. 方程 $|\lambda\boldsymbol{E} - \boldsymbol{A}| = 0$ 的解也称为 $\boldsymbol{A}$ 的**特征根**. 而 $\lambda\boldsymbol{E} - \boldsymbol{A}$ 称为 $\boldsymbol{A}$ 的**特征矩阵**.

由上面的定义和讨论知，$\boldsymbol{A}$ 的特征根与 $\boldsymbol{A}$ 的特征值是相同的. 对 n 阶方阵 $\boldsymbol{A}$，$\boldsymbol{A}$ 的特征多项式是 n 次多项式，由代数基本定理知，$\boldsymbol{A}$ 有 n 个特征值 (包括重数). 对于实数矩阵 $\boldsymbol{A}$，若 λ 是它的实特征值，则 $(\lambda\boldsymbol{E} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{\theta}$ 必有实非零解. 于是对于实数矩阵的实特征值，我们只考虑属于它的实特征向量. 而对于实数矩阵的复特征值，我们必须考虑属于它的复特征向量. 综合上面的讨论，我们给出求矩阵 $\boldsymbol{A}$ 的全部特征值和特征向量的计算步骤如下：

根：$\lambda_1^{s_1}, \lambda_2^{s_2}, \cdots, \lambda_n^{s_n}$，其中 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 两两不同，$s_1 + s_2 + \cdots + s_r = n$.

(1) ① 计算行列式 $|\lambda\boldsymbol{E} - \boldsymbol{A}|$，并 ② 求出 $|\lambda\boldsymbol{E} - \boldsymbol{A}| = 0$ 的全部根，即 $\boldsymbol{A}$ 的特征值；

(2) ③ 对于每个特征值 λ_i，求齐次线性方程组

$$(\lambda_i\boldsymbol{E} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{\theta}$$

的一个基础解系 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_{s_i}$；

(3) 写出 $\boldsymbol{A}$ 属于 λ_i 的全部特征向量为

$$k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2 + \cdots + k_{s_i}\boldsymbol{\alpha}_{s_i},$$

其中 $k_1, k_2, \cdots, k_{s_i}$ 为不全为零的任意常数. 而所求得的所有特征值的所有特征向量，就是 $\boldsymbol{A}$ 的全部特征向量.

注　在下一节我们将会看到，对于重特征值 λ，虽然有多个属于 λ 的线性无关的特征向量，但其个数不会超过 λ 的重数. 对于单特征值，有且只有一个线性无关的特征向量.

下面我们给出一些求矩阵特征值和特征向量的例子.

例 4.2.1　求矩阵

$$A=\begin{pmatrix}5&-2&1\\0&4&0\\1&-2&5\end{pmatrix}$$

的全部特征值和特征向量.

解　由

$$|\lambda E-A|=\begin{vmatrix}\lambda-5&2&-1\\0&\lambda-4&0\\-1&2&\lambda-5\end{vmatrix}=\begin{vmatrix}\lambda-6&2&-1\\0&\lambda-4&0\\\lambda-6&2&\lambda-5\end{vmatrix}=(\lambda-6)(\lambda-4)^2$$

得 A 的两个特征值为：$\lambda=6,4$(二重).

对于 $\lambda=6$，解齐次方程组 $(6E-A)x=\theta$, 由

$$\begin{pmatrix}1&2&-1\\0&2&0\\-1&2&1\end{pmatrix}\xrightarrow[r_3-2r_2]{r_3+r_1}\begin{pmatrix}1&2&-1\\0&2&0\\0&0&0\end{pmatrix}\xrightarrow[\frac{1}{2}r_2]{r_1-r_2}\begin{pmatrix}1&0&-1\\0&1&0\\0&0&0\end{pmatrix}$$

求得该齐次方程组的一个基础解系为：$\alpha_1=(1,0,1)^{\mathrm{T}}$. 故属于特征值 6 的全部特征向量为：$k_1\alpha_1$, 其中 k_1 为任意非零常数.

对于 $\lambda=4$，解齐次方程组 $(4E-A)x=\theta$, 由

$$\begin{pmatrix}-1&2&-1\\0&0&0\\-1&2&-1\end{pmatrix}\to\begin{pmatrix}1&-2&1\\0&0&0\\0&0&0\end{pmatrix}$$

求得该齐次方程组的一个基础解系为

$$\alpha_2=\begin{pmatrix}2\\1\\0\end{pmatrix},\quad \alpha_3=\begin{pmatrix}-1\\0\\1\end{pmatrix}.$$

故属于特征值 4 的全部特征向量为：$k_2\alpha_2+k_3\alpha_3$, 其中 k_2,k_3 为不全为零的任意常数.　□

例 4.2.2　求矩阵

$$A=\begin{pmatrix}3&0&1\\2&0&2\\3&-4&5\end{pmatrix}$$

的全部特征值和特征向量.

解　由

$$|\lambda E-A|=\begin{vmatrix}\lambda-3&0&-1\\-2&\lambda&-2\\-3&4&\lambda-5\end{vmatrix}=\begin{vmatrix}\lambda-4&0&-1\\\lambda-4&\lambda&-2\\\lambda-4&4&\lambda-5\end{vmatrix}=(\lambda-4)(\lambda-2)^2$$

得 $\boldsymbol{A}$ 的两个特征值为: $\lambda=4,2$(二重)

对于 $\lambda=4$，解齐次方程组 $(4\boldsymbol{E}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{\theta}$, 由

$$\begin{pmatrix} 1 & 0 & -1 \\ -2 & 4 & -2 \\ -3 & 4 & -1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}$$

求得该齐次方程组的一个基础解系为：$\boldsymbol{\alpha}_1=(1,1,1)^{\mathrm{T}}$. 故属于特征值 4 的全部特征向量为：$k_1\boldsymbol{\alpha}_1$, 其中 k_1 为任意非零常数.

对于 $\lambda=2$，解齐次方程组 $(2\boldsymbol{E}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{\theta}$, 由

$$\begin{pmatrix} -1 & 0 & -1 \\ -2 & 2 & -2 \\ -3 & 4 & -3 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

求得该齐次方程组的一个基础解系为：$\boldsymbol{\alpha}_2=(-1,0,1)^{\mathrm{T}}$. 故属于特征值 2 的全部特征向量为：$k_2\boldsymbol{\alpha}_2$, 其中 k_2 为任意非零常数. □

例 4.2.3　求矩阵

$$\boldsymbol{A}=\begin{pmatrix} 2 & -1 & 1 \\ 1 & 1 & 1 \\ 1 & -1 & 2 \end{pmatrix}$$

的全部特征值和特征向量.

解　由

$$|\lambda\boldsymbol{E}-\boldsymbol{A}|=\begin{vmatrix} \lambda-2 & 1 & -1 \\ -1 & \lambda-1 & -1 \\ -1 & 1 & \lambda-2 \end{vmatrix}=(\lambda-1)(\lambda^2-4\lambda+5)$$

得 $\boldsymbol{A}$ 的特征值为：$\lambda=1,2\pm i$.

对于 $\lambda=1$，解齐次方程组 $(\boldsymbol{E}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{\theta}$, 由

$$\begin{pmatrix} -1 & 1 & -1 \\ -1 & 0 & -1 \\ -1 & 1 & -1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

求得该齐次方程组的一个基础解系为：$\boldsymbol{\alpha}_1=(-1,0,1)^{\mathrm{T}}$. 故属于特征值 1 的全部特征向量为：$k_1\boldsymbol{\alpha}_1$, 其中 k_1 为任意非零实常数.

对于 $\lambda=2+i$，解齐次方程组 $((2+i)\boldsymbol{E}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{\theta}$, 由

$$\begin{pmatrix} i & 1 & -1 \\ -1 & 1+i & -1 \\ -1 & 1 & i \end{pmatrix} \xrightarrow[(-i)r_1]{\substack{r_2-ir_1 \\ r_3-ir_1}} \begin{pmatrix} 1 & -i & i \\ 0 & 1 & -1+i \\ 0 & 1-i & 2i \end{pmatrix} \xrightarrow{\substack{r_1+ir_2 \\ r_3-(1-i)r_2}} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1+i \\ 0 & 0 & 0 \end{pmatrix}$$

求得该齐次方程组的一个基础解系为: $\boldsymbol{\alpha}_2=(1,1-i,1)^{\mathrm{T}}$. 故属于特征值 $2+i$ 的全部特征向量为: $k_2\boldsymbol{\alpha}_2$, 其中 k_2 为任意非零复常数.

对于 $\lambda=2-i$, 对刚得到的结果 $((2+i)\boldsymbol{E}-\boldsymbol{A})\boldsymbol{\alpha}_2=\boldsymbol{\theta}$ 两边取共轭得 $((2-i)\boldsymbol{E}-\boldsymbol{A})\overline{\boldsymbol{\alpha}}_2=\boldsymbol{\theta}$. 故 $\overline{\boldsymbol{\alpha}}_2=(1,1+i,1)^{\mathrm{T}}$ 是 $((2-i)\boldsymbol{E}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{\theta}$ 的一个非零解.

又易知 $\mathrm{r}((2-i)\boldsymbol{E}-\boldsymbol{A})=2$, 于是 $((2-i)\boldsymbol{E}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{\theta}$ 的基础解系只有一个解, 故 $\overline{\boldsymbol{\alpha}}_2=(1,1+i,1)^{\mathrm{T}}$ 是 $((2-i)\boldsymbol{E}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{\theta}$ 的一个基础解系. 从而属于特征值 $2-i$ 的全部特征向量为: $k_3\overline{\boldsymbol{\alpha}}_2$, 其中 k_3 为任意非零复常数. □

我们有时候需要利用矩阵的运算来求特征值和特征向量，如下例.

例 9. 求矩阵 $E+xy^T$ 的特征值与特征向量，其中 E 为 n 阶单位阵，
$x=(x_1\ x_2\ \cdots\ x_n)^T$, $y=(y_1\ y_2\ \cdots\ y_n)^T$, $n\geq 2$.

解. 我们只需考虑 xy^T 的特征值与特征向量。

当 $x=0$ 或 $y=0$ 时，$E+xy^T=E$. 此时有 n 重特征值 1，
任意 n-维非零向量均是属于特征值 1 的特征向量。

现在假设 $x\neq 0$ 且 $y\neq 0$. 此时 $xy^T\neq 0$, 从而 $r\left(xy^T\right)\geq 1$.
又因为 $r\left(xy^T\right)\leq\min\left\{r(x),r\left(y^T\right)\right\}=1$, 故有 $r\left(xy^T\right)=1$.
由于 $n\geq 2$, 故 $\left|xy^T\right|=0$, 从而 xy^T 有特征值 0.
因为 $r\left(0E-xy^T\right)=r\left(xy^T\right)=1$,
故 xy^T 属于特征值 0 的线性无关的特征向量有 $n-1$ 个，
设为 $\xi_1,\xi_2,\ldots,\xi_{n-1}$. 这表明 0 至少是 xy^T 的 $n-1$ 重根。

若 λ 是 xy^T 不等于 0 的根，则由

$$tr\left(xy^T\right)=x_1y_1+x_2y_2+\cdots+x_ny_n=y^Tx=\lambda+(n-1)\cdot 0$$

可得

$$\lambda=y^Tx.$$

由于 $\left(xy^T\right)x=x\left(y^Tx\right)=\left(y^Tx\right)x=\lambda x$, 且 $x\neq 0$,
故属于 $\lambda=y^Tx$ 的所有特征向量是 kx，k为任意非零常数。

xy^T 的特征值	对应的特征向量	$E+xy^T$的特征值
0 (n 重)	$k_1\xi_1+k_2\xi_2+\cdots+k_{n-1}\xi_{n-1}$	1 (n 重)
y^Tx	kx	$1+y^Tx$
0 ($n-1$重)	$k_1\xi_1+k_2\xi_2+\cdots+k_{n-1}\xi_{n-1}$	1 ($n-1$ 重)

当 $\boldsymbol{x}\neq\boldsymbol{\theta},\boldsymbol{y}\neq\boldsymbol{\theta}$ 且 $\boldsymbol{y}^{\mathrm{T}}\boldsymbol{x}\neq 0$ 时，特征值为 $\lambda=1+\boldsymbol{y}^{\mathrm{T}}\boldsymbol{x}$(单重) 和 $\lambda=1$ ($n-1$ 重)，其中属于 $\lambda=1+\boldsymbol{y}^{\mathrm{T}}\boldsymbol{x}$ 的特征向量为 $k\boldsymbol{x}$，$k\neq 0$，而属于 $\lambda=1$ 的特征向量为 $k_1\boldsymbol{\xi}_1+\cdots+k_{n-1}\boldsymbol{\xi}_{n-1}$，其中 $\boldsymbol{\xi}_1,\cdots,\boldsymbol{\xi}_{n-1}$ 为 $\boldsymbol{x}\boldsymbol{y}^{\mathrm{T}}\boldsymbol{\xi}=\boldsymbol{\theta}$ 的基础解系，$k_1,\cdots,k_{n-1}$ 不全为零. 当 $\boldsymbol{x}\neq\boldsymbol{\theta},\boldsymbol{y}\neq\boldsymbol{\theta}$ 且 $\boldsymbol{y}^{\mathrm{T}}\boldsymbol{x}=0$ 时，特征值为 $\lambda=1$ (n 重)，对应的特征向量为 $k_1\boldsymbol{\xi}_1+\cdots+k_{n-1}\boldsymbol{\xi}_{n-1}$，$k_1,\cdots,k_{n-1}$ 不全为零. □

下面我们来看一看当两个矩阵之间有某种关系时，它们的特征值和特征向量有什么关系.

例 4.2.5 设矩阵

$$\boldsymbol{A}=\begin{pmatrix}1&1&1\\-2&1&2\\2&1&0\end{pmatrix}.$$

求 $\boldsymbol{A}$ 的全部特征值和 $\boldsymbol{B}=5\boldsymbol{A}$ 的全部特征值.

解 由

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

得 $\boldsymbol{A}$ 的特征值为 $1,-1,2$.

$$\boldsymbol{B}=5\boldsymbol{A}=\begin{pmatrix}5&5&5\\-10&5&10\\10&5&0\end{pmatrix}，由$$

$$|\lambda\boldsymbol{E}-\boldsymbol{B}|=\begin{vmatrix}\lambda-5&-5&-5\\10&\lambda-5&-10\\-10&-5&\lambda\end{vmatrix}=\begin{vmatrix}\lambda-5&-5&-5\\0&\lambda-10&\lambda-10\\-10&-5&\lambda\end{vmatrix}=(\lambda^2-25)(\lambda-10)$$

得 $\boldsymbol{B}$ 的特征值为 $5,-5,10$. □

我们看到，上述例子中 $\boldsymbol{B}$ 的特征值与 $\boldsymbol{A}$ 的特征值的运算关系，正好是矩阵的运算关系. 这种关系不是偶然的，请看下面的定理.

矩阵的运算与矩阵的特征值及特征向量 定理1.

定理 4.2.2 *若 $f(x)$ 为 x 的多项式，矩阵 $\boldsymbol{A}$ 有特征值 λ，则 $f(\boldsymbol{A})$ 有特征值 $f(\lambda)$.*

证明 设 $f(x)=a_mx^m+a_{m-1}x^{m-1}+\cdots+a_1x+a_0$，$\boldsymbol{\xi}$ 为 $\boldsymbol{A}$ 的属于特征值 λ 的特征向量，则有 $\boldsymbol{A\xi}=\lambda\boldsymbol{\xi}$，于是有

$$\begin{aligned}f(\boldsymbol{A})\boldsymbol{\xi}&=(a_m\boldsymbol{A}^m+a_{m-1}\boldsymbol{A}^{m-1}+\cdots+a_1\boldsymbol{A}+a_0\boldsymbol{E})\boldsymbol{\xi}\\&=a_m\boldsymbol{A}^m\boldsymbol{\xi}+a_{m-1}\boldsymbol{A}^{m-1}\boldsymbol{\xi}+\cdots+a_1\boldsymbol{A\xi}+a_0\boldsymbol{\xi}\\&=a_m\lambda^m\boldsymbol{\xi}+a_{m-1}\lambda^{m-1}\boldsymbol{\xi}+\cdots+a_1\lambda\boldsymbol{\xi}+a_0\boldsymbol{\xi}\\&=(a_m\lambda^m+a_{m-1}\lambda^{m-1}+\cdots+a_1\lambda+a_0)\boldsymbol{\xi}\\&=f(\lambda)\boldsymbol{\xi}.\end{aligned}$$

故 $f(\lambda)$ 为 $f(\boldsymbol{A})$ 的特征值，且 $\boldsymbol{\xi}$ 也是 $f(\boldsymbol{A})$ 的属于 $f(\lambda)$ 的特征向量. □

特别的，若 n 阶方阵 A 的特征值是：λ_1, λ_2, …, λ_n,
那么 $kE+A$ 的特征值就是：$k+\lambda_1, k+\lambda_2, \dots, k+\lambda_n$.

(重要)

注 1　若定理 4.2.2 中矩阵 $\boldsymbol{A}$ 的所有特征值为 $\lambda_1, \cdots, \lambda_n$ (包括相同的特征值)，则 $f(\boldsymbol{A})$ 的所有特征值为 $f(\lambda_1), \cdots, f(\lambda_n)$. 结论的证明参见后面若尔当标准形和奇异值分解一节.

定理2. **注 2**　若 n 阶可逆方阵 $\boldsymbol{A}$ 的所有特征值为 $\lambda_1, \cdots, \lambda_n$ (包括相同的特征值)，则 $\lambda_i \neq 0$, $i = 1, 2, \cdots, n$，且矩阵 $\boldsymbol{A}^{-1}$ 的所有特征值为 $\lambda_1^{-1}, \cdots, \lambda_n^{-1}$. $\frac{1}{\lambda_1}, \cdots, \frac{1}{\lambda_n}$.

事实上，若设 $\boldsymbol{\xi}_i$ 为 $\boldsymbol{A}$ 的属于特征值 λ_i 的特征向量，则有

$$\boldsymbol{A}\boldsymbol{\xi}_i = \lambda_i \boldsymbol{\xi}_i, \quad i = 1, 2, \cdots, n.$$

因为 $\boldsymbol{A}$ 可逆，其特征向量 $\boldsymbol{\xi}_i \neq \boldsymbol{\theta}$, $i = 1, 2, \cdots, n$，故有 $\boldsymbol{\xi}_i = \lambda_i \boldsymbol{A}^{-1}\boldsymbol{\xi}_i$，从而 $\lambda_i \neq 0$ 且 $\boldsymbol{A}^{-1}\boldsymbol{\xi}_i = \lambda_i^{-1}\boldsymbol{\xi}_i$，此即 $\boldsymbol{A}^{-1}$ 的特征值为 λ_i^{-1}.

注 3　不可逆方阵 $\boldsymbol{A}$ 必有 0 特征值.

例 4.2.6　已知矩阵

$$\boldsymbol{A} = \begin{pmatrix} 1 & 1 & 1 \\ -2 & 1 & 2 \\ 2 & 1 & 0 \end{pmatrix}$$

的特征值为 $1, -1, 2$, 求 $\boldsymbol{B} = \boldsymbol{A}^2 + 2\boldsymbol{A} + \boldsymbol{E}$ 和 $\boldsymbol{C} = \boldsymbol{A}^2$ 的全部特征值.

解　显然 $1, -1, 2$ 是 $\boldsymbol{A}$ 的全部特征值，由定理 4.2.2 知，$\boldsymbol{B} = f(\boldsymbol{A}) = \boldsymbol{A}^2 + 2\boldsymbol{A} + \boldsymbol{E}$ 有特征值 $f(\lambda) = \lambda^2 + 2\lambda + 1$，因为 $f(1) = 4, f(-1) = 0, f(2) = 9$，故 $4, 0, 9$ 为 $\boldsymbol{B}$ 的特征值, 且是 $\boldsymbol{B}$ 的全部特征值.

同样，$1^2, (-1)^2, 2^2$，即 1(二重) 和 4 也是 $\boldsymbol{C}$ 的全部特征值. □

利用矩阵之间的关系和其中一个矩阵的特征值去计算另一个矩阵的特征值是计算矩阵特征值的又一个重要方法. 为此，要继续讨论矩阵特征值的一些共有性质.

定理3.

矩阵的相似不变量

定理 3. 设矩阵 A 相似矩阵 B, $B = P^{-1}AP$. 则 A 与 B 具有相同的特征多项式以及特征值。

证明. 只需证明 A 与 B 具有相同的特征多项式。

$$|\lambda E - B| = |\lambda P^{-1}EP - P^{-1}AP| = |P^{-1}(\lambda E)P - P^{-1}AP|$$
$$= |P^{-1}(\lambda E - A)P| = |P^{-1}|\,|\lambda E - A|\,|P| = |\lambda E - A|\,.$$

即 □

$$\boldsymbol{E} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \boldsymbol{B} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix},$$

两者的特征多项式都是 $(\lambda - 1)^2$, 但对于任何二阶可逆矩阵 $\boldsymbol{P}$, 都成立 $\boldsymbol{P}^{-1}\boldsymbol{E}\boldsymbol{P} = \boldsymbol{E} \neq \boldsymbol{B}$, 可知 $\boldsymbol{E}$ 与 $\boldsymbol{B}$ 并不相似.

定义 4.2.3 (迹)　定义

$$\mathrm{tr}\,(\boldsymbol{A}) = \sum_{i=1}^{n} a_{ii}$$

主对角线元素之和

为矩阵 $\boldsymbol{A} = (a_{ij})_{n \times n}$ 的迹.

定理4.

定理 4.2.4 若 n 阶矩阵 $\boldsymbol{A}$ 的特征值为 $\lambda_1, \cdots, \lambda_n$，则有

$$\operatorname{tr}(\boldsymbol{A}) = \sum_{i=1}^{n} \lambda_i, \qquad |\boldsymbol{A}| = \prod_{i=1}^{n} \lambda_i.$$

证明 因为

$$\begin{aligned} |\lambda \boldsymbol{E} - \boldsymbol{A}| &= (\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_n) \\ &= \lambda^n - (\lambda_1 + \lambda_2 + \cdots + \lambda_n)\lambda^{n-1} + \cdots + (-1)^n \prod_{i=1}^{n} \lambda_i, \end{aligned}$$

取 $\lambda = 0$ 可得 $|-\boldsymbol{A}| = (-1)^n \prod_{i=1}^{n} \lambda_i$，即 $|\boldsymbol{A}| = \prod_{i=1}^{n} \lambda_i$.

进一步比较上式两边 λ 的 $n-1$ 次项系数，由于

$$|\lambda \boldsymbol{E} - \boldsymbol{A}| = \begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{vmatrix},$$

将行列式按第一行展开，可知只有第一项

$$(\lambda - a_{11}) \begin{vmatrix} \lambda - a_{22} & -a_{23} & \cdots & -a_{2n} \\ -a_{32} & \lambda - a_{33} & \cdots & -a_{3n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n2} & -a_{n3} & \cdots & \lambda - a_{nn} \end{vmatrix}$$

含 λ^{n-1} 的项. 继续展开可得只有 $(\lambda - a_{11})(\lambda - a_{22}) \cdots (\lambda - a_{nn})$ 才含 λ^{n-1} 的项，故知 $|\lambda \boldsymbol{E} - \boldsymbol{A}| = \lambda^n - (a_{11} + a_{22} + \cdots + a_{nn})\lambda^{n-1} + \cdots$，比较 λ^{n-1} 的系数可得 $\operatorname{tr}(\boldsymbol{A}) = \sum_{i=1}^{n} a_{ii} = \sum_{i=1}^{n} \lambda_i$.

$\underbrace{+(-1)^n|A|}$ 取 $\lambda=0$ 得常数项

$\square$

推论1. **推论 4.2.5** 相似矩阵有相同的迹和相同的行列式.

证明 由定理 4.2.3 知相似矩阵有相同的特征值，设为 $\lambda_1, \cdots, \lambda_n$，再由定理 4.2.4 知它们的迹和行列式分别为 $\sum_{i=1}^{n} \lambda_i$ 和 $\prod_{i=1}^{n} \lambda_i$. $\square$

下面再看一些例子.

例 4.2.7 设 $\begin{pmatrix} 3 & 2 & 2 \\ 2 & a & -1 \\ b & 3 & 1 \end{pmatrix}$ 与 $\begin{pmatrix} -2 & -8 & 6 \\ 10 & 12 & -3 \\ 5 & 10 & -5 \end{pmatrix}$ 相似，求 a, b 的值.

解 由于相似矩阵有相同的迹和行列式，故由 $3 + a + 1 = -2 + 12 + (-5)$ 可得 $a = 1$. 将 $a = 1$ 代入矩阵，再由行列式相等，得

$$\begin{vmatrix} 3 & 2 & 2 \\ 2 & 1 & -1 \\ b & 3 & 1 \end{vmatrix} = \begin{vmatrix} -2 & -8 & 6 \\ 10 & 12 & -3 \\ 5 & 10 & -5 \end{vmatrix} = 20,$$

即有 $4(5 - b) = 20$，解得 $b = 0$. 故有 $a = 1, b = 0$. $\square$

例 4. 设 A^*是 3 阶矩阵 A 的伴随矩阵，A^*的特征值为$-1,2,-2$，试求 $A+E$ 的特征值。

解. 根据定理 4 以及$|A^*| = |A|^{n-1}$可得

$$|A| = \pm\sqrt{(-1)\cdot 2\cdot(-2)} = \pm 2.$$

易见 A^* 可逆。根据定理 2，$(A^*)^{-1}$ 的特征值是:

$$-1, \frac{1}{2}, -\frac{1}{2}.$$

由于 $A = |A|(A^*)^{-1}$，故 A 的特征值是:

$$-2,1,-1 \quad 或 \quad 2,-1,1.$$

再根据 $kE+A$ 与 A 的特征值之间的关系知，$A+E$ 的特征值是:

$$-1, 2, 0 \quad 或 \quad 3, 0, 2.$$

例 4.2.9　设 $\boldsymbol{A}$ 为 3 阶矩阵，$\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\boldsymbol{\xi}_3$ 为 3 个线性无关的向量，且有关系: $\boldsymbol{A}\boldsymbol{\xi}_1 = -3\boldsymbol{\xi}_1 + 2\boldsymbol{\xi}_2 - \boldsymbol{\xi}_3$，$\boldsymbol{A}\boldsymbol{\xi}_2 = 6\boldsymbol{\xi}_1 + \boldsymbol{\xi}_2 + 2\boldsymbol{\xi}_3$，$\boldsymbol{A}\boldsymbol{\xi}_3 = \boldsymbol{\xi}_1 + \boldsymbol{\xi}_2 + 3\boldsymbol{\xi}_3$，求矩阵 $\boldsymbol{A}$ 的特征值与特征向量.

解　设 $\boldsymbol{P} = (\boldsymbol{\xi}_1\ \ \boldsymbol{\xi}_2\ \ \boldsymbol{\xi}_3)$，则有

$$\begin{aligned}\boldsymbol{AP} &\overset{=(A\xi_1\ A\xi_2\ A\xi_3)}{=} (-3\boldsymbol{\xi}_1 + 2\boldsymbol{\xi}_2 - \boldsymbol{\xi}_3\ \ 6\boldsymbol{\xi}_1 + \boldsymbol{\xi}_2 + 2\boldsymbol{\xi}_3\ \ \boldsymbol{\xi}_1 + \boldsymbol{\xi}_2 + 3\boldsymbol{\xi}_3) \\ &= (\boldsymbol{\xi}_1\ \ \boldsymbol{\xi}_2\ \ \boldsymbol{\xi}_3)\begin{pmatrix} -3 & 6 & 1 \\ 2 & 1 & 1 \\ -1 & 2 & 3 \end{pmatrix} = \boldsymbol{PB}.\end{aligned}$$

又因为 $\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\boldsymbol{\xi}_3$ 线性无关，故 $\boldsymbol{P}$ 可逆，于是有 $\boldsymbol{P}^{-1}\boldsymbol{AP} = \boldsymbol{B}$，即 $\boldsymbol{A} \sim \boldsymbol{B}$.

现在求 $\boldsymbol{B}$ 的特征值和特征向量. 由

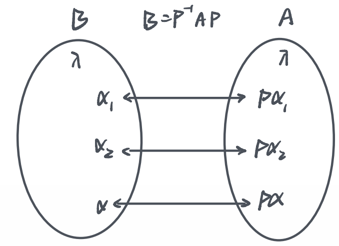
$$\begin{aligned}|\lambda\boldsymbol{E} - \boldsymbol{B}| &= \begin{vmatrix} \lambda+3 & -6 & -1 \\ -2 & \lambda-1 & -1 \\ 1 & -2 & \lambda-3 \end{vmatrix} = \begin{vmatrix} \lambda+3 & -6 & -1 \\ -\lambda-5 & \lambda+5 & 0 \\ 1 & -2 & \lambda-3 \end{vmatrix} \\ &= (\lambda+5)(\lambda-2)(\lambda-4) = 0\end{aligned}$$

解得特征值为：$\lambda = -5, 2, 4$.

对 $\lambda = -5$, 由 $\begin{pmatrix} -2 & -6 & -1 \\ -2 & -6 & -1 \\ 1 & -2 & -8 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & -4.6 \\ 0 & 1 & 1.7 \\ 0 & 0 & 0 \end{pmatrix}$ 得特征向量 $k_{11}\boldsymbol{\alpha}_{11} = \begin{pmatrix} 46 \\ -17 \\ 10 \end{pmatrix}$.

对 $\lambda = 2$，由 $\begin{pmatrix} 5 & -6 & -1 \\ -2 & 1 & -1 \\ 1 & -2 & -1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ 得特征向量 $k_{21}\boldsymbol{\alpha}_{21} = \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}$.

对 $\lambda = 4$，由 $\begin{pmatrix} 7 & -6 & -1 \\ -2 & 3 & -1 \\ 1 & -2 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}$ 得特征向量 $k_{31}\boldsymbol{\alpha}_{31} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.



$B\alpha = \lambda\alpha \Leftrightarrow PB\alpha = \lambda P\alpha \Leftrightarrow AP\alpha = \lambda P\alpha$

 $\Leftrightarrow A(P\alpha) = \lambda(P\alpha)$

因为由 $\boldsymbol{B\alpha} = \lambda\boldsymbol{\alpha}$ 可得 $\boldsymbol{AP\alpha} = \boldsymbol{PB\alpha} = \lambda\boldsymbol{P\alpha}$，故 $\boldsymbol{A}$ 有特征值 $\lambda = -5, 2, 4$，对应的特征向量为 $\boldsymbol{P\alpha}_1, \boldsymbol{P\alpha}_2, \boldsymbol{P\alpha}_3$. □

下面从分块矩阵的角度来考虑，也可得到一些有用的结论.

定理 4.2.6 设 $\boldsymbol{A}$ 是一个块对角矩阵

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{A}_1 & & & \\ & \boldsymbol{A}_2 & & \\ & & \ddots & \\ & & & \boldsymbol{A}_m \end{pmatrix}$$

则 $\boldsymbol{A}$ 的特征多项式是 $\boldsymbol{A}_1, \boldsymbol{A}_2, \cdots, \boldsymbol{A}_m$ 的特征多项式的乘积，于是 $\boldsymbol{A}_1, \boldsymbol{A}_2, \cdots, \boldsymbol{A}_m$ 的所有特征值就是 $\boldsymbol{A}$ 的所有特征值.

证明 将单位矩阵 $\boldsymbol{E}$ 按分块形式写成

$$\boldsymbol{E} = \begin{pmatrix} \boldsymbol{E}_1 & & & \\ & \boldsymbol{E}_2 & & \\ & & \ddots & \\ & & & \boldsymbol{E}_m \end{pmatrix}$$

则

$$\lambda\boldsymbol{E} - \boldsymbol{A} = \begin{pmatrix} \lambda\boldsymbol{E}_1 - \boldsymbol{A}_1 & & & \\ & \lambda\boldsymbol{E}_2 - \boldsymbol{A}_2 & & \\ & & \ddots & \\ & & & \lambda\boldsymbol{E}_m - \boldsymbol{A}_m \end{pmatrix}$$

因此

$$|\lambda\boldsymbol{E} - \boldsymbol{A}| = |\lambda\boldsymbol{E}_1 - \boldsymbol{A}_1||\lambda\boldsymbol{E}_2 - \boldsymbol{A}_2|\cdots|\lambda\boldsymbol{E}_m - \boldsymbol{A}_m|.$$

□

例 4.2.10 设 $\boldsymbol{A} \in \mathbf{R}^{m\times n}, \boldsymbol{B} \in \mathbf{R}^{n\times m}$, 证明

$$\lambda^n|\lambda\boldsymbol{E}_m - \boldsymbol{AB}| = \lambda^m|\lambda\boldsymbol{E}_n - \boldsymbol{BA}|.$$ 矩阵的特征多项式

$\begin{pmatrix} AB & 0 \\ B & 0 \end{pmatrix} = \left|\lambda E_{m+n} - \begin{pmatrix} AB & 0 \\ B & 0 \end{pmatrix}\right| = \left|\begin{pmatrix} \lambda E_m & 0 \\ 0 & \lambda E_n \end{pmatrix} - \begin{pmatrix} AB & 0 \\ B & 0 \end{pmatrix}\right| = \left|\begin{pmatrix} \lambda E_m - AB & 0 \\ -B & \lambda E_n \end{pmatrix}\right| = |\lambda E_n||\lambda E_m - AB| = \lambda^n|\lambda E_m - AB|$

证明 容易验证

$$\begin{pmatrix} \boldsymbol{E}_m & -\boldsymbol{A} \\ \boldsymbol{O} & \boldsymbol{E}_n \end{pmatrix}\begin{pmatrix} \boldsymbol{AB} & \boldsymbol{O} \\ \boldsymbol{B} & \boldsymbol{O} \end{pmatrix} = \begin{pmatrix} \boldsymbol{O} & \boldsymbol{O} \\ \boldsymbol{B} & \boldsymbol{BA} \end{pmatrix}\begin{pmatrix} \boldsymbol{E}_m & -\boldsymbol{A} \\ \boldsymbol{O} & \boldsymbol{E}_n \end{pmatrix}.$$

因 $\begin{pmatrix} \boldsymbol{E}_m & \boldsymbol{A} \\ \boldsymbol{O} & \boldsymbol{E}_n \end{pmatrix}$ 可逆, 由上式可知 $\begin{pmatrix} \boldsymbol{AB} & \boldsymbol{O} \\ \boldsymbol{B} & \boldsymbol{O} \end{pmatrix}$ 和 $\begin{pmatrix} \boldsymbol{O} & \boldsymbol{O} \\ \boldsymbol{B} & \boldsymbol{BA} \end{pmatrix}$ 是相似的, 从而有

$$\left|\lambda\boldsymbol{E}_{m+n} - \begin{pmatrix} \boldsymbol{AB} & \boldsymbol{O} \\ \boldsymbol{B} & \boldsymbol{O} \end{pmatrix}\right| = \left|\lambda\boldsymbol{E}_{m+n} - \begin{pmatrix} \boldsymbol{O} & \boldsymbol{O} \\ \boldsymbol{B} & \boldsymbol{BA} \end{pmatrix}\right|,$$

即

$$\left|\begin{pmatrix} \lambda \boldsymbol{E}_m - \boldsymbol{AB} & \boldsymbol{O} \\ -\boldsymbol{B} & \lambda \boldsymbol{E}_n \end{pmatrix}\right| = \left|\begin{pmatrix} \lambda \boldsymbol{E}_m & \boldsymbol{O} \\ -\boldsymbol{B} & \lambda \boldsymbol{E}_n - \boldsymbol{BA} \end{pmatrix}\right|.$$

故

$$\lambda^n|\lambda \boldsymbol{E}_m - \boldsymbol{AB}| = \lambda^m|\lambda \boldsymbol{E}_n - \boldsymbol{BA}|. \qquad \square$$

由例 4.2.10 可知, m 阶方阵 $\boldsymbol{AB}$ 与 n 阶方阵 $\boldsymbol{BA}$ 有相同的非零特征值, 且 $\mathrm{tr}(\boldsymbol{AB}) = \mathrm{tr}(\boldsymbol{BA})$.

4.3　矩阵可对角化的条件

下面进一步讨论什么样的矩阵可以相似于和怎样相似于对角矩阵的问题.

定义5、**定义 4.3.1** (可对角化)　若方阵 $\boldsymbol{A}$ 相似于一个对角矩阵，则称 $\boldsymbol{A}$ **可对角化**.

并非所有方阵都是可对角化，请看下面的例子.

例 4.3.1　说明矩阵

$$\boldsymbol{A} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$$

不可对角化.

解　由

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

可得特征值 $\lambda = 1$(二重)，再假设矩阵 $\boldsymbol{A}$ 可对角化，即存在可逆矩阵 $\boldsymbol{P}$，有 $\boldsymbol{P}^{-1}\boldsymbol{AP} = \mathrm{diag}\,(s_1, s_2)$，或者 $\boldsymbol{AP} = \boldsymbol{P}\cdot\mathrm{diag}\,(s_1, s_2)$. 因为相似矩阵有相同的特征值以及$\mathrm{diag}\,(s_1, s_2)$ 的特征值为 s_1, s_2, 所以 $s_1 = s_2 = 1$，于是 $\boldsymbol{P}^{-1}\boldsymbol{AP} = \boldsymbol{E}$. 从而得 $\boldsymbol{A} = \boldsymbol{E}$，与假设矛盾. 故 $\boldsymbol{A}$ 不可对角化. $\square$

既然有不可对角化的矩阵，那么到底什么样的矩阵是可对角化的？下面的定理给出了矩阵可对角化的一个充要条件，以及相应的相似变换矩阵和对角矩阵的构成特点.

定理6、**定理 4.3.1**　**n 阶矩阵可对角化的充要条件是有 n 个线性无关的特征向量；且对角矩阵的主对角线由特征值（可按任意次序）构成，相似变换矩阵由属于相应特征值的特征向量构成.**

相似变换矩阵有无穷多个
对角阵有 $n!$ 个

证明　设 n 阶矩阵为 $\boldsymbol{A}$，先证明必要性：

证明. (⇒) 若 A 可对角化, 则存在可逆阵 $P=(\xi_1\ \xi_2\ \cdots\ \xi_n)$ 使得

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix},$$

$$A\,(\xi_1\ \xi_2\ \cdots\ \xi_n) = (\xi_1\ \xi_2\ \cdots\ \xi_n)\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}.$$

$$(A\xi_1\ A\xi_2\ \cdots\ A\xi_n) = (\lambda_1\xi_1\ \lambda_2\xi_2\ \cdots\ \lambda_n\xi_n)$$

$$A\xi_i = \lambda_i\xi_i,\ 1 \le i \le n.$$

这表明相似变换矩阵 P 的 n 个列 $\xi_1, \xi_2, \ldots, \xi_n$ 是 A 的 n 个线性无关的特征向量。

(⇐) 设 A 有 n 个线性无关的特征向量 $\xi_1, \xi_2, \ldots, \xi_n$, $A\xi_i = \lambda_i\xi_i$, $1 \le i \le n$. 令 $P=(\xi_1\ \xi_2\ \cdots\ \xi_n)$，则 P 可逆，且有

$$A\left(\xi_1\ \xi_2\ \cdots\ \xi_n\right)=\left(A\xi_1\ A\xi_2\ \cdots\ A\xi_n\right),$$

$$\left(\xi_1\ \xi_2\ \cdots\ \xi_n\right)\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{pmatrix}=\left(\lambda_1\xi_1\ \lambda_2\xi_2\ \cdots\ \lambda_n\xi_n\right).$$

由于 $A\xi_i=\lambda_i\xi_i, 1\le i\le n$. 故有

$$A\left(\xi_1\ \xi_2\ \cdots\ \xi_n\right)=\left(\xi_1\ \xi_2\ \cdots\ \xi_n\right)\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{pmatrix},$$

显然 $s_1, s_2, \cdots$ …的特征向量，由 $\boldsymbol{P}$ 可逆

可得 n 个特征…

再证充分性…的特征值为 $s_1, s_2, \cdots, s_n$.

则 $\boldsymbol{P}=(\boldsymbol{\xi}_1\ \boldsymbol{\xi}_2$ …

$\boldsymbol{AP}=$ … $s_1, s_2, \cdots, s_n)$,

亦即

$$P^{-1}AP=\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{pmatrix}.$$

因此 A 可对角化。

故有 $\boldsymbol{P}^{-1}\boldsymbol{AP}=$ d…

进一步令 $\boldsymbol{C}=$ …为 $(1,\cdots,n)$ 的一个排列，则

$\boldsymbol{C}^{\mathrm{T}}\boldsymbol{C}=\boldsymbol{E}$，故 $\boldsymbol{C}^{\mathrm{T}}$ …

$\boldsymbol{D}^{-1}\boldsymbol{AD}=$ … $=\operatorname{diag}(s_{i_1}, s_{i_2}, \cdots, s_{i_n})$

即对角线的特征值… □

定理6. **定理 4.3.2　属于不同特征值的特征向量线性无关.**

证明. 令 $\lambda_1,\lambda_2,\ldots,\lambda_r$ 是 A 的两两不同的特征值，与之相应的特征向量是 $\xi_1,\xi_2,\ldots,\xi_r$. 设

…$,\cdots,\boldsymbol{\xi}_r$. 现在

当… $_2\boldsymbol{\xi}_2=\boldsymbol{\theta}$，即

$$k_1\xi_1+k_2\xi_2+\cdots+k_r\xi_r=0. \tag{1}$$

用矩阵 A 左乘以 (1) 式两边可得：

$$k_1\lambda_1\xi_1+k_2\lambda_2\xi_2+\cdots+k_r\lambda_r\xi_r=0. \tag{2}$$

在 (1) 式两边同时乘以 λ_r 可得：

$$k_1\lambda_r\xi_1+k_2\lambda_r\xi_2+\cdots+k_r\lambda_r\xi_r=0. \tag{3}$$

由 (3) − (2) 可得：

$$k_1(\lambda_r-\lambda_1)\xi_1+k_2(\lambda_r-\lambda_2)\xi_2+\cdots+k_{r-1}(\lambda_r-\lambda_{r-1})\xi_{r-1}=0. \tag{4}$$

上述第…为 $\boldsymbol{\xi}_1\ne\boldsymbol{\theta}, \lambda_1\ne\lambda_2$，故得到

$k_1=0$…

假… $_2\boldsymbol{\xi}_2+\cdots+k_m\boldsymbol{\xi}_m=\boldsymbol{\theta}$，则有

… $=\boldsymbol{\theta}$

现在对 r 归纳证明命题成立。

当 $r=2$ 时，由 (4) 式可得 $k_1(\lambda_2-\lambda_1)\xi_1=0$.

由于 $\lambda_2-\lambda_1\ne0$，$\xi_1\ne0$，故 $k_1=0$.

将 $k_1=0$ 代入 (1) 式可得 $k_2=0$，即 ξ_1,ξ_2 线性无关。

假设命题对 $r-1$ 成立。

则因为 $\lambda_r-\lambda_i\ne0$，$\xi_i\ne0$，$1\le i\le r-1$，由归纳假定及 (4) 式可得

$$k_1=\cdots=k_{r-1}=0.$$

再将 $k_1=\cdots=k_{r-1}=0$ 代入 (1) 式可得 $k_r=0$. 故 $\xi_1,\xi_2,\ldots,\xi_r$ 线性无关。

第一式… $\boldsymbol{\xi}_{m-1}=\boldsymbol{\theta}$.

由假设…因为 $\lambda_1,\lambda_2,\cdots,\lambda_m$ 互

不相同… $_m$ 线性无关. 故 $r=m$

时结论… □

从…条件.

推论2. **推论 4.3.3　若 n 阶矩阵有 n 个互不相同的特征值，则矩阵可对角化.**

例 4.3.2　将矩阵

$$\boldsymbol{A}=\begin{pmatrix}1 & 1 & 1\\ -2 & 1 & 2\\ 2 & 1 & 0\end{pmatrix}$$

对角化.

解　在例 4.2.5 中求得 $\boldsymbol{A}$ 的特征值为: $\lambda = 1, -1, 2$.

对于 $\lambda = 1$，解齐次方程组 $(\boldsymbol{E} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{\theta}$, 由

$$\begin{pmatrix} 0 & -1 & -1 \\ 2 & 0 & -2 \\ -2 & -1 & 1 \end{pmatrix} \xrightarrow[]{\substack{r_3+r_2 \\ \frac{1}{2}r_2}} \begin{pmatrix} 0 & -1 & -1 \\ 1 & 0 & -1 \\ 0 & -1 & -1 \end{pmatrix} \xrightarrow[]{\substack{r_3-r_1 \\ (-1)r_1 \\ r_1 \leftrightarrow r_2}} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

求得该齐次方程组的一个基础解系为: $\boldsymbol{\alpha}_1 = (1, -1, 1)^{\mathrm{T}}$.

对于 $\lambda = -1$，解齐次方程组 $(-\boldsymbol{E} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{\theta}$, 由

$$\begin{pmatrix} -2 & -1 & -1 \\ 2 & -2 & -2 \\ -2 & -1 & -1 \end{pmatrix} \xrightarrow[]{\substack{r_1+r_2 \\ r_3+r_2 \\ \frac{1}{2}r_2}} \begin{pmatrix} 0 & -3 & -3 \\ 1 & -1 & -1 \\ 0 & -3 & -3 \end{pmatrix} \xrightarrow[\substack{(-\frac{1}{3})r_1 \\ r_1 \leftrightarrow r_2}]{\substack{r_3-r_1 \\ r_2-\frac{1}{3}r_1}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

求得该齐次方程组的一个基础解系为: $\boldsymbol{\alpha}_2 = (0, -1, 1)^{\mathrm{T}}$.

对于 $\lambda = 2$，解齐次方程组 $(2\boldsymbol{E} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{\theta}$, 由

$$\begin{pmatrix} 1 & -1 & -1 \\ 2 & 1 & -2 \\ -2 & -1 & 2 \end{pmatrix} \xrightarrow[]{\substack{r_2-2r_1 \\ r_3+2r_1}} \begin{pmatrix} 1 & -1 & -1 \\ 0 & 3 & 0 \\ 0 & -3 & 0 \end{pmatrix} \xrightarrow[]{\substack{r_3+r_2 \\ r_1+\frac{1}{3}r_2 \\ \frac{1}{3}r_2}} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

求得该齐次方程组的一个基础解系为: $\boldsymbol{\alpha}_3 = (1, 0, 1)^{\mathrm{T}}$.

令 $\boldsymbol{P} = \begin{pmatrix} 1 & 0 & 1 \\ -1 & -1 & 0 \\ 1 & 1 & 1 \end{pmatrix}$，则有 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \operatorname{diag}(1, -1, 2)$. □

具有重特征值的矩阵的对角化问题要复杂一些，下面来讨论这种情况.

定理7.

定理 4.3.4　若 $\lambda_1, \lambda_2, \cdots, \lambda_m$ 是矩阵 $\boldsymbol{A}$ 的不同特征值, 而 $\boldsymbol{A}$ 的属于 λ_i 的线性无关的特征向量为 $\boldsymbol{\alpha}_{i1}, \boldsymbol{\alpha}_{i2}, \cdots, \boldsymbol{\alpha}_{is_i} (i = 1, 2, \cdots, m)$, 则向量组

$$\boldsymbol{\alpha}_{11}, \cdots, \boldsymbol{\alpha}_{1s_1}, \boldsymbol{\alpha}_{21}, \cdots, \boldsymbol{\alpha}_{2s_2}, \cdots, \boldsymbol{\alpha}_{m1}, \cdots, \boldsymbol{\alpha}_{ms_m}$$

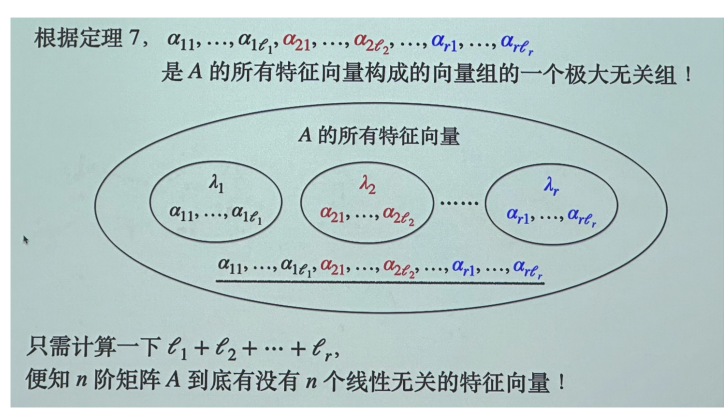
线性无关.

证明　考虑如下式子

$$\underline{k_{11}\boldsymbol{\alpha}_{11} + \cdots + k_{1s_1}\boldsymbol{\alpha}_{1s_1}} + \underline{k_{21}\boldsymbol{\alpha}_{21} + \cdots + k_{2s_2}\boldsymbol{\alpha}_{2s_2}} + \cdots + \underline{k_{m1}\boldsymbol{\alpha}_{m1} + \cdots + k_{ms_m}\boldsymbol{\alpha}_{ms_m}} = \boldsymbol{\theta}.$$

设

$$\begin{cases} \boldsymbol{\beta}_1 = k_{11}\boldsymbol{\alpha}_{11} + \cdots + k_{1s_1}\boldsymbol{\alpha}_{1s_1}, \\ \boldsymbol{\beta}_2 = k_{21}\boldsymbol{\alpha}_{21} + \cdots + k_{2s_2}\boldsymbol{\alpha}_{2s_2}, \\ \quad \cdots\cdots \\ \boldsymbol{\beta}_m = k_{m1}\boldsymbol{\alpha}_{m1} + \cdots + k_{ms_m}\boldsymbol{\alpha}_{ms_m}, \end{cases}$$



则有

$$\boldsymbol{\beta}_1 + \boldsymbol{\beta}_2 + \cdots + \boldsymbol{\beta}_m = \boldsymbol{\theta},$$

且由定理 4.2.1 知 $A\beta_i = A(k_{i1}\alpha_{i1} + \cdots + k_{il_i}\alpha_{il_i}) = \lambda_i(k_{i1}\alpha_{i1} + \cdots + k_{il_i}\alpha_{il_i}) = \lambda_i\beta_i$

$$\boldsymbol{A}\boldsymbol{\beta}_1 = \lambda_1\boldsymbol{\beta}_1,\ \boldsymbol{A}\boldsymbol{\beta}_2 = \lambda_2\boldsymbol{\beta}_2,\ \cdots,\ \boldsymbol{A}\boldsymbol{\beta}_m = \lambda_m\boldsymbol{\beta}_m.$$

若 $\boldsymbol{\beta}_{i_1}, \cdots, \boldsymbol{\beta}_{i_r}$ 为非零向量，其余为零向量，则有

$$\boldsymbol{\beta}_{i_1} + \cdots + \boldsymbol{\beta}_{i_r} = \boldsymbol{\theta},$$

即这些向量线性相关. 但由定理 4.3.2 可知 $\boldsymbol{\beta}_{i_1}, \cdots, \boldsymbol{\beta}_{i_r}$ 线性无关，产生矛盾. 故必有 $\boldsymbol{\beta}_1 = \boldsymbol{\beta}_2 = \cdots = \boldsymbol{\beta}_m = \boldsymbol{\theta}$.

$\beta_{i_1} \cdots \beta_{i_r}$ 是 A 的属于不同特征值的特征向量.

对每个 $i \in \{1, 2, \cdots, m\}$，由于 $\boldsymbol{\alpha}_{i1}, \boldsymbol{\alpha}_{i2}, \cdots, \boldsymbol{\alpha}_{is_i}$ 是线性无关的特征向量，有

$$\boldsymbol{\beta}_i = k_{i1}\boldsymbol{\alpha}_{i1} + \cdots + k_{is_i}\boldsymbol{\alpha}_{is_i} = \boldsymbol{\theta},$$

可得

$$k_{i1} = k_{i2} = \cdots = k_{is_i} = 0,\ i = 1, 2, \cdots, m.$$

这就证得向量组 $\boldsymbol{\alpha}_{11}, \cdots, \boldsymbol{\alpha}_{1s_1}, \boldsymbol{\alpha}_{21}, \cdots, \boldsymbol{\alpha}_{2s_2}, \cdots, \boldsymbol{\alpha}_{m1}, \cdots, \boldsymbol{\alpha}_{ms_m}$ 线性无关. □

定理8. **定理 4.3.5** 设 λ_0 是 n 阶方阵 $\boldsymbol{A}$ 的 k 重特征值，则 $\boldsymbol{A}$ 的属于特征值 λ_0 的线性无关的特征向量个数不超过 k.

$s_1 + s_2 + \cdots + s_r = n$
$l_1 + l_2 + \cdots + l_r = n.$
那么 $l_i = s_i$ 吗?

证明. 设 $\xi_1, \xi_2, \ldots, \xi_\ell$ 是 A 属于 λ_0 的特征向量的极大无关组。

由于 $r\{\xi_1, \ldots, \xi_\ell, e_1, \ldots, e_n\} = n$，故 $\xi_1, \xi_2, \ldots, \xi_\ell, e_1, e_2, \ldots, e_n$ 中包含 $\xi_1, \xi_2, \ldots, \xi_\ell$ 的极大的线性无关组必定包含 n 个向量，记为

$$\xi_1, \ldots, \xi_\ell, \xi_{\ell+1}, \ldots, \xi_n,$$

其中 $\xi_{\ell+1}, \ldots, \xi_n$ 取自单位向量 $e_1, e_2, \ldots, e_n$.

$\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$

关时，删去该

则 $\mathrm{r}(\boldsymbol{P}) = n$,

令 $P = (\xi_1\ \cdots\ \xi_\ell\ \xi_{\ell+1}\ \cdots\ \xi_n)$，则 P 可逆。于是

$$\begin{aligned} B = P^{-1}AP &= P^{-1}A(\xi_1\ \cdots\ \xi_\ell\ \ \xi_{\ell+1}\ \cdots\ \xi_n) \\ &= P^{-1}(A\xi_1\ \cdots\ A\xi_\ell\ \ A\xi_{\ell+1}\ \cdots\ A\xi_n) \\ &= P^{-1}(\lambda_0\xi_1\ \cdots\ \lambda_0\xi_\ell\ \ \alpha_{\ell+1}\ \cdots\ \alpha_n) \\ &= (\lambda_0\underline{P^{-1}\xi_1}\ \cdots\ \underline{\lambda_0P^{-1}\xi_\ell}\ \ \underline{P^{-1}\alpha_{\ell+1}}\ \cdots\ \underline{P^{-1}\alpha_n}) \end{aligned}$$

对任意 ξ_i，由于

$$Pe_i = (\xi_1\ \cdots\ \xi_i\ \cdots\ \xi_n)\begin{pmatrix} 0 \\ \cdots \\ 1 \\ \cdots \\ 0 \end{pmatrix} = \xi_i,$$

故有 $P^{-1}\xi_i = e_i$. 于是

$$B = (\lambda_0 e_1\ \cdots\ \lambda_0 e_\ell\ \beta_{\ell+1}\ \cdots\ \beta_n) = \begin{pmatrix} \lambda_0 E_\ell & B_{12} \\ O & B_{22} \end{pmatrix}.$$

从而,

$$|\lambda E - A| = |\lambda E - B| = (\lambda - \lambda_0)^\ell\,|\lambda E_{n-\ell} - B_{22}|.$$

定理 9. n 阶矩阵 A 可对角化的充要条件是，对任意 s_i 重根 λ_i，

$$r(\lambda_i E - A) = n - s_i.$$

定理9. **定理 4.3.6** n 阶方阵 $\boldsymbol{A}$ 可对角化的充要条件是每个 k_i 重特征值 λ_i 对应的特征矩阵 $\lambda_i\boldsymbol{E} - \boldsymbol{A}$ 的秩为 $n - k_i$.

证明 设 $\lambda_1, \lambda_2, \cdots, \lambda_s$ 为 $\boldsymbol{A}$ 的全部不同特征值，注意到 $(\lambda_i\boldsymbol{E} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{\theta}$ 的基础解系包含 $n - (n - k_i)$，即 k_i 个线性无关的特征向量，而 k_i 之和是 n，由定理 4.3.4 可知，$\boldsymbol{A}$ 有 n 个线性无关的特征向量，则 $\boldsymbol{A}$ 与对角矩阵相似. 反之，若 $\boldsymbol{A}$ 相似于对角矩阵，则 $\boldsymbol{A}$ 有 n 个线性无关的特征向量，而 $k_i \geqslant 1$ 且 $\sum\limits_{i=1}^{s} k_i = n$，由定理 4.3.5，$\boldsymbol{A}$ 的属于 λ_i 的特征向量的极

大无关组恰好有 k_i 个特征向量. 这个极大无关组又是 $(\lambda_i \boldsymbol{E}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{\theta}$ 的基础解系，从而 $\lambda_i \boldsymbol{E}-\boldsymbol{A}$ 的秩为 $n-k_i$. □

由此, 我们得到在 $\boldsymbol{A}$ 有重特征值时 $\boldsymbol{A}$ 可否对角化的判定方法:

对每个重特征值 λ_i, 求矩阵 $\lambda_i \boldsymbol{E}-\boldsymbol{A}$ 的秩 r_i, 若对每个重特征值 $\lambda_i, n-r_i$ 都等于 λ_i 的重数, 则 $\boldsymbol{A}$ 可以对角化; 否则 $\boldsymbol{A}$ 不可对角化.

我们将矩阵 $\boldsymbol{A}$ 对角化的过程按步骤描述如下.

(1) 解特征方程 $|\lambda \boldsymbol{E}-\boldsymbol{A}|=0$ 得特征值 $\lambda=\lambda_1(s_1$ 重$),\cdots,\lambda_m(s_m$ 重$)$;

(2) 对每个特征值 $\lambda_i, i=1,2,\cdots,m$, 解齐次方程组

$$(\lambda_i \boldsymbol{E}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{\theta},$$

得一个基础解系 $\boldsymbol{\alpha}_{i1},\cdots,\boldsymbol{\alpha}_{ir_i}$.

若有某个 i 使得 $r_i<s_i$, 则矩阵 $\boldsymbol{A}$ 不可对角化;

(3) 当所有的 $r_i=s_i, i=1,2,\cdots,m$，则令

$$\boldsymbol{P}=(\boldsymbol{\alpha}_{11}\ \cdots\ \boldsymbol{\alpha}_{1s_1}\ \boldsymbol{\alpha}_{21}\ \cdots\ \boldsymbol{\alpha}_{2s_2}\ \cdots\ \boldsymbol{\alpha}_{m1}\ \cdots\ \boldsymbol{\alpha}_{ms_m}),$$

即得 $$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{\Lambda}=\operatorname{diag}(\underbrace{\lambda_1,\cdots,\lambda_1}_{s_1\text{ 个}},\underbrace{\lambda_2,\cdots,\lambda_2}_{s_2\text{ 个}},\cdots,\underbrace{\lambda_m,\cdots,\lambda_m}_{s_m\text{ 个}}).$$

由于每个齐次线性方程组的基础解系都不是唯一的, 因而 $\boldsymbol{P}$ 的取法也不是唯一的. 但是, 由于 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ 的主对角线上的元素是 $\boldsymbol{A}$ 的全体特征值, 因此除了主对角线上元素的次序不同外, $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ 是唯一确定的.

例 4.3.3　问

$$\boldsymbol{A}=\begin{pmatrix}2&-1&2\\5&-3&3\\-1&0&-2\end{pmatrix}$$

是否可对角化, 为什么?

解　由

$$|\lambda \boldsymbol{E}-\boldsymbol{A}|=\begin{vmatrix}\lambda-2&1&-2\\-5&\lambda+3&-3\\1&0&\lambda+2\end{vmatrix}=(\lambda+1)^3=0$$

得三重特征值 $\lambda=-1$, 显然 $-\boldsymbol{E}-\boldsymbol{A}\neq\boldsymbol{O}$, 因而其秩 $r\geqslant 1$,

$$n-r=3-r\leqslant 2<3,$$

故 $\boldsymbol{A}$ 不可能与对角矩阵相似. □

例 4.3.4　将上节例 4.2.1 的矩阵

$$\boldsymbol{A}=\begin{pmatrix}5&-2&1\\0&4&0\\1&-2&5\end{pmatrix}$$

对角化.

例 8. 证明：若方阵 A 满足 $A^2=E$，则 A 可对角化。

证明. 由 $A^2=E$ 可得：$(E-A)(-E-A)=(-E-A)(E-A)=0.$

若 $A=E$ 或 $-E$，则显然 A 可对角化。
若 $A\neq E,-E$，则 $-E-A\neq 0$，$E-A\neq 0$.
这样由上式可知，1 和 -1都是 A 的特征值。
因为 $-E-A$ 的非零列是 $(E-A)x=0$ 的非零解，
故 A至少有 $r(-E-A)$个属于特征值 1的线性无关的特征向量。
同理，A至少有 $r(E-A)$ 个属于特征值 -1的线性无关的特征向量。由于

秩的和≥和的秩

$$r(-E-A)+r(E-A)=\underline{r(E+A)+r(E-A)\geq r(2E)}=n,$$

故 A 有 n 个线性无关的特征向量，从而可以对角化。

可得 $\boldsymbol{A}$ 有 n 个线性无关的特征向量，可以对角化. □

例19、

4.4 正交矩阵与施密特正交化方法

为了进一步讨论矩阵对角化和矩阵的分解以及后续内容的需要，我们要引入正交矩阵，首先引入向量的内积与正交的有关概念. 本节内容除了定理 4.4.4 外，其余均在实数范围内考虑.

定义6、**定义 4.4.1** (向量内积) 设 $\boldsymbol{\alpha}$、$\boldsymbol{\beta}$ 为 n 维向量，用列向量表示为 $\boldsymbol{\alpha}=(a_1,a_2,\cdots,a_n)^{\mathrm{T}}$，$\boldsymbol{\beta}=(b_1,b_2,\cdots,b_n)^{\mathrm{T}}$. 若 $\boldsymbol{\alpha}$、$\boldsymbol{\beta}$ 为实向量，则称 $a_1b_1+a_2b_2+\cdots+a_nb_n$ 为 $\boldsymbol{\alpha}$、$\boldsymbol{\beta}$ 的**实内积**；若 $\boldsymbol{\alpha}$、$\boldsymbol{\beta}$ 为复向量，则称 $a_1\bar{b}_1+a_2\bar{b}_2+\cdots+a_n\bar{b}_n$ 为 $\boldsymbol{\alpha}$、$\boldsymbol{\beta}$ 的**复内积**. 统称为**向量的内积**，记为 $(\boldsymbol{\alpha},\boldsymbol{\beta})$，并称 $\|\boldsymbol{\alpha}\|=\sqrt{(\boldsymbol{\alpha},\boldsymbol{\alpha})}$ 为向量 $\boldsymbol{\alpha}$ 的**长度或模**；称模为 1 的向量为**单位向量**.

实内积无序，复内积有序

显然，n 维实向量的内积和长度就是 3 维向量的点乘和长度的推广. 借用矩阵的乘法，实向量 $\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}$ 的内积还可以表示为 $(\boldsymbol{\alpha},\boldsymbol{\beta})=\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{\beta}$. 复向量 $\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}$ 的内积还可以表示为 $(\boldsymbol{\alpha},\boldsymbol{\beta})=\boldsymbol{\alpha}^{\mathrm{T}}\overline{\boldsymbol{\beta}}$.

由定义 4.4.1 容易得到实内积的如下性质：(内积公理)

(1) $(\boldsymbol{\alpha},\boldsymbol{\beta})=(\boldsymbol{\beta},\boldsymbol{\alpha})$;

(2) $(k\boldsymbol{\alpha},\boldsymbol{\beta})=k(\boldsymbol{\alpha},\boldsymbol{\beta})$; $\forall k\in\mathbb{R}$.

(3) $(\boldsymbol{\alpha}+\boldsymbol{\gamma},\boldsymbol{\beta})=(\boldsymbol{\alpha},\boldsymbol{\beta})+(\boldsymbol{\gamma},\boldsymbol{\beta})$;

(4) $(\boldsymbol{\alpha},\boldsymbol{\alpha})\geqslant 0$；$(\boldsymbol{\alpha},\boldsymbol{\alpha})=0$ 当且仅当 $\boldsymbol{\alpha}=\boldsymbol{\theta}$.

实内积、可以看成一个函数：

$f:\langle\alpha,\beta\rangle\mapsto a_1b_1+a_2b_2+\cdots+a_nb_n$

若一个二元实值函数满足上述4条性质，则称 f 为内积函数

定义7、**定义 4.4.2** (向量夹角、向量正交) 若 $(\boldsymbol{\alpha},\boldsymbol{\beta})=0$，则称 $\boldsymbol{\alpha}$ 和 $\boldsymbol{\beta}$ **正交或垂直**. 若 $\boldsymbol{\alpha}$、$\boldsymbol{\beta}$ 均为非零实向量，则称 $\arccos\left(\dfrac{(\boldsymbol{\alpha},\boldsymbol{\beta})}{\|\boldsymbol{\alpha}\|\cdot\|\boldsymbol{\beta}\|}\right)$ 为向量 $\boldsymbol{\alpha}$ 和 $\boldsymbol{\beta}$ 的**夹角**.

柯西-布涅科夫斯基不等式： 对任意向量 α,β，总有 $|(\alpha,\beta)|\leq\|\alpha\|\cdot\|\beta\|$.

证明：$\beta=0$ 时，显然成立.

$\beta\neq 0$ 时，对任意实数 t，

由内积性质：$(\alpha+t\beta,\alpha+t\beta)\geqslant 0\Rightarrow(\alpha,\alpha)+2(\alpha,\beta)t+(\beta,\beta)t^2\geqslant 0.$

$\Delta=[2(\alpha,\beta)]^2-4(\alpha,\alpha)(\beta,\beta)\leq 0$

$(\alpha,\beta)^2\leq(\alpha,\alpha)(\beta,\beta)$，有 $|(\alpha,\beta)|\leq\|\alpha\|\|\beta\|$

显然，零向量 $\boldsymbol{\theta}$ 与任意同维向量正交. 另外由

$$\left\|\frac{\boldsymbol{\alpha}}{\|\boldsymbol{\alpha}\|}\pm\frac{\boldsymbol{\beta}}{\|\boldsymbol{\beta}\|}\right\|^2=\left(\frac{\boldsymbol{\alpha}}{\|\boldsymbol{\alpha}\|}\pm\frac{\boldsymbol{\beta}}{\|\boldsymbol{\beta}\|},\frac{\boldsymbol{\alpha}}{\|\boldsymbol{\alpha}\|}\pm\frac{\boldsymbol{\beta}}{\|\boldsymbol{\beta}\|}\right)\geqslant 0$$

可得 $\left|\dfrac{(\boldsymbol{\alpha},\boldsymbol{\beta})}{\|\boldsymbol{\alpha}\|\cdot\|\boldsymbol{\beta}\|}\right|\leqslant 1$，于是向量夹角的定义是合理的.

例 4.4.1　若有两个不同的实向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$ 满足 $\|\boldsymbol{\alpha}_1\|=\|\boldsymbol{\alpha}_2\|>0$，试证明 $\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2$ 与 $\boldsymbol{\alpha}_1-\boldsymbol{\alpha}_2$ 正交.

证明　由 $\|\boldsymbol{\alpha}_1\|=\|\boldsymbol{\alpha}_2\|>0$ 可知 $(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_1)=(\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_2)>0$，故有

$$\begin{aligned}(\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_1-\boldsymbol{\alpha}_2)&=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_1)+(\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_1)-(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2)-(\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_2)\\&=(\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_1)-(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2)=0,\end{aligned}$$

即 $\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2$ 与 $\boldsymbol{\alpha}_1-\boldsymbol{\alpha}_2$ 正交.　□

例 4.4.2　方程组 $\boldsymbol{Ax}=\boldsymbol{\theta}$ 的解集即为与 $\boldsymbol{A}$ 的所有行向量正交的向量的集合.

解　将矩阵 $\boldsymbol{A}$ 写成按行分块形式，

$$\boldsymbol{A}=\begin{pmatrix}\boldsymbol{\alpha}_1^{\mathrm{T}}\\ \vdots\\ \boldsymbol{\alpha}_n^{\mathrm{T}}\end{pmatrix},$$

则 $\boldsymbol{Ax}=\boldsymbol{\theta}$ 即为

$$\boldsymbol{Ax}=\begin{pmatrix}\boldsymbol{\alpha}_1^{\mathrm{T}}\\ \vdots\\ \boldsymbol{\alpha}_n^{\mathrm{T}}\end{pmatrix}\boldsymbol{x}=\begin{pmatrix}\boldsymbol{\alpha}_1^{\mathrm{T}}\boldsymbol{x}\\ \vdots\\ \boldsymbol{\alpha}_n^{\mathrm{T}}\boldsymbol{x}\end{pmatrix}=\begin{pmatrix}0\\ \vdots\\ 0\end{pmatrix},$$

即 $(\boldsymbol{\alpha}_i,\boldsymbol{x})=\boldsymbol{\alpha}_i^{\mathrm{T}}\boldsymbol{x}=0,\ i=1,2,\cdots,n$.　□

定义7. **定义 4.4.3**（正交向量组、法正交组）　若一个不含零向量的向量组中的向量两两正交，则称该向量组为**正交向量组**；若一个正交向量组中的向量均为单位向量，则该向量组称为**标准正交向量组**，简称**法正交组**.

例 4.4.3　易验证 $\mathbf{R}^n$ 中的基本向量组 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n$ 是法正交组.　□

定理10. **定理 4.4.1**　正交向量组必线性无关.

证明. 设 $\alpha_1,\alpha_2,\ldots,\alpha_\ell$ 是正交向量组。若有

$$k_1\alpha_1+k_2\alpha_2+\cdots+k_\ell\alpha_\ell=0,$$

则对任意的 i，$1\le i\le\ell$，用 α_i 跟上式两端分别做内积可得

$$(\alpha_i,k_1\alpha_1+k_2\alpha_2+\cdots+k_\ell\alpha_\ell)=(\alpha_i,0)\iff k_i(\alpha_i,\alpha_i)=0\iff k_i=0,$$

故 $\alpha_1,\alpha_2,\ldots,\alpha_\ell$ 线性无关。

$\boldsymbol{\alpha}_m,\boldsymbol{\alpha}_i)=k_1(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_i)$

$2,\cdots,m$. 因为 $\boldsymbol{\alpha}_i\neq$

无关.　□

一个极大无关组可以表示其所在向量组的全部向量，但需要解一个方程组. 如果一个极大无关组还是一个正交向量组，则用它来表示其所在向量组的其他向量就相当简单，只要用内积就可以计算出表示式中的系数. 那么能否由一个线性无关向量组构造出一个与之等价的正交向量组呢？下面的定理给出了肯定的回答，并给出了构造的方法.

注意： 若向量 β 可由正交向量组 $\alpha_1,\alpha_2,\ldots,\alpha_\ell$ 线性表示，比如 $\beta=k_1\alpha_1+k_2\alpha_2+\cdots+k_\ell\alpha_\ell$，则易知

$$k_i=\frac{(\alpha_i,\beta)}{(\alpha_i,\alpha_i)},\quad 1\le i\le\ell.$$

定理11. **定理 4.4.2** (施密特 (Schmidt) 正交化) **由线性无关向量组 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n$ 可构造出与之等价的正交向量组 $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \cdots, \boldsymbol{\xi}_n$. 并且 $\boldsymbol{\xi}_i$ 可以表示成 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_i, i = 1, \cdots, n$ 的线性组合.**

取 $\xi_1 = \alpha_1$

$\xi_2 = k_2\alpha_2 + k_1\alpha_1$ 取 $k_2 = 1$.

$\xi_2 = \alpha_2 + k_1\xi_1$

$(\xi_2, \xi_1) = 0 \Rightarrow (\xi_1, \alpha_2 + k_1\xi_1) = 0$

$(\xi_1, \alpha_2) + k_1(\xi_1, \xi_1) = 0$

$k_1 = -\frac{(\alpha_2, \xi_1)}{(\xi_1, \xi_1)} = -\frac{(\alpha_2, \xi_1)}{\|\xi_1\|^2}$

证明 下面我们用构造的方法证明. 我们设法用 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n$ 构造出正交向量组 $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \cdots, \boldsymbol{\xi}_n$, 构造过程如下.

第一步, 取 $\boldsymbol{\xi}_1 = \boldsymbol{\alpha}_1 \neq \boldsymbol{\theta}$.

第二步, 取 $\boldsymbol{\xi}_2 = \boldsymbol{\alpha}_2 - k_{21}\boldsymbol{\xi}_1$, 其中 $k_{21} = \frac{(\boldsymbol{\alpha}_2, \boldsymbol{\xi}_1)}{\|\boldsymbol{\xi}_1\|^2}$, 则首先 $\boldsymbol{\xi}_2 \neq \boldsymbol{\theta}$, 因为 $\boldsymbol{\alpha}_2 - k_{21}\boldsymbol{\xi}_1 = \boldsymbol{\theta}$

这与 α_1, α_2 线性无关矛盾。同时，ξ_1, ξ_2 正交：

$$(\xi_2, \xi_1) = \left(\alpha_2 - \frac{(\alpha_2, \xi_1)}{\|\xi_1\|^2}\xi_1, \xi_1\right) = (\alpha_2, \xi_1) - \frac{(\alpha_2, \xi_1)}{\|\xi_1\|^2}(\xi_1, \xi_1) = 0.$$

假设已构造出与 $\alpha_1, \alpha_2, \ldots, \alpha_\ell$ 等价的正交组 $\xi_1, \xi_2, \ldots, \xi_\ell$，我们现在构造 $\xi_{\ell+1}$：令

$$\xi_{\ell+1} = \alpha_{\ell+1} - k_1\xi_1 - \cdots - k_\ell\xi_\ell,$$

则 $\xi_{\ell+1} \neq 0$，否则 $\alpha_{\ell+1}$ 可由 $\xi_1, \xi_2, \ldots, \xi_\ell$ 线性表示，
从而可由 $\alpha_1, \alpha_2, \ldots, \alpha_\ell$ 线性表示，这与 $\alpha_1, \alpha_2, \ldots, \alpha_{\ell+1}$ 线性无关矛盾。

…是正交向量组且 $\boldsymbol{\xi}_2$ 是

令

$$k_i = \frac{(\alpha_{\ell+1}, \xi_i)}{\|\xi_i\|^2}, \quad 1 \le i \le \ell,$$

则有

$$\begin{aligned}(\xi_{\ell+1}, \xi_i) &= (\alpha_{\ell+1} - k_1\xi_1 - \cdots - k_\ell\xi_\ell, \xi_i) \\ &= (\alpha_{\ell+1}, \xi_i) - k_i(\xi_i, \xi_i) \\ &= (\alpha_{\ell+1}, \xi_i) - \frac{(\alpha_{\ell+1}, \xi_i)}{\|\xi_i\|^2}(\xi_i, \xi_i) = 0.\end{aligned}$$

…$\boldsymbol{\xi}_{i-1} = \boldsymbol{\theta}$, (1)

…的线性组合, 式 (1) 就可改写成

…$\boldsymbol{\alpha}_{i-1} = \boldsymbol{\theta}$,

…1. 故 $\{\boldsymbol{\xi}_1, \cdots, \boldsymbol{\xi}_i\}$ 是正交向量组且

…且是正交向量组, 且 $\boldsymbol{\xi}_i, i = 1, 2, \cdots, n$ 是 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_i$ 的线性组合.

由前面的过程可知, 向量组 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n\}$ 与向量组 $\{\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \cdots, \boldsymbol{\xi}_n\}$ 可互相线性表示, 故两向量组等价. □

注 由线性无关向量组 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n$ 可构造出与之等价的法正交组 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_n$. 并且 $\boldsymbol{\beta}_i$ 可以表示成 $\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_i, i = 1, \cdots, n$ 的线性组合.

例10. **例 4.4.4** 将 3 个线性无关 4 维向量组 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3\}$ 标准正交化, 其中

$$\boldsymbol{\alpha}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \quad \boldsymbol{\alpha}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \quad \boldsymbol{\alpha}_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}.$$

施密特 (Schmidt) 正交化方法一般步骤：

1. 令 $\xi_1 = \alpha_1$.
2. 若已构造出与 $\alpha_1, \alpha_2, \ldots, \alpha_\ell$ 等价的正交向量组 $\xi_1, \xi_2, \ldots, \xi_\ell$，则令

$$\xi_{\ell+1} = \alpha_{\ell+1} - \frac{(\alpha_{\ell+1}, \xi_1)}{\|\xi_1\|^2}\xi_1 - \cdots - \frac{(\alpha_{\ell+1}, \xi_\ell)}{\|\xi_\ell\|^2}\xi_\ell.$$

解　先做正交化，令

$$\boldsymbol{\xi}_1=\boldsymbol{\alpha}_1=\begin{pmatrix}1\\0\\0\\1\end{pmatrix},\quad \boldsymbol{\xi}_2=\boldsymbol{\alpha}_2-\frac{(\boldsymbol{\alpha}_2,\boldsymbol{\xi}_1)}{\|\boldsymbol{\xi}_1\|^2}\boldsymbol{\xi}_1=\begin{pmatrix}1\\1\\0\\1\end{pmatrix}-\frac{2}{2}\begin{pmatrix}1\\0\\0\\1\end{pmatrix}=\begin{pmatrix}0\\1\\0\\0\end{pmatrix},$$

$$\boldsymbol{\xi}_3=\boldsymbol{\alpha}_3-\frac{(\boldsymbol{\alpha}_3,\boldsymbol{\xi}_1)}{\|\boldsymbol{\xi}_1\|^2}\boldsymbol{\xi}_1-\frac{(\boldsymbol{\alpha}_3,\boldsymbol{\xi}_2)}{\|\boldsymbol{\xi}_2\|^2}\boldsymbol{\xi}_2=\begin{pmatrix}1\\1\\1\\0\end{pmatrix}-\frac{1}{2}\begin{pmatrix}1\\0\\0\\1\end{pmatrix}-\frac{1}{1}\begin{pmatrix}0\\1\\0\\0\end{pmatrix}=\begin{pmatrix}1/2\\0\\1\\-1/2\end{pmatrix}.$$

再做标准化，

$$\boldsymbol{\beta}_1=\frac{1}{\|\boldsymbol{\xi}_1\|}\boldsymbol{\xi}_1=\frac{1}{\sqrt{2}}\boldsymbol{\xi}_1=\begin{pmatrix}1/\sqrt{2}\\0\\0\\1/\sqrt{2}\end{pmatrix},$$

$$\boldsymbol{\beta}_2=\frac{1}{\|\boldsymbol{\xi}_2\|}\boldsymbol{\xi}_2=\boldsymbol{\xi}_2=\begin{pmatrix}0\\1\\0\\0\end{pmatrix},$$

$$\boldsymbol{\beta}_3=\frac{1}{\|\boldsymbol{\xi}_3\|}\boldsymbol{\xi}_3=\sqrt{\frac{2}{3}}\boldsymbol{\xi}_3=\begin{pmatrix}1/\sqrt{6}\\0\\2/\sqrt{6}\\-1/\sqrt{6}\end{pmatrix}.$$

得标准正交向量组为 $\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3\}$.　□

定义9、**定义 4.4.4** (正交矩阵)　若实方阵 $\boldsymbol{A}$ 满足 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\boldsymbol{E}$，则称 $\boldsymbol{A}$ 为**正交矩阵**.

注　若 $\boldsymbol{A},\boldsymbol{B}$ 是同阶的正交矩阵，则 $\boldsymbol{AB}$ 也是正交矩阵.

对于单位矩阵 $\boldsymbol{E}$，显然有 $\boldsymbol{E}^{\mathrm{T}}\boldsymbol{E}=\boldsymbol{E}$，故 $\boldsymbol{E}$ 为正交矩阵.

容易验证下列矩阵

$$\begin{pmatrix}\cos\theta & \sin\theta\\-\sin\theta & \cos\theta\end{pmatrix},\quad\begin{pmatrix}-1&0&0\\0&1&0\\0&0&-1\end{pmatrix},\quad\begin{pmatrix}\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{3}}&\frac{1}{\sqrt{6}}\\-\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{3}}&\frac{1}{\sqrt{6}}\\0&-\frac{1}{\sqrt{3}}&\frac{2}{\sqrt{6}}\end{pmatrix}$$

都是正交矩阵.

例11、**例 4.4.5**　设 $\boldsymbol{A}=\boldsymbol{E}-2\boldsymbol{v}\boldsymbol{v}^{\mathrm{T}}\in\mathbf{R}^{n\times n}$，其中 $\boldsymbol{v}\in\mathbf{R}^n,\boldsymbol{v}^{\mathrm{T}}\boldsymbol{v}=1$. 验证 $\boldsymbol{A}$ 是对称矩阵和正交矩阵. 且当 $\|\boldsymbol{p}\|=\|\boldsymbol{q}\|,\boldsymbol{p}\neq\boldsymbol{q},\boldsymbol{p},\boldsymbol{q}\in\mathbf{R}^n,\boldsymbol{v}=\dfrac{\boldsymbol{p}-\boldsymbol{q}}{\|\boldsymbol{p}-\boldsymbol{q}\|}$ 时，有

$$\boldsymbol{Ap}=\boldsymbol{q},\qquad \boldsymbol{Aq}=\boldsymbol{p}.$$

解 因为

$$\boldsymbol{A}^{\mathrm{T}}=(\boldsymbol{E}-2\boldsymbol{v}\boldsymbol{v}^{\mathrm{T}})^{\mathrm{T}}=\boldsymbol{E}-2(\boldsymbol{v}\boldsymbol{v}^{\mathrm{T}})^{\mathrm{T}}=\boldsymbol{E}-2\boldsymbol{v}\boldsymbol{v}^{\mathrm{T}}=\boldsymbol{A},$$
$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=(\boldsymbol{E}-2\boldsymbol{v}\boldsymbol{v}^{\mathrm{T}})(\boldsymbol{E}-2\boldsymbol{v}\boldsymbol{v}^{\mathrm{T}})=\boldsymbol{E}-2\boldsymbol{v}\boldsymbol{v}^{\mathrm{T}}-2\boldsymbol{v}\boldsymbol{v}^{\mathrm{T}}+4\boldsymbol{v}(\boldsymbol{v}^{\mathrm{T}}\boldsymbol{v})\boldsymbol{v}^{\mathrm{T}}=\boldsymbol{E},$$

所以 $\boldsymbol{A}$ 为对称和正交矩阵.

当 $\boldsymbol{v}=\dfrac{\boldsymbol{p}-\boldsymbol{q}}{\|\boldsymbol{p}-\boldsymbol{q}\|}$ 时，有

$$\begin{aligned}\boldsymbol{A}\boldsymbol{p}&=\left(\boldsymbol{E}-2\frac{(\boldsymbol{p}-\boldsymbol{q})(\boldsymbol{p}-\boldsymbol{q})^{\mathrm{T}}}{\|\boldsymbol{p}-\boldsymbol{q}\|^2}\right)\boldsymbol{p}\\&=\boldsymbol{p}-2\frac{(\boldsymbol{p}-\boldsymbol{q})(\boldsymbol{p}^{\mathrm{T}}\boldsymbol{p}-\boldsymbol{q}^{\mathrm{T}}\boldsymbol{p})}{(\boldsymbol{p}-\boldsymbol{q})^{\mathrm{T}}(\boldsymbol{p}-\boldsymbol{q})}\\&=\boldsymbol{p}-2\frac{(\boldsymbol{p}^{\mathrm{T}}\boldsymbol{p}-\boldsymbol{q}^{\mathrm{T}}\boldsymbol{p})(\boldsymbol{p}-\boldsymbol{q})}{2(\boldsymbol{p}^{\mathrm{T}}\boldsymbol{p}-\boldsymbol{q}^{\mathrm{T}}\boldsymbol{p})}\\&=\boldsymbol{p}-(\boldsymbol{p}-\boldsymbol{q})=\boldsymbol{q}.\end{aligned}$$

同理有 $\boldsymbol{A}\boldsymbol{q}=\boldsymbol{p}$. □

关于正交矩阵，有下面一些基本性质.

定理12、**定理 4.4.3** 对于方阵 $\boldsymbol{A}$，下列条件互为等价:

(1) $\boldsymbol{A}$ 为正交矩阵;

(2) $\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{A}^{-1}$;

(3) $\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{E}$;

(4) $\boldsymbol{A}$ 的列向量构成标准正交列向量组;

(5) $\boldsymbol{A}$ 的行向量构成标准正交行向量组.

证明. 显然，(1) $\Longleftrightarrow$ (2) $\Longleftrightarrow$ (3).

下证 (1) $\Longleftrightarrow$ (4). 设 $A=(\xi_1\ \xi_2\ \cdots\ \xi_n)$，则有

$$A^TA=\begin{pmatrix}\xi_1^T\\ \xi_2^T\\ \cdots\\ \xi_n^T\end{pmatrix}(\xi_1\ \xi_2\ \cdots\ \xi_n)=\begin{pmatrix}\xi_1^T\xi_1&\xi_1^T\xi_2&\cdots&\xi_1^T\xi_n\\ \xi_2^T\xi_1&\xi_2^T\xi_2&\cdots&\xi_2^T\xi_n\\ \cdots&\cdots&\cdots&\cdots\\ \xi_n^T\xi_1&\xi_n^T\xi_2&\cdots&\xi_n^T\xi_n\end{pmatrix}.$$

(1) $\Rightarrow$ (4): 若 A 是正交矩阵，则有

$$A^TA=\begin{pmatrix}1&&&\\&1&&\\&&\ddots&\\&&&1\end{pmatrix}=\begin{pmatrix}\xi_1^T\xi_1&\xi_1^T\xi_2&\cdots&\xi_1^T\xi_n\\ \xi_2^T\xi_1&\xi_2^T\xi_2&\cdots&\xi_2^T\xi_n\\ \cdots&\cdots&\cdots&\cdots\\ \xi_n^T\xi_1&\xi_n^T\xi_2&\cdots&\xi_n^T\xi_n\end{pmatrix}.$$

故 $\xi_i^T\xi_i=1,\ 1\le i\le n$; $\xi_i^T\xi_j=0,\ 1\le i<j\le n$，即 $\xi_1,\xi_2,\ldots,\xi_n$ 是标准正交组。

(4) $\Rightarrow$ (1): 若 $\xi_1,\xi_2,\ldots,\xi_n$ 是标准正交组，则有

$$A^TA=\begin{pmatrix}\xi_1^T\xi_1&\xi_1^T\xi_2&\cdots&\xi_1^T\xi_n\\ \xi_2^T\xi_1&\xi_2^T\xi_2&\cdots&\xi_2^T\xi_n\\ \cdots&\cdots&\cdots&\cdots\\ \xi_n^T\xi_1&\xi_n^T\xi_2&\cdots&\xi_n^T\xi_n\end{pmatrix}=\begin{pmatrix}1&&&\\&1&&\\&&\ddots&\\&&&1\end{pmatrix}.$$

故 A 是正交矩阵。

$\delta_{ij}=\begin{cases}1, & i=j\\ 0, & i\neq j\end{cases}$. 故 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 为法正交列向量组. 反之则有 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=(\boldsymbol{\alpha}_i^{\mathrm{T}}\boldsymbol{\alpha}_j)_{n\times n}=(\delta_{ij})_{n\times n}=\boldsymbol{E}$，故 (1) 成立. 从而得 (1)$\Longleftrightarrow$(4). 同理 (3)$\Longleftrightarrow$(5). 故 (1)、(2)、$\cdots$、(5) 相互等价. □

定理13、**定理 4.4.4** 设 $\boldsymbol{A}$ 为 n 阶正交矩阵，λ 为 $\boldsymbol{A}$ 的特征值 $(\lambda\in\mathbb{C})$，$\boldsymbol{\alpha}$ 为 n 维列向量，则有

(1) $|\boldsymbol{A}|^2=1$;

(2) $(\boldsymbol{A}\boldsymbol{\alpha})^{\mathrm{T}}(\overline{\boldsymbol{A}\boldsymbol{\alpha}})=\boldsymbol{\alpha}^{\mathrm{T}}\overline{\boldsymbol{\alpha}}$;

(3) $\|\lambda\|=1$.

证明　(1) 由正交矩阵的定义, $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\boldsymbol{E}$, 得

$$|\boldsymbol{A}|^2=|\boldsymbol{A}||\boldsymbol{A}|=|\boldsymbol{A}^{\mathrm{T}}||\boldsymbol{A}|=|\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}|=|\boldsymbol{E}|=1.$$

(2) 注意到 $\overline{\boldsymbol{A}}$ 表示对 $\boldsymbol{A}$ 的每个元素取共轭, $\overline{\boldsymbol{\xi}}$ 表示对 $\boldsymbol{\xi}$ 的每个元素取共轭，由矩阵运算性质可得

复向量内积　　A是实阵

$$(\boldsymbol{A\alpha})^{\mathrm{T}}(\overline{\boldsymbol{A\alpha}})=\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}(\overline{\boldsymbol{A}}\,\overline{\boldsymbol{\alpha}})=\boldsymbol{\alpha}^{\mathrm{T}}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})\overline{\boldsymbol{\alpha}}=\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{E}\,\overline{\boldsymbol{\alpha}}=\boldsymbol{\alpha}^{\mathrm{T}}\overline{\boldsymbol{\alpha}}.$$ 模的平方不变

(3) 设 $\boldsymbol{\xi}$ 为 $\boldsymbol{A}$ 的属于 λ 的特征向量，则有 $\boldsymbol{A\xi}=\lambda\boldsymbol{\xi}$, 于是

$$(\boldsymbol{A\xi})^{\mathrm{T}}(\overline{\boldsymbol{A\xi}})=(\lambda\boldsymbol{\xi})^{\mathrm{T}}(\overline{\lambda\boldsymbol{\xi}})=\overset{\lambda\bar{\lambda}\xi^T\bar{\xi}}{\|\lambda\|^2\boldsymbol{\xi}^{\mathrm{T}}\overline{\boldsymbol{\xi}}}.$$

另一方面，注意到 $\boldsymbol{\xi}\neq\boldsymbol{\theta}$ 并应用 (2) 的结果得

$$(\boldsymbol{A\xi})^{\mathrm{T}}(\overline{\boldsymbol{A\xi}})=\boldsymbol{\xi}^{\mathrm{T}}\overline{\boldsymbol{\xi}}\neq 0.$$

所以 $\|\lambda\|^2\boldsymbol{\xi}^{\mathrm{T}}\overline{\boldsymbol{\xi}}=\boldsymbol{\xi}^{\mathrm{T}}\overline{\boldsymbol{\xi}}\neq 0$. 两边除以 $\boldsymbol{\xi}^{\mathrm{T}}\overline{\boldsymbol{\xi}}$ 即得 $\|\lambda\|=1$.　$\begin{pmatrix}\frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2}\\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2}\end{pmatrix}$ 实正交阵但无实特征值.　□

4.5　实对称矩阵的对角化

一个实矩阵 $\boldsymbol{A}$ 具有对称性，即 $\left\{\begin{array}{l}\overline{A}=A\\ \boldsymbol{A}^{\mathrm{T}}=\boldsymbol{A}\end{array}\right.$，称它为**实对称矩阵**. 一般矩阵不一定能对角化，但实对称矩阵一定可以对角化，而且该矩阵还有许多很好的性质. 本节内容除定理 4.5.1 的证明过程考虑复数外，均在实数范围内考虑. 下面先讨论实对称矩阵的一些特有性质.

定理14、**定理 4.5.1**　实对称矩阵的特征值均为实数.

证明. 设 A 为实对称阵，λ 为 A 的特征值，ξ 为 A 的属于 λ 的特征量。于是，

$$(\xi,A\xi)=\xi^T\overline{A\xi}=\xi^T\overline{\lambda\xi}=\bar{\lambda}\xi^T\bar{\xi},$$

$$(\xi,A\xi)=\xi^T\overline{A\xi}=\xi^T\overline{A\xi}=\xi^TA^T\bar{\xi}=(A\xi)^T\bar{\xi}=\lambda\xi^T\bar{\xi}.$$

比较可得 $\bar{\lambda}\xi^T\bar{\xi}=\lambda\xi^T\bar{\xi}$. 由于 $\xi^T\bar{\xi}$ 是不等于零的实数，故有

$$\bar{\lambda}=\lambda.$$

. 再设 $\boldsymbol{A}$ 有特征值 λ 和属于 λ 的特征向

$\boldsymbol{A\xi}=\lambda\overline{\boldsymbol{\xi}}^{\mathrm{T}}\boldsymbol{\xi}$，从而

$^{\mathrm{T}}\boldsymbol{\xi}=(\overline{\lambda\boldsymbol{\xi}})^{\mathrm{T}}\boldsymbol{\xi}=\overline{\lambda}\,\overline{\boldsymbol{\xi}}^{\mathrm{T}}\boldsymbol{\xi}.$

寺征值 λ 为实数.　□

定理15、**定理 4.5.2**　实对称矩阵的属于不同特征值的特征向量相互正交.

证明. 设 λ_1,λ_2 是实对称矩阵 A 的两个不同特征值，ξ_1,ξ_2 为相应的特征向量，亦即 $A\xi_1=\lambda_1\xi_1$，$A\xi_2=\lambda_2\xi_2$. 于是，

$$\xi_2^TA\xi_1=\lambda_1\xi_2^T\xi_1,\quad\Longrightarrow\quad \xi_1^TA\xi_2=\lambda_1\xi_1^T\xi_2.$$

另一方面，注意到

$$\xi_1^TA\xi_2=\lambda_2\xi_1^T\xi_2,$$

我们有

$$(\lambda_1-\lambda_2)\,\xi_1^T\xi_2=0\quad\Longrightarrow\quad\xi_1^T\xi_2=0.$$

5.1 知 λ_1,λ_2 为实数，故有属

$\lambda_1\boldsymbol{\xi}_2^{\mathrm{T}}\boldsymbol{\xi}_1=\lambda_1\boldsymbol{\xi}_1^{\mathrm{T}}\boldsymbol{\xi}_2,$

互正交.　□

下面讨论实对称矩阵的可对角化问题.

引理1、**引理 4.5.1**　设有实 n 维单位列向量 $\boldsymbol{\beta}$, 则必能找到 $n-1$ 个向量与 $\boldsymbol{\beta}$ 一起构成由 n 个向量组成的标准正交向量组.

证明. 由于 $r\{\xi, e_1, \ldots, e_n\} = n$,

故 $\xi, e_1, e_2, \ldots, e_n$ 中包含 ξ 的极大的线性无关组必定包含 n 个向量，记为

$$\xi, \xi_2, \xi_3, \ldots, \xi_n,$$

其中 $\xi_2, \xi_3 \ldots, \xi_n$ 取自单位向量 $e_1, e_2, \ldots, e_n$.

使用施密特正交化方法对 $\xi, \xi_2, \xi_3, \ldots, \xi_n$ 实施正交化（取 $\eta_1 = \xi$）可得一标准正交组:

$$\xi, \eta_2, \eta_3, \ldots, \eta_n$$

定理16. **定理 4.5.3** 若 $\boldsymbol{A}$ 是实对称矩阵，则存在同阶的正交矩阵 $\boldsymbol{P}$ 使得 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{AP}$ 是实对角矩阵，从而实对称矩阵可对角化.

证明. 对 n 进行归纳。当 $n=1$ 时结论是显然的。
假设当 $n=m$ 时结论成立。现在设 $n=m+1$，λ_1 是 A的特征值。
根据定理 15，λ_1 是实数。因此，存在属于 λ_1 的实单位特征向量 ξ_1.
由引理 1，存在 $\xi_2, \ldots, \xi_n$ 使得 $Q = (\xi_1\ \xi_2\ \cdots\ \xi_n)$ 是正交阵。于是，

$$Q^TAQ = \begin{pmatrix} \xi_1^T \\ \xi_2^T \\ \cdots \\ \xi_n^T \end{pmatrix} A\,(\xi_1\ \xi_2\ \cdots\ \xi_n) = \begin{pmatrix} \xi_1^T \\ \xi_2^T \\ \cdots \\ \xi_n^T \end{pmatrix} (\lambda_1\xi_1\ A\xi_2\ \cdots\ A\xi_n).$$

$$Q^TAQ = \begin{pmatrix} \xi_1^T \\ \xi_2^T \\ \cdots \\ \xi_n^T \end{pmatrix} (\lambda_1\xi_1\ A\xi_2\ \cdots\ A\xi_n) = \begin{pmatrix} \lambda_1 & \xi_1^TA\xi_2 & \cdots & \xi_1^TA\xi_n \\ 0 & \xi_2^TA\xi_2 & \cdots & \xi_2^TA\xi_n \\ \cdots & \cdots & \cdots & \cdots \\ 0 & \xi_n^TA\xi_2 & \cdots & \xi_n^TA\xi_n \end{pmatrix}$$

$$= \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \xi_2^TA\xi_2 & \cdots & \xi_2^TA\xi_n \\ \cdots & \cdots & \cdots & \cdots \\ 0 & \xi_n^TA\xi_2 & \cdots & \xi_n^TA\xi_n \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & A_1 \end{pmatrix},$$

其中，A_1 是 $m=n-1$ 阶实对称阵。

根据归纳假定，存在 m 阶正交矩阵 Q_1 使得 $Q_1^TA_1Q_1 = D$ 为对角形。令

$$P = Q\begin{pmatrix} 1 & 0 \\ 0 & Q_1 \end{pmatrix},$$

则 P 为正交阵，且有

$$P^TAP = \begin{pmatrix} 1 & 0 \\ 0 & Q_1 \end{pmatrix}^T Q^TAQ \begin{pmatrix} 1 & 0 \\ 0 & Q_1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 \\ 0 & Q_1^T \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & A_1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & Q_1 \end{pmatrix}$$

$$= \begin{pmatrix} \lambda_1 & 0 \\ 0 & Q_1^TA_1Q_1 \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & D \end{pmatrix}.$$

设 A 实对称矩阵，如何求正交矩阵 P 使得 P^TAP 是对角形?

1. 计算行列式 $|\lambda E - A|$；

2. 求 $|\lambda E - A| = 0$ 的根，设为 $\lambda_1^{s_1}, \lambda_2^{s_2}, \ldots, \lambda_r^{s_r}$, $\lambda_1, \lambda_2, \ldots, \lambda_r$ 两两不同；

3. 对每个特征值 λ_i, $1 \le i \le r$, 求 $(\lambda_i E - A)\,x = 0$ 的一个基础解系:

$$\alpha_{i1}, \alpha_{i2}, \ldots, \alpha_{i\ell_i}.$$

用施密特正交化方法将 $\alpha_{i1}, \alpha_{i2}, \ldots, \alpha_{i\ell_i}$ 标准正交化得:

$$\xi_{i1}, \xi_{i2}, \ldots, \xi_{i\ell_i}.$$

4. 令 $P = (\xi_{11}\ \cdots\ \xi_{1\ell_1}\ \xi_{21}\ \cdots\ \xi_{2\ell_2}\ \cdots\ \xi_{r1}\ \cdots\ \xi_{r\ell_r})$, 则 P 为正交阵，且 $P^TAP = P^{-1}AP$ 为对角形。

$$= \begin{pmatrix} \lambda_1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{Q}_2^{\mathrm{T}}\boldsymbol{BQ}_2 \end{pmatrix} = \begin{pmatrix} & \\ \boldsymbol{0} & \boldsymbol{\Lambda} \end{pmatrix},$$

故 $n = m+1$ 时结论也成立.

进一步由定理 4.3.1 可知，$\lambda_1, \lambda_2, \cdots, \lambda_n$ 为 $\boldsymbol{A}$ 的特征值. □

例12. **例 4.5.1** 设

$$\boldsymbol{A} = \begin{pmatrix} 2 & 1 & 2 \\ 1 & 3 & 1 \\ 2 & 1 & 2 \end{pmatrix},$$

求正交矩阵 $\boldsymbol{P}$，使 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{AP}$ 为对角矩阵.

解　由

$$|\lambda \boldsymbol{E}-\boldsymbol{A}|=\begin{vmatrix}\lambda-2 & -1 & -2\\ -1 & \lambda-3 & -1\\ -2 & -1 & \lambda-2\end{vmatrix}=\begin{vmatrix}\lambda & 0 & -\lambda\\ -1 & \lambda-3 & -1\\ -2 & -1 & \lambda-2\end{vmatrix}=\lambda(\lambda-2)(\lambda-5),$$

解得特征值为 $\lambda=5,2,0$.

对 $\lambda=5$, 由 $\begin{pmatrix}3 & -1 & -2\\ -1 & 2 & -1\\ -2 & -1 & 3\end{pmatrix}\to\begin{pmatrix}1 & 0 & -1\\ 0 & 1 & -1\\ 0 & 0 & 0\end{pmatrix}$ 得特征向量 $\boldsymbol{\xi}_1=\begin{pmatrix}1\\1\\1\end{pmatrix}$.

对 $\lambda=2$, 由 $\begin{pmatrix}0 & -1 & -2\\ -1 & -1 & -1\\ -2 & -1 & 0\end{pmatrix}\to\begin{pmatrix}1 & 0 & -1\\ 0 & 1 & 2\\ 0 & 0 & 0\end{pmatrix}$ 得特征向量 $\boldsymbol{\xi}_2=\begin{pmatrix}1\\-2\\1\end{pmatrix}$.

对 $\lambda=0$, 由 $\begin{pmatrix}-2 & -1 & -2\\ -1 & -3 & -1\\ -2 & -1 & -2\end{pmatrix}\to\begin{pmatrix}1 & 0 & 1\\ 0 & 1 & 0\\ 0 & 0 & 0\end{pmatrix}$ 得特征向量 $\boldsymbol{\xi}_3=\begin{pmatrix}-1\\0\\1\end{pmatrix}$.

由定理 4.5.2 知, $\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\boldsymbol{\xi}_3$ 已两两正交，因此，只要将它们单位化. 取

$$\boldsymbol{\eta}_1=\frac{1}{\sqrt{\boldsymbol{\xi}_1^{\mathrm{T}}\boldsymbol{\xi}_1}}\boldsymbol{\xi}_1=\frac{1}{\sqrt{3}}\begin{pmatrix}1\\1\\1\end{pmatrix},\quad \boldsymbol{\eta}_2=\frac{1}{\sqrt{\boldsymbol{\xi}_2^{\mathrm{T}}\boldsymbol{\xi}_2}}\boldsymbol{\xi}_2=\frac{1}{\sqrt{6}}\begin{pmatrix}1\\-2\\1\end{pmatrix},$$

$$\boldsymbol{\eta}_3=\frac{1}{\sqrt{\boldsymbol{\xi}_3^{\mathrm{T}}\boldsymbol{\xi}_3}}\boldsymbol{\xi}_3=\frac{1}{\sqrt{2}}\begin{pmatrix}-1\\0\\1\end{pmatrix},$$

再令

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

则有 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{P}=\boldsymbol{E},\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\mathrm{diag}\,(5,2,0)$. □

例13、**例 4.5.2**　设

$$\boldsymbol{A}=\begin{pmatrix}1 & -2 & 2\\ -2 & -2 & 4\\ 2 & 4 & -2\end{pmatrix},$$

求正交矩阵 $\boldsymbol{P}$，使 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}$ 为对角矩阵.

解 由

$$|\lambda \boldsymbol{E}-\boldsymbol{A}|=\begin{vmatrix}\lambda-1 & 2 & -2\\ 2 & \lambda+2 & -4\\ -2 & -4 & \lambda+2\end{vmatrix}=\begin{vmatrix}\lambda-1 & 2 & -2\\ 2 & \lambda+2 & -4\\ 0 & \lambda-2 & \lambda-2\end{vmatrix}=(\lambda-2)^2(\lambda+7),$$

解得特征值为 $\lambda=2$(二重)，-7.

对 $\lambda=2$, 由 $\begin{pmatrix}1 & 2 & -2\\ 2 & 4 & -4\\ -2 & -4 & 4\end{pmatrix}\to\begin{pmatrix}1 & 2 & -2\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}$ 得线性无关特征向量

$$\boldsymbol{\xi}_1=\begin{pmatrix}-2\\ 1\\ 0\end{pmatrix},\quad \boldsymbol{\xi}_2=\begin{pmatrix}2\\ 0\\ 1\end{pmatrix},$$

标准正交化得$\boldsymbol{\eta}_1=\dfrac{1}{\sqrt{5}}\begin{pmatrix}-2\\ 1\\ 0\end{pmatrix}$，$\boldsymbol{\eta}_2=\dfrac{1}{3\sqrt{5}}\begin{pmatrix}2\\ 4\\ 5\end{pmatrix}$.

对 $\lambda=-7$, 由 $\begin{pmatrix}-8 & 2 & -2\\ 2 & -5 & -4\\ -2 & -4 & -5\end{pmatrix}\to\begin{pmatrix}1 & 0 & 1/2\\ 0 & 1 & 1\\ 0 & 0 & 0\end{pmatrix}$ 得单位特征向量

$$\boldsymbol{\eta}_3=\frac{1}{3}\begin{pmatrix}-1\\ -2\\ 2\end{pmatrix}.$$

再令

$$\boldsymbol{P}=(\boldsymbol{\eta}_1\boldsymbol{\eta}_2\boldsymbol{\eta}_3)=\begin{pmatrix}-\dfrac{2}{\sqrt{5}} & \dfrac{2}{3\sqrt{5}} & -\dfrac{1}{3}\\ \dfrac{1}{\sqrt{5}} & \dfrac{4}{3\sqrt{5}} & -\dfrac{2}{3}\\ 0 & \dfrac{5}{3\sqrt{5}} & \dfrac{2}{3}\end{pmatrix},$$

则有 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{P}=\boldsymbol{E}$, $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\mathrm{diag}\,(2,2,-7)$. □

例14. **例 4.5.3** 设 3 阶实对称矩阵 $\boldsymbol{A}$ 的秩为 2，$\lambda=6$ 为 $\boldsymbol{A}$ 的二重特征值，若 $\boldsymbol{\alpha}_1=(1,0,1)^{\mathrm{T}}$, $\boldsymbol{\alpha}_2=(1,3,-2)^{\mathrm{T}}$ 都是 $\boldsymbol{A}$ 的属于 $\lambda=6$ 的特征向量，求矩阵 $\boldsymbol{A}$.

解 因为 $\mathrm{r}(\boldsymbol{A})=2$，所以 $|\boldsymbol{A}|=0$, 故 $\lambda=0$ 为 $\boldsymbol{A}$ 的特征值. 设 $\boldsymbol{\alpha}_3$ 是属于特征值 0 的特征向量，则由 $\boldsymbol{A}$ 为实对称矩阵的性质可知: $\boldsymbol{\alpha}_3$ 与属于特征值 6 的特征向量 $\boldsymbol{\alpha}_1$, $\boldsymbol{\alpha}_2$ 正交.

令 $\boldsymbol{\alpha}_3=(x_1,x_2,x_3)^{\mathrm{T}}$，则由正交性得方程组

$$\begin{cases}(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_3)=x_1+x_3=0,\\ (\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3)=x_1+3x_2-2x_3=0.\end{cases}$$

解此方程组可得基础解系: $\boldsymbol{\alpha}_3=(-1,1,1)^{\mathrm{T}}$. 现在我们有

$= (6\alpha_1\ 6\alpha_2\ 0\alpha_3) = (\alpha_1\ \alpha_2\ \alpha_3)\begin{pmatrix}6&&\\&6&\\&&0\end{pmatrix}$

$$\boldsymbol{A}(\boldsymbol{\alpha}_1\ \ \boldsymbol{\alpha}_2\ \ \boldsymbol{\alpha}_3)=(6\boldsymbol{\alpha}_1\ \ 6\boldsymbol{\alpha}_2\ \ \boldsymbol{\theta}),$$

$\Rightarrow AP=P\Lambda$

即

$$\boldsymbol{A}\begin{pmatrix}1&1&-1\\0&3&1\\1&-2&1\end{pmatrix}=\begin{pmatrix}6&6&0\\0&18&0\\6&-12&0\end{pmatrix}.$$

故

$$\boldsymbol{A}=\begin{pmatrix}6&6&0\\0&18&0\\6&-12&0\end{pmatrix}\begin{pmatrix}1&1&-1\\0&3&1\\1&-2&1\end{pmatrix}^{-1}=\begin{pmatrix}6&6&0\\0&18&0\\6&-12&0\end{pmatrix}\begin{pmatrix}5/9&1/9&4/9\\1/9&2/9&-1/9\\-1/3&1/3&1/3\end{pmatrix}$$

$$=\begin{pmatrix}4&2&2\\2&4&-2\\2&-2&4\end{pmatrix}.$$

□

4.6* 若尔当 (Jordan) 标准形和奇异值分解

由 4.3 节可知有些矩阵可对角化，而有些矩阵不可对角化，这些不可对角化的矩阵是否也能相似变换为较简单的形式? 当我们将对角化放宽成块对角化时，任意矩阵就都可相似简化成块对角矩阵，并且对角块也可以有很简单的形式.

定理 4.6.1　*复数域上任意方阵 $\boldsymbol{A}$ 均可相似于一个***若尔当标准形**

$$\boldsymbol{J}=\begin{pmatrix}\boldsymbol{J}_1&&&\\&\boldsymbol{J}_2&&\\&&\ddots&\\&&&\boldsymbol{J}_s\end{pmatrix},$$

其中

$$\boldsymbol{J}_i=\begin{pmatrix}\lambda_i&1&0&\cdots&0\\0&\lambda_i&1&\cdots&0\\&&\ddots&&\\0&\cdots&0&\lambda_i&1\\0&\cdots&\cdots&0&\lambda_i\end{pmatrix},\quad i=1,2,\cdots,s,$$

*称为***若尔当块**.

证明　证明略. □

利用上述定理的结论，我们可以很容易地得出如下关于矩阵多项式 $f(\boldsymbol{A})$ 的特征值与 $\boldsymbol{A}$ 的特征值的关系.

定理 4.6.2 若 n 阶方阵 $\boldsymbol{A}$ 的全部特征值为 $\lambda_1,\cdots,\lambda_n$(可能有相同的特征值)，则 $\boldsymbol{A}$ 的矩阵多项式 $f(\boldsymbol{A})$ 的全部特征值为 $f(\lambda_1),\cdots,f(\lambda_n)$, 其中 $f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$.

证明 由定理 4.6.1 可知，存在可逆矩阵 $\boldsymbol{P}$ 使得 $\boldsymbol{B}=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ 为若尔当标准形，它是一个上三角形矩阵，即

$$\boldsymbol{B}=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\begin{pmatrix}\lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \cdots & * \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n\end{pmatrix}.$$

因为 $\boldsymbol{A}\sim\boldsymbol{B}$, 所以 $f(\boldsymbol{A})\sim f(\boldsymbol{B})$，即 $f(\boldsymbol{A})$ 的所有特征值与 $f(\boldsymbol{B})$ 的相同.

而易知

$$\begin{pmatrix}\lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \cdots & * \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n\end{pmatrix}^i=\begin{pmatrix}\lambda_1^i & * & \cdots & * \\ 0 & \lambda_2^i & \cdots & * \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n^i\end{pmatrix},$$

故有

$$f(\boldsymbol{B})=\begin{pmatrix}f(\lambda_1) & * & \cdots & * \\ 0 & f(\lambda_2) & \cdots & * \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & f(\lambda_n)\end{pmatrix},$$

即 $f(\boldsymbol{B})$ 的所有特征值为 $f(\lambda_1),\cdots,f(\lambda_n)$，此即 $f(\boldsymbol{A})$ 的全部特征值. □

将一个矩阵 $\boldsymbol{A}$ 简化成对角矩阵，除了相似变换 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ 外，也可以要求左、右矩阵放宽为不同的可逆矩阵，这是初等变换，考虑到正交矩阵的优良性质，也可以要求左、右矩阵为不同的正交矩阵，也能保证化为对角矩阵. 下面就证明这一点.

定理 4.6.3(奇异值分解 SVD) 任意 $m\times n$ 阶实矩阵 $\boldsymbol{A}$ 均可分解成 $\boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^{\mathrm{T}}$ 的形式，其中 $\boldsymbol{U},\boldsymbol{V}$ 为 m 阶和 n 阶正交矩阵，$\boldsymbol{\Sigma}=\mathrm{diag}\,(\sigma_1,\cdots,\sigma_r,0,\cdots,0)_{m\times n}$，$r$ 为矩阵的秩，且可保证 $\sigma_1\geqslant\sigma_2\geqslant\cdots\geqslant\sigma_r>0$，$\sigma_1,\sigma_2,\cdots,\sigma_r$ 称为奇异值.

证明 构造 $m+n$ 阶实对称矩阵如下

$$\boldsymbol{C}=\begin{pmatrix}\boldsymbol{O} & \boldsymbol{A} \\ \boldsymbol{A}^{\mathrm{T}} & \boldsymbol{O}\end{pmatrix},$$

故存在 $m+n$ 个正交的特征向量. 设 $\boldsymbol{w}$ 为属于特征值 $\lambda\neq 0$ 的特征向量，即

$$\boldsymbol{C}\boldsymbol{w}=\lambda\boldsymbol{w}.$$

令

$$\boldsymbol{w}=\begin{pmatrix}\boldsymbol{\xi} \\ \boldsymbol{\eta}\end{pmatrix},$$

其中 ξ 为 m 维向量, $\boldsymbol{\eta}$ 为 n 维向量, 则有

$$\boldsymbol{A}\boldsymbol{\eta}=\lambda\boldsymbol{\xi},\quad \boldsymbol{A}^{\mathrm{T}}\boldsymbol{\xi}=\lambda\eta,\quad \boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\eta}=\lambda\boldsymbol{\eta}^{\mathrm{T}}\boldsymbol{\eta}=\lambda\boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{\xi}.$$

进一步有 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{\eta}=\lambda^2\boldsymbol{\eta}$，$\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\xi}=\lambda^2\boldsymbol{\xi}$.

显然 $\boldsymbol{\eta}$ 是 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 的属于 λ^2 的特征向量，而 $\boldsymbol{\xi}$ 是 $\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}$ 的属于 λ^2 的特征向量.

考虑 $\boldsymbol{v}=\begin{pmatrix}-\boldsymbol{\xi}\\ \boldsymbol{\eta}\end{pmatrix}$ 有 $\boldsymbol{C}\boldsymbol{v}=-\lambda\boldsymbol{v}$，故 $\boldsymbol{C}$ 的非零特征值成对出现.

设 $\mathrm{r}(\boldsymbol{A})=r$，则 $\mathrm{r}(\boldsymbol{C})=2r$, 从而 $\boldsymbol{C}$ 的非零特征值为 $\sigma_1\geqslant\cdots\geqslant\sigma_r>0,-\sigma_1,\cdots,-\sigma_r$.

由于 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 是实对称矩阵，且易知 $\mathrm{r}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})=r$，故可设属于特征值 $\sigma_1^2,\cdots,\sigma_r^2,0,\cdots,0$ 的特征向量 $\boldsymbol{\eta}_1,\cdots,\boldsymbol{\eta}_r,\cdots,\boldsymbol{\eta}_n$ 为标准正交的. 令

$$\boldsymbol{V}=(\boldsymbol{\eta}_1\ \ \boldsymbol{\eta}_2\ \ \cdots\boldsymbol{\eta}_n),$$

则 $\boldsymbol{V}$ 为 n 阶正交矩阵，并有

$$\boldsymbol{A}\boldsymbol{V}=(\sigma_1\boldsymbol{\xi}_1\ \ \cdots\ \ \sigma_r\boldsymbol{\xi}_r\ \ \boldsymbol{\theta}\ \ \cdots\ \ \boldsymbol{\theta})$$

又

$$\boldsymbol{V}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{V}=\boldsymbol{V}^{\mathrm{T}}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})\boldsymbol{V}=\mathrm{diag}(\sigma_1^2,\cdots,\sigma_r^2,0,\cdots,0),$$

可得 $\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\cdots,\boldsymbol{\xi}_r$ 为标准正交的. 进一步扩充成 m 个向量构成的标准正交向量组 $\boldsymbol{\xi}_1,\cdots,\boldsymbol{\xi}_r,\cdots,\boldsymbol{\xi}_m$，令 $\boldsymbol{U}=(\boldsymbol{\xi}_1\ \ \cdots\ \ \boldsymbol{\xi}_r\ \ \cdots\ \ \boldsymbol{\xi}_m)$, 则 $\boldsymbol{U}$ 为 m 阶正交矩阵，并有

$$\boldsymbol{A}\boldsymbol{V}=\boldsymbol{U}\boldsymbol{\Sigma},$$

故得

$$\boldsymbol{A}=\boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^{\mathrm{T}}.$$

□

4.7* 应用于解常系数线性齐次微分方程组

矩阵对角化的化简可用于求解微分方程组.

定义 4.7.1 (常系数线性齐次微分方程组)　称

$$\begin{cases} y_1'(x)=a_{11}y_1(x)+\cdots+a_{1n}y_n(x),\\ y_2'(x)=a_{21}y_1(x)+\cdots+a_{2n}y_n(x),\\ \cdots\cdots\\ y_n'(x)=a_{n1}y_1(x)+\cdots+a_{nn}y_n(x)\end{cases}$$

为**常系数线性齐次微分方程组**.

例 4.7.1　解一阶常系数线性微分方程组

$$\begin{cases} y_1'=3y_1,\\ y_2'=-2y_2,\\ y_3'=4y_3.\end{cases}$$

解 对方程组的各个方程 $y_1' = 3y_1, y_2' = -2y_2, y_3' = 4y_3$ 分别求解，可得

$$y_1 = C_1e^{3x}, \quad y_2 = C_2e^{-2x}, \quad y_3 = C_3e^{4x},$$

故方程组的通解为

$$\begin{aligned}\boldsymbol{y} &= \begin{pmatrix} C_1e^{3x} \\ C_2e^{-2x} \\ C_3e^{4x} \end{pmatrix} \\ &= C_1e^{3x}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + C_2e^{-2x}\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + C_3e^{4x}\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \\ &= C_1e^{3x}\boldsymbol{e}_1 + C_2e^{-2x}\boldsymbol{e}_2 + C_3e^{4x}\boldsymbol{e}_3.\end{aligned}$$

□

定理 4.7.1 *常系数线性齐次微分方程组 $\boldsymbol{y}' = \boldsymbol{A}\boldsymbol{y}$($\boldsymbol{A}$ 为实对称矩阵) 有如下通解: $\boldsymbol{y} = C_1e^{\lambda_1 x}\boldsymbol{p}_1 + C_2e^{\lambda_2 x}\boldsymbol{p}_2 + \cdots + C_ne^{\lambda_n x}\boldsymbol{p}_n$，其中 $C_1, C_2, \cdots, C_n$ 为任意常数，$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \mathrm{diag}\,(\lambda_1, \lambda_2, \cdots, \lambda_n), \boldsymbol{P} = (\boldsymbol{p}_1\ \boldsymbol{p}_2\ \cdots\ \boldsymbol{p}_n)$.*

证明 因为 $\boldsymbol{A}$ 为实对称矩阵，故存在可逆矩阵 $\boldsymbol{P}$ 将 $\boldsymbol{A}$ 对角化，对角矩阵的对角线为 $\boldsymbol{A}$ 的特征值，即

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \mathrm{diag}\,(\lambda_1, \lambda_2, \cdots, \lambda_n) = \boldsymbol{\Lambda}, \ \ \boldsymbol{P} = (\boldsymbol{p}_1\ \boldsymbol{p}_2\ \cdots\ \boldsymbol{p}_n).$$

令 $\boldsymbol{z} = \boldsymbol{P}^{-1}\boldsymbol{y}$，则 $\boldsymbol{y} = \boldsymbol{P}\boldsymbol{z}$ 于是

$$\boldsymbol{z}' = (\boldsymbol{P}^{-1}\boldsymbol{y})' = \boldsymbol{P}^{-1}\boldsymbol{y}' = \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{y} = (\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P})\boldsymbol{P}^{-1}\boldsymbol{y} = \boldsymbol{\Lambda}\boldsymbol{z},$$

即

$$\begin{cases} z_1' = \lambda_1 z_1, \\ z_2' = \lambda_2 z_2, \\ \cdots\cdots \\ z_n' = \lambda_n z_n. \end{cases}$$

分别解方程 $z_i' = \lambda_i z_i, i = 1, 2, \cdots, n$ 得 $z_i = C_ie^{\lambda_i x}$，其中 $C_1, C_2, \cdots, C_n$ 为任意常数. 故有通解

$$\boldsymbol{y} = \boldsymbol{P}\boldsymbol{z} = (\boldsymbol{p}_1\ \boldsymbol{p}_2\ \cdots\ \boldsymbol{p}_n)\begin{pmatrix} C_1e^{\lambda_1 x} \\ \vdots \\ C_ne^{\lambda_n x} \end{pmatrix} = C_1e^{\lambda_1 x}\boldsymbol{p}_1 + C_2e^{\lambda_2 x}\boldsymbol{p}_2 + \cdots + C_ne^{\lambda_n x}\boldsymbol{p}_n.$$

□

例 4.7.2 解一阶常系数线性微分方程组

$$\begin{cases} y_1' = 2y_1 + y_2 + 2y_3, \\ y_2' = y_1 + 3y_2 + y_3, \\ y_3' = 2y_1 + y_2 + 2y_3. \end{cases}$$

解　令

$$\boldsymbol{A}=\begin{pmatrix}2&1&2\\1&3&1\\2&1&2\end{pmatrix},$$

在例 4.5.1 中已求得特征值为 $\lambda_1=5,\lambda_2=2,\lambda_3=0$，对应的特征向量分别为

$$\boldsymbol{p}_1=\begin{pmatrix}1\\1\\1\end{pmatrix},\quad \boldsymbol{p}_2=\begin{pmatrix}1\\-2\\1\end{pmatrix},\quad \boldsymbol{p}_3=\begin{pmatrix}-1\\0\\1\end{pmatrix}.$$

故方程组的通解为

$$\boldsymbol{y}=C_1e^{5x}\boldsymbol{p}_1+C_2e^{2x}\boldsymbol{p}_2+C_3\boldsymbol{p}_3=C_1e^{5x}\begin{pmatrix}1\\1\\1\end{pmatrix}+C_2e^{2x}\begin{pmatrix}1\\-2\\1\end{pmatrix}+C_3\begin{pmatrix}-1\\0\\1\end{pmatrix}.$$

□

习　题　四

1. 求下列矩阵的特征值与特征向量:

(1) $\begin{pmatrix}-2&1&-2\\-2&1&-2\\-14&8&3\end{pmatrix}$;　　(2) $\begin{pmatrix}4&-3&6\\0&1&0\\-3&3&-5\end{pmatrix}$;

(3) $\begin{pmatrix}6&-1&5\\-4&3&-5\\-7&-3&-2\end{pmatrix}$;　　(4) $\begin{pmatrix}7&-4&-4\\3&-1&-1\\5&-4&-2\end{pmatrix}$.

2. 求下列 n 阶矩阵的特征值与特征向量 ($n\geqslant 2$):

(1) $\begin{pmatrix}0&1&1&\cdots&1\\1&0&1&\cdots&1\\1&1&0&\cdots&1\\\vdots&\vdots&\vdots&&\vdots\\1&1&1&\cdots&0\end{pmatrix}$;　　(2) $\begin{pmatrix}0&\cdots&0&1\\0&\cdots&1&0\\\vdots&&\vdots&\vdots\\1&\cdots&0&0\end{pmatrix}$.

3. 求矩阵

$$\begin{pmatrix}4&-3&-1\\4&-3&-3a-1\\-1&1&a+3\end{pmatrix}$$

的特征值与特征向量.

4. 证明方阵 $\boldsymbol{A}$ 与 $\boldsymbol{A}^{\mathrm{T}}$ 有相同的特征值. 进一步, 是否对应相同的特征向量, 若成立则给出证明, 否则给出反例.

2(2). 令 $A_n=\left|\lambda E-A\right|$, 则当 $n=2k$ 时,

$$A_n=\begin{vmatrix}\lambda&&&&&-1\\&\ddots&&&\iddots&\\&&\lambda&-1&&\\&&-1&\lambda&&\\&\iddots&&&\ddots&\\-1&&&&&\lambda\end{vmatrix}$$

故, $A_2=\lambda^2-1$,

$$A_n=\left(\lambda^2-1\right)A_{n-2}=\cdots=\left(\lambda^2-1\right)^{k-1}A_2=\left(\lambda^2-1\right)^{k-1}\left(\lambda^2-1\right)=(\lambda-1)^k(\lambda+1)^k$$

当 $n=2k+1$ 时,

$$A_n=A_{2k+1}=\begin{vmatrix}\lambda&&&&&-1\\&\ddots&&&\iddots&\\&&\lambda&&-1&\\&&&\lambda-1&&\\&&-1&&\lambda&\\&\iddots&&&&\ddots&\\-1&&&&&\lambda\end{vmatrix}$$

$$=(\lambda-1)(-1)^{(k+1)+(k+1)}A_{2k}$$

$$=(\lambda-1)^{k+1}(\lambda+1)^k$$

6. 由于

$$\lambda_0 E_{m+n}-\begin{pmatrix} O & A \\ A^T & O \end{pmatrix}=\begin{pmatrix} \lambda_0 E_m & -A \\ -A^T & \lambda_0 E_n \end{pmatrix}$$

$$\begin{pmatrix} E_m & O \\ \frac{1}{\lambda_0}A^T & E_n \end{pmatrix}\begin{pmatrix} \lambda_0 E_m & -A \\ -A^T & \lambda_0 E_n \end{pmatrix}=\begin{pmatrix} \lambda_0 E_m & -A \\ O & \lambda_0 E_n-\frac{1}{\lambda_0}A^T A \end{pmatrix}$$

故

$$\begin{vmatrix} E_m & O \\ \frac{1}{\lambda_0}A^T & E_n \end{vmatrix}\begin{vmatrix} \lambda_0 E_m & -A \\ -A^T & \lambda_0 E_n \end{vmatrix}=\begin{vmatrix} \lambda_0 E_m & -A \\ O & \lambda_0 E_n-\frac{1}{\lambda_0}A^T A \end{vmatrix}.$$

亦即

$$\left|\lambda_0 E_{m+n}-\begin{pmatrix} O & A \\ A^T & O \end{pmatrix}\right|=\begin{vmatrix} \lambda_0 E_m & -A \\ -A^T & \lambda_0 E_n \end{vmatrix}$$
$$=\begin{vmatrix} \lambda_0 E_m & -A \\ O & \lambda_0 E_n-\frac{1}{\lambda_0}A^T A \end{vmatrix}$$
$$=\left|\lambda_0 E_m\right|\left|\lambda_0 E_n-\frac{1}{\lambda_0}A^T A\right|$$
$$=\lambda_0^m\cdot\frac{1}{\lambda_0^n}\left|\lambda_0^2 E_n-A^T A\right|=0,$$

故

$$\left|\lambda_0^2 E_n-A^T A\right|=0.$$

5. 设 $\xi\neq 0$, $\eta\neq 0$ 满足 $A\xi=\lambda_0\xi$, $B\eta=\gamma_0\eta$. 则

$$\begin{pmatrix} \xi \\ 0 \end{pmatrix}\neq 0,\quad \begin{pmatrix} 0 \\ \eta \end{pmatrix}\neq 0,$$

且

$$\begin{pmatrix} A & O \\ O & B \end{pmatrix}\begin{pmatrix} \xi \\ 0 \end{pmatrix}=\begin{pmatrix} A\xi \\ 0 \end{pmatrix}=\begin{pmatrix} \lambda_0\xi \\ 0 \end{pmatrix}=\lambda_0\begin{pmatrix} \xi \\ 0 \end{pmatrix},$$

$$\begin{pmatrix} A & O \\ O & B \end{pmatrix}\begin{pmatrix} 0 \\ \eta \end{pmatrix}=\begin{pmatrix} 0 \\ B\eta \end{pmatrix}=\begin{pmatrix} 0 \\ \gamma_0\eta \end{pmatrix}=\gamma_0\begin{pmatrix} 0 \\ \eta \end{pmatrix}.$$

这表明 λ_0,γ_0 是矩阵

$$\begin{pmatrix} A & O \\ O & B \end{pmatrix}$$

的特征值。

5. 设 λ_0 为 $\boldsymbol{A}$ 的特征值, γ_0 为 $\boldsymbol{B}$ 的特征值，证明 λ_0 与 γ_0 均为 $\begin{pmatrix} \boldsymbol{A} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{B} \end{pmatrix}$ 的特征值.

6. 设 $\boldsymbol{A}\in\mathbf{R}^{m\times n}$, λ_0 为 $\begin{pmatrix} \boldsymbol{O} & \boldsymbol{A} \\ \boldsymbol{A}^{\mathrm{T}} & \boldsymbol{O} \end{pmatrix}$ 的非零特征值, 证明 λ_0^2 为 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 的特征值.

7. 设 $\boldsymbol{\xi}$ 既是矩阵 $\boldsymbol{A}$ 的特征向量，也是矩阵 $\boldsymbol{B}$ 的特征向量，证明: $\boldsymbol{\xi}$ 也是 $\boldsymbol{A}+\boldsymbol{B}$ 和 $\boldsymbol{AB}$ 的特征向量.

8. 已知矩阵

$$\begin{pmatrix} -4 & 2 & -1 \\ 1 & a & -4 \\ -3 & b & -2 \end{pmatrix}$$

有特征值 $-1,-2,-3$，求 a,b 的值.

9. 由于 A 的全部特征值是 $1,2,-3$，故存在可逆阵 P 使得

$$A=P^{-1}\begin{pmatrix} 1 & & \\ & 2 & \\ & & -3 \end{pmatrix}P.$$

于是,

$$A^{-1}=P^{-1}\begin{pmatrix} 1 & & \\ & 2 & \\ & & -3 \end{pmatrix}^{-1}P=P^{-1}\begin{pmatrix} 1 & & \\ & \frac{1}{2} & \\ & & -\frac{1}{3} \end{pmatrix}P,$$

$$A^{-1}+A=P^{-1}\begin{pmatrix} 1+1 & & \\ & \frac{1}{2}+2 & \\ & & -\frac{1}{3}-3 \end{pmatrix}P.$$

这说明 $A^{-1}+A$ 的特征值是 2, 5/2, −10/3.

9. 已知 3 阶矩阵 $\boldsymbol{A}$ 有特征值 $1,2,-3$，求 $\boldsymbol{A}^{-1}+\boldsymbol{A}$ 的特征值.

10. 设 3 阶矩阵 $\boldsymbol{A}$ 有 3 个不同的特征值 $\lambda_1,\lambda_2,\lambda_3$，对应特征向量为 $\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\boldsymbol{\xi}_3$, $\boldsymbol{B}=\boldsymbol{A}^3-(\lambda_1+\lambda_2+\lambda_3)\boldsymbol{A}^2+(\lambda_1\lambda_2+\lambda_2\lambda_3+\lambda_3\lambda_1)\boldsymbol{A}$，证明 $\boldsymbol{\xi}_1,\boldsymbol{\xi}_2,\boldsymbol{\xi}_3$ 的任意非零线性组合都是 $\boldsymbol{B}$ 的特征向量.

11. 已知 3 阶矩阵 $\boldsymbol{A}$ 有特征值 $2,3,-3$，对应特征向量为 $\boldsymbol{\xi}_1=(1,1,2)^{\mathrm{T}}$, $\boldsymbol{\xi}_2=(2,0,1)^{\mathrm{T}}$, $\boldsymbol{\xi}_3=(1,-2,-3)^{\mathrm{T}}$，求 $\boldsymbol{A}$.

12. 证明: 若 $\boldsymbol{A},\boldsymbol{B}$ 为同阶方阵, 则 $\boldsymbol{AB}$ 和 $\boldsymbol{BA}$ 有相同的特征值.

13. 设 $\boldsymbol{\xi}_1$ 与 $\boldsymbol{\xi}_2$ 为 $\boldsymbol{A}$ 的特征向量，且 $\boldsymbol{\xi}_1+\boldsymbol{\xi}_2$ 仍然是 $\boldsymbol{A}$ 的特征向量，证明 $\boldsymbol{\xi}_1,\boldsymbol{\xi}_2$ 属于相同的特征值.

12. 由于

$$\begin{pmatrix} E_n & -A \\ O & E_n \end{pmatrix}\begin{pmatrix} AB & O \\ B & O \end{pmatrix}\begin{pmatrix} E_n & -A \\ O & E_n \end{pmatrix}^{-1}=\begin{pmatrix} O & O \\ B & BA \end{pmatrix},$$

故

$$\left|\lambda E_{2n}-\begin{pmatrix} AB & O \\ B & O \end{pmatrix}\right|=\left|\lambda E_{2n}-\begin{pmatrix} O & O \\ B & BA \end{pmatrix}\right|.$$

亦即,

$$\begin{vmatrix} \lambda E_n-AB & O \\ -B & \lambda E_n \end{vmatrix}=\begin{vmatrix} \lambda E_n & O \\ -B & \lambda E_n-BA \end{vmatrix},$$

$$\lambda^n\left|\lambda E_n-AB\right|=\lambda^n\left|\lambda E_n-BA\right|.$$

14. 判断下列矩阵是否可对角化，若可以则对角化，否则说明原因:

(1) $\begin{pmatrix} -2 & 1 & -1 \\ 0 & -2 & -1 \\ 0 & 0 & -2 \end{pmatrix}$; (2) $\begin{pmatrix} 4 & -2 & 1 \\ 5 & -3 & -2 \\ 2 & -2 & 0 \end{pmatrix}$;

(3) $\begin{pmatrix} 5 & -2 & 1 \\ 1 & 2 & -1 \\ -3 & 3 & -1 \end{pmatrix}$; (4) $\begin{pmatrix} 1 & -2 & 3 \\ 2 & -3 & 3 \\ 2 & -2 & 2 \end{pmatrix}$.

15. 已知矩阵

$$\boldsymbol{A}=\begin{pmatrix} 6 & -5 & -3 \\ 10 & -9 & -6 \\ -6 & 6 & 5 \end{pmatrix},$$

求 $\boldsymbol{A}^n$.

16. 已知矩阵 $\boldsymbol{A}=\begin{pmatrix} 3 & 1 & 0 \\ 1 & 3 & 0 \\ 1 & b-1 & a \end{pmatrix}$ 与 $\boldsymbol{B}=\begin{pmatrix} 4 & 0 & 0 \\ 0 & 4 & b \\ 0 & 2 & 2a \end{pmatrix}$ 相似，求 $\boldsymbol{P}$，使得 $\boldsymbol{P}^{-1}\boldsymbol{AP}=\boldsymbol{B}$.

17. 已知 n 阶矩阵 $\boldsymbol{A}$ 有关系式 $\boldsymbol{A}^2-3\boldsymbol{A}+2\boldsymbol{E}=\boldsymbol{O}$，试证 $\boldsymbol{A}$ 可对角化.

18. 设 $\boldsymbol{A}$ 为 n 阶方阵, $\mathrm{r}(\boldsymbol{A})=r_1$, $\mathrm{r}(\boldsymbol{A}+\boldsymbol{E})=r_2$, $\mathrm{r}(\boldsymbol{A}+2\boldsymbol{E})=r_3$，且 $r_1+r_2+r_3=2n$，证明 $\boldsymbol{A}$ 可对角化.

19. 求向量 $\boldsymbol{\alpha}_1=(1,2,4,-2)^{\mathrm{T}}$, $\boldsymbol{\alpha}_2=(1,-1,1,-1)^{\mathrm{T}}$ 的内积及各个向量的长度.

18. 若 $A=-kE$, 结论显然成立。故可设 $A\neq -kE$, $k=0,1,2$.

此时，$r(A+kE)=r_{k+1}>0$, $k=0,1,2$.

若 $r_1=n$, 则有 $0<r_2<n$, $0<r_3<n$.

这表明 $-1,-2$ 是 A 的特征值，且

A 有 $n-r_2$ 个属于特征值 -1 的线性无关的特征向量,

有 $n-r_3$ 个属于特征值 -2 的线性无关的特征向量。

由于 $(n-r_2)+(n-r_3)=r_1=n$, A 有 n 个线性无关的特征向量，从而 A 可以对角化。

同理可证，当 $r_2=n$ 或者 $r_3=n$ 时，A 可以对角化。

若 $\forall k=0,1,2$, $0<r_{k+1}<n$, 则 $0,-1,-2$ 是 A 的特征值。

由于

$$(n-r_1)+(n-r_2)+(n-r_3)=n,$$

故 A 有 n 个线性无关的特征向量，从而 A 可以对角化。

20. 求向量 $\boldsymbol{\alpha}_1=(1,2,2)^{\mathrm{T}},\boldsymbol{\alpha}_2=(2,-1,2)^{\mathrm{T}}$ 之间的夹角.

21. 证明: $\|\boldsymbol{\beta}_1+\boldsymbol{\beta}_2\|^2+\|\boldsymbol{\beta}_1-\boldsymbol{\beta}_2\|^2=2(\|\boldsymbol{\beta}_1\|^2+\|\boldsymbol{\beta}_2\|^2)$.

22. 设 $\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_r\}$ 为 n 维正交向量组，$\boldsymbol{Q}\in\mathbf{R}^{n\times n}$ 为正交矩阵，$\boldsymbol{\beta}_i=\boldsymbol{Q}\boldsymbol{\alpha}_i,i=1,2,\cdots,r$，证明 $\{\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_r\}$ 也为正交向量组.

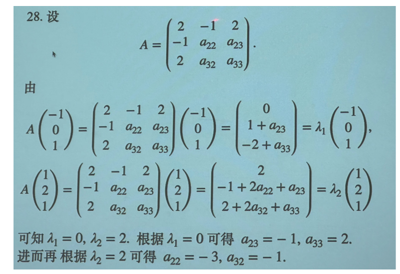
23. 将向量组 $\boldsymbol{\alpha}_1=(1,1,1,0)^{\mathrm{T}},\boldsymbol{\alpha}_2=(2,1,0,2)^{\mathrm{T}},\boldsymbol{\alpha}_3=(0,0,1,1)^{\mathrm{T}}$ 正交化.

24. 将向量组 $\boldsymbol{\alpha}_1=(1,1,0)^{\mathrm{T}},\boldsymbol{\alpha}_2=(0,1,1)^{\mathrm{T}},\boldsymbol{\alpha}_3=(1,0,1)^{\mathrm{T}}$ 标准正交化.

25. 证明: 若 $\boldsymbol{A},\boldsymbol{B}$ 是同阶的正交矩阵，且 $|\boldsymbol{A}|=-|\boldsymbol{B}|$，则有 $|\boldsymbol{A}+\boldsymbol{B}|=0$.

26. 用正交矩阵作为相似变换矩阵将下列实对称矩阵对角化:

(1) $\begin{pmatrix}3&2&1\\2&6&2\\1&2&3\end{pmatrix}$;　　(2) $\begin{pmatrix}2&2&0\\2&0&2\\0&2&2\end{pmatrix}$.


28. 设

$$A=\begin{pmatrix}2&-1&2\\-1&a_{22}&a_{23}\\2&a_{32}&a_{33}\end{pmatrix}.$$

由

$$A\begin{pmatrix}-1\\0\\1\end{pmatrix}=\begin{pmatrix}2&-1&2\\-1&a_{22}&a_{23}\\2&a_{32}&a_{33}\end{pmatrix}\begin{pmatrix}-1\\0\\1\end{pmatrix}=\begin{pmatrix}0\\1+a_{23}\\-2+a_{33}\end{pmatrix}=\lambda_1\begin{pmatrix}-1\\0\\1\end{pmatrix},$$

$$A\begin{pmatrix}1\\2\\1\end{pmatrix}=\begin{pmatrix}2&-1&2\\-1&a_{22}&a_{23}\\2&a_{32}&a_{33}\end{pmatrix}\begin{pmatrix}1\\2\\1\end{pmatrix}=\begin{pmatrix}2\\-1+2a_{22}+a_{23}\\2+2a_{32}+a_{33}\end{pmatrix}=\lambda_2\begin{pmatrix}1\\2\\1\end{pmatrix}$$

可知 $\lambda_1=0,\lambda_2=2$. 根据 $\lambda_1=0$ 可得 $a_{23}=-1,a_{33}=2$.
进而再根据 $\lambda_2=2$ 可得 $a_{22}=-3,a_{32}=-1$.

27. 已知 3 阶实对称矩阵 $\boldsymbol{A}$ 每一行的和均为 3，且其特征值均为正整数，$|\boldsymbol{A}|=3$，求矩阵 $\boldsymbol{A}$.

28. 已知 3 阶实对称矩阵 $\boldsymbol{A}$ 的第一列为 $(2,-1,2)^{\mathrm{T}}$，有两个特征向量为 $(-1,0,1)^{\mathrm{T}}$, $(1,2,1)^{\mathrm{T}}$, 求 $\boldsymbol{A}$.

29*. 设 n 阶实对称矩阵 $\boldsymbol{A}$ 满足 $\boldsymbol{A}^2=\boldsymbol{A}$，并设 $\mathrm{r}(\boldsymbol{A})=r<n$，证明 $\boldsymbol{A}$ 可以表示为 $\boldsymbol{A}=\boldsymbol{U}\boldsymbol{U}^{\mathrm{T}}$，其中 $\boldsymbol{U}\in\mathbf{R}^{n\times r},\boldsymbol{U}^{\mathrm{T}}\boldsymbol{U}=\boldsymbol{E}_r$.

30*. 解一阶常系数线性微分方程组

$$\begin{cases}y_1'=-2y_1+2y_2+y_3,\\y_2'=-y_1+y_2,\\y_3'=3y_1-3y_2+y_3.\end{cases}$$