高等工程數學（7）

廈門大學
XIAMEN UNIVERSITY

信息学院
（国家示范性软件学院）
School of Informatics

黄 煒
博士，副教授
Dr. Wei Huang

JORDAN标准型

矩阵论（7）

厦門大學 XIAMEN UNIVERSITY
信息学院（国家示范性软件学院）School of Informatics
黄 炜 博士·副教授 Dr. Wei Huang

2 方阵的相似化简

本章 1.5 节中已经说明，一个 n 阶方阵可相似对角化的充要条件是有 n 个线性无关的特征向量. 但一般的 n 阶方阵不一定有 n 个线性无关的特征向量，因而就不一定能相似于对角矩阵. 本章将论述复数域上任一 n 阶方阵都可相似于 Jordan 矩阵，它的特殊情形就是对角矩阵，并应用这一结论来证明矩阵论中一个重要定理，即 Cayley-Hamilton 定理. 如果要求限制在酉相似变换，那么任一 n 阶复方阵都相似于上三角矩阵. 对于实数域上 n 阶方阵来说，我们将证明它可正交相似于上 Hessenberg 矩阵.

2.1 Jordan标准型

如何判断一个 n 阶方阵不能相似于对角矩阵？这就要从此方阵的特征值的几何重数与代数重数的关系入手.

定理 2.1-1 任何方阵的特征值的几何重数都不大于其代数重数.

证 设 $\lambda_1,\lambda_2,\cdots,\lambda_s$ 是 n 阶方阵 $\boldsymbol{A}$ 的所有不同特征值，则 $\boldsymbol{A}$ 的特征多项式 $\det(\boldsymbol{A}-\lambda\boldsymbol{I})$ 可以表示为

$$\det(\boldsymbol{A}-\lambda\boldsymbol{I})=(\lambda_1-\lambda)^{n_1}\cdots(\lambda_s-\lambda)^{n_s},$$

且

$$n_1+\cdots+n_s=n,$$

其中 $n_i(i=1,2,\cdots,s)$ 是特征值 λ_i 的代数重数.

不失一般性，就 λ_1 来证明其几何重数不大于 n_1. 设 λ_1 的几何重数是 k，则 $\boldsymbol{A}$ 的属于特征值 λ_1 的线性无关特征向量有 k 个：$\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_k$. 再添加 $n-k$ 个向量 $\boldsymbol{\alpha}_{k+1},\cdots,\boldsymbol{\alpha}_n$ 使向量组 $\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n$ 是 n 维向量空间 $\mathbf{C}^n$ 的一个基.

2.1 Jordan标准型

令 $P=[\alpha_1 \ \cdots \ \alpha_k \ \vdots \ \alpha_{k+1} \ \cdots \ \alpha_n]$，则 P 是 n 阶可逆矩阵，且

$$AP=[A\alpha_1 \ \cdots \ A\alpha_k \ \vdots \ A\alpha_{k+1} \ \cdots \ A\alpha_n]=[\lambda_1\alpha_1 \ \cdots \ \lambda_1\alpha_k \ \vdots \ A\alpha_{k+1} \ \cdots \ A\alpha_n]$$

$$=[\alpha_1 \ \cdots \ \alpha_k \ \vdots \ \alpha_{k+1} \ \cdots \ \alpha_n]\begin{array}{c} \left[\begin{array}{ccc|c} \lambda_1 & & & \\ & \ddots & & A_{12} \\ & & \lambda_1 & \\ \hline & O & & A_{22} \end{array}\right] \begin{array}{l} \left.\begin{array}{c} \\ \\ \\ \end{array}\right\} k \\ \left.\right\} n-k \end{array} \\ \begin{array}{cc} \underbrace{\qquad\qquad}_{k} & \underbrace{\quad}_{n-k} \end{array} \end{array}=P\left[\begin{array}{ccc|c} \lambda_1 & & & \\ & \ddots & & A_{12} \\ & & \lambda_1 & \\ \hline & O & & A_{22} \end{array}\right].$$

因 P 可逆，故

$$P^{-1}AP=\left[\begin{array}{ccc|c} \lambda_1 & & & \\ & \ddots & & A_{12} \\ & & \lambda_1 & \\ \hline & O & & A_{22} \end{array}\right],$$

从而由相似矩阵有相同的特征多项式得

$$\det(A-\lambda I)=(\lambda_1-\lambda)^k\cdot\det(A_{22}-\lambda I_{n-k}).$$

由上式知 $(\lambda_1-\lambda)^k$ 必能整除 $\det(A-\lambda I)$，故 $k\leqslant n_1$，即 λ_1 的几何重数不大于其代数重数. ∎

这个定理表明，只要 A 的一个特征值的几何重数小于其代数重数，则 A 就没有 n 个线性无关的特征向量，因而不可相似于对角矩阵.

2.1 Jordan标准型

例 1 证明 $\boldsymbol{A}=\begin{bmatrix}-3 & 3 & -2\\ -7 & 6 & -3\\ 1 & -1 & 2\end{bmatrix}$不可相似对角化.

证 由于

$$|\boldsymbol{A}-\lambda\boldsymbol{I}|=\begin{vmatrix}-3-\lambda & 3 & -2\\ -7 & 6-\lambda & -3\\ 1 & -1 & 2-\lambda\end{vmatrix}=\begin{vmatrix}2-\lambda & 3 & -2\\ 2-\lambda & 6-\lambda & -3\\ -(2-\lambda) & -1 & 2-\lambda\end{vmatrix}$$

$$=(2-\lambda)\begin{vmatrix}1 & 3 & -2\\ 0 & 3-\lambda & -1\\ 0 & 2 & -\lambda\end{vmatrix}=(1-\lambda)(2-\lambda)^2,$$

所以 $\boldsymbol{A}$ 有两个不同的特征值 1 和 2，且特征值 2 的代数重数是 2.

对特征值 2，由

$$\boldsymbol{A}-2\boldsymbol{I}=\begin{bmatrix}-5 & 3 & -2\\ -7 & 4 & -3\\ 1 & -1 & 0\end{bmatrix},$$

知 $r(\boldsymbol{A}-2\boldsymbol{I})=2$，故特征值 2 的几何重数是 $3-2=1$.

因此，特征值 2 的几何重数小于其代数重数，故 $\boldsymbol{A}$ 不可相似对角化. ∎

2.1 Jordan标准型

定义 2.1-1 形如

$$\begin{bmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{bmatrix}_{r\times r}$$

的 $r(\geqslant 1)$ 阶方阵称为 **Jordan 块**，其中 λ 是实数或复数. Jordan 块的特点是主对角线元素都相同，且次对角线元素规定是 1，其余元素为 0. 由若干个(包括单个)Jordan 块所构成的分块对角矩阵称为 **Jordan 矩阵.**

例如
$$\begin{bmatrix} -2 & 1 & \\ & -2 & 1 \\ & & -2 \end{bmatrix},\quad \begin{bmatrix} 0 & 1 & & \\ & 0 & 1 & \\ & & 0 & 1 \\ & & & 0 \end{bmatrix},\quad \begin{bmatrix} 1-\mathrm{i} & 1 & \\ & 1-\mathrm{i} & 1 \\ & & 1-\mathrm{i} \end{bmatrix}$$

都是 Jordan 块(如果把它们看作是单个 Jordan 块的方阵，则它们也都是 Jordan 矩阵).
而

2.1 Jordan标准型

$$\begin{bmatrix} 1 & & \\ & 2 & \\ & & 1-\mathrm{i} \end{bmatrix},\ \begin{bmatrix} 2 & 1 & \\ & 2 & \\ & & 3 \end{bmatrix},\ \begin{bmatrix} 1 & 1 & & & \\ & 1 & & & \\ & & \mathrm{i} & 1 & \\ & & & \mathrm{i} & 1 \\ & & & & \mathrm{i} \end{bmatrix},\ \begin{bmatrix} 1 & 1 & & & \\ & 1 & & & \\ & & -2 & & \\ & & & -2 & 1 \\ & & & & -2 \end{bmatrix}$$

都是 Jordan 矩阵.

常把 Jordan 矩阵中有相同主对角线元素的 Jordan 块紧靠在一起，这些(包括单个) Jordan 块构成的 Jordan 矩阵称为原 Jordan 矩阵的子 Jordan 矩阵. 于是上述的最后一个 Jordan 矩阵 $\boldsymbol{J}$ 可以看成是由两个子 Jordan 矩阵

$$\boldsymbol{J}_1=\begin{bmatrix} 1 & 1 \\ & 1 \end{bmatrix},\quad \boldsymbol{J}_2=\begin{bmatrix} -2 & & \\ & -2 & 1 \\ & & -2 \end{bmatrix}$$

构成的，其中 $\boldsymbol{J}_1$ 由一个 Jordan 块构成，而 $\boldsymbol{J}_2$ 是由同一个主对角线元素 -2 的两个 Jordan 块构成的，并记为 $\boldsymbol{J}=\mathrm{diag}(\boldsymbol{J}_1,\boldsymbol{J}_2)$.

对角矩阵可以看成是由 1 阶 Jordan 块所构成的 Jordan 矩阵，所以 Jordan 矩阵也包含了对角矩阵.

2.1 Jordan标准型

本节的目的是论证复数域上任一 n 阶方阵都相似于一个 Jordan 矩阵，其中子 Jordan 矩阵的个数等于该方阵的不同特征值的个数，而 Jordan 块的个数等于该方阵的线性无关的特征向量个数.

以 3 阶方阵 $\boldsymbol{A}$ 为例，讨论 $\boldsymbol{A}$ 相似于 Jordan 矩阵

$$\boldsymbol{J}=\begin{bmatrix}\lambda_0 & 1 & \\ & \lambda_0 & 1 \\ & & \lambda_0\end{bmatrix}$$

的问题.

设有一个可逆矩阵 $\boldsymbol{P}=[\boldsymbol{x}_1 \quad \boldsymbol{x}_2 \quad \boldsymbol{x}_3]$ 使 $\boldsymbol{P}^{-1}\boldsymbol{AP}=\boldsymbol{J}$，那么如何确定 λ_0 和可逆矩阵 $\boldsymbol{P}$ 呢？首先由本章 1.5 节中定理 1.5-5 知，$\boldsymbol{A}$ 与 $\boldsymbol{J}$ 有相同的特征值. 显然 $\boldsymbol{J}$ 的特征值是 λ_0，λ_0，λ_0，即 λ_0 是 $\boldsymbol{J}$ 的 3 重特征值，因而 λ_0 也应是 $\boldsymbol{A}$ 的 3 重特征值. 其次，由 $\boldsymbol{P}^{-1}\boldsymbol{AP}=\boldsymbol{J}$ 得 $\boldsymbol{AP}=\boldsymbol{PJ}$，即有

$$\boldsymbol{AP}=\boldsymbol{A}[\boldsymbol{x}_1 \vdots \boldsymbol{x}_2 \vdots \boldsymbol{x}_3]=[\boldsymbol{Ax}_1 \vdots \boldsymbol{Ax}_2 \vdots \boldsymbol{Ax}_3]=[\boldsymbol{x}_1 \vdots \boldsymbol{x}_2 \vdots \boldsymbol{x}_3]\begin{bmatrix}\lambda_0 & 1 & \\ & \lambda_0 & 1 \\ & & \lambda_0\end{bmatrix}$$

$$=[\lambda_0\boldsymbol{x}_1 \vdots \lambda_0\boldsymbol{x}_2+\boldsymbol{x}_1 \vdots \lambda_0\boldsymbol{x}_3+\boldsymbol{x}_2].$$

2.1 Jordan标准型

对 $\lambda_1=-3$，解线性方程组 $(\boldsymbol{A}-(-3)\boldsymbol{I})\boldsymbol{x}=\boldsymbol{0}$，即

$$\begin{bmatrix} 3 & 1 & 1 & -1 \\ 1 & 3 & -1 & 1 \\ 1 & -1 & 3 & 1 \\ -1 & 1 & 1 & 3 \end{bmatrix}\boldsymbol{x}=\boldsymbol{0},$$

得其一个基础解系 $\boldsymbol{x}_1=[1,-1,-1,1]^{\mathrm{T}}$，单位化后得

$$\boldsymbol{\varepsilon}_1=\left[\frac{1}{2},-\frac{1}{2},-\frac{1}{2},\frac{1}{2}\right]^{\mathrm{T}}.$$

对 $\lambda_2=\lambda_3=\lambda_4=1$，解线性方程组 $(\boldsymbol{A}-1\cdot\boldsymbol{I})\boldsymbol{x}=\boldsymbol{0}$，即

$$\begin{bmatrix} -1 & 1 & 1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & -1 & 1 \\ -1 & 1 & 1 & -1 \end{bmatrix}\boldsymbol{x}=\boldsymbol{0},$$

于是，$\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{x}_3$ 应满足关系式

$$\begin{cases} \boldsymbol{A}\boldsymbol{x}_1=\lambda_0\boldsymbol{x}_1, \\ \boldsymbol{A}\boldsymbol{x}_2=\lambda_0\boldsymbol{x}_2+\boldsymbol{x}_1, \\ \boldsymbol{A}\boldsymbol{x}_3=\lambda_0\boldsymbol{x}_3+\boldsymbol{x}_2. \end{cases} \tag{2.1-1}$$

(2.1-1)式可以改写成下式：

2.1 Jordan标准型

$$\begin{cases}(\boldsymbol{A}-\lambda_0\boldsymbol{I})\boldsymbol{x}_1=\boldsymbol{0},\\(\boldsymbol{A}-\lambda_0\boldsymbol{I})\boldsymbol{x}_2=\boldsymbol{x}_1,\\(\boldsymbol{A}-\lambda_0\boldsymbol{I})\boldsymbol{x}_3=\boldsymbol{x}_2.\end{cases}\tag{2.1-2}$$

由于 $\boldsymbol{P}$ 可逆，所以 $\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{x}_3$ 线性无关，故 $\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{x}_3$ 均是非零向量. 从而由(2.1-2)式的第一式知，$\boldsymbol{x}_1$ 是 $\boldsymbol{A}$ 的属于特征值 λ_0 的特征向量. 而 $\boldsymbol{x}_2,\boldsymbol{x}_3$ 均不是 $\boldsymbol{A}$ 的特征向量(因为(2.1-2)式的第二和第三式之右端向量都是非零向量). 为了由 $\boldsymbol{x}_1$ 通过(2.1-2)式的第二式解出 $\boldsymbol{x}_2$，由本章 1.4 节中定理 1.4-1 知，此非齐次线性方程组的增广矩阵的秩应等于其系数矩阵的秩，

$$r([\boldsymbol{A}-\lambda_0\boldsymbol{I} \,\vdots\, \boldsymbol{x}_1])=r(\boldsymbol{A}-\lambda_0\boldsymbol{I}).\tag{2.1-3}$$

这就是说，为了能从 $\boldsymbol{x}_1$ 得到 $\boldsymbol{x}_2$，$\boldsymbol{A}$ 的对应于特征值 λ_0 的特征向量 $\boldsymbol{x}_1$ 不能随便选取，而必须满足条件式(2.1-3).

此外，由于 $\boldsymbol{x}_1$ 满足(2.1-2)式的第一式，所以由(2.1-2)式的第二式知，当 $\boldsymbol{x}_2$ 是这个非齐次线性方程组的解时，它还应满足

$$(\boldsymbol{A}-\lambda_0\boldsymbol{I})^2\boldsymbol{x}_2=\boldsymbol{0},\quad 但(\boldsymbol{A}-\lambda_0\boldsymbol{I})\boldsymbol{x}_2\neq\boldsymbol{0}.\tag{2.1-4}$$

类似地，要从 $\boldsymbol{x}_2$ 通过(2.1-2)式的第三式解出 $\boldsymbol{x}_3$，$\boldsymbol{x}_2$ 不能是满足(2.1-4)式的任一解，它还必须满足以下条件

$$r([\boldsymbol{A}-\lambda_0\boldsymbol{I} \,\vdots\, \boldsymbol{x}_2])=r(\boldsymbol{A}-\lambda_0\boldsymbol{I}),\tag{2.1-5}$$

2.1 Jordan标准型

并且此时 $\boldsymbol{x}_3$ 还满足

$$(\boldsymbol{A}-\lambda_0\boldsymbol{I})^3\boldsymbol{x}_3=\boldsymbol{0},\quad 但(\boldsymbol{A}-\lambda_0\boldsymbol{I})^2\boldsymbol{x}_3\neq\boldsymbol{0}. \tag{2.1-6}$$

再则，容易验证，由解线性方程组(2.1-2)所得的非零向量 $\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{x}_3$ 必定线性无关. 事实上，设

$$k_1\boldsymbol{x}_1+k_2\boldsymbol{x}_2+k_3\boldsymbol{x}_3=\boldsymbol{0}, \tag{2.1-7}$$

用 $(\boldsymbol{A}-\lambda_0\boldsymbol{I})^2$ 左乘(2.1-7)式两边，得

$$\begin{aligned}(\boldsymbol{A}-\lambda_0\boldsymbol{I})^2(k_1\boldsymbol{x}_1+k_2\boldsymbol{x}_2+k_3\boldsymbol{x}_3)&=k_1(\boldsymbol{A}-\lambda_0\boldsymbol{I})^2\boldsymbol{x}_1+k_2(\boldsymbol{A}-\lambda_0\boldsymbol{I})^2\boldsymbol{x}_2+k_3(\boldsymbol{A}-\lambda_0\boldsymbol{I})^2\boldsymbol{x}_3\\&=k_3(\boldsymbol{A}-\lambda_0\boldsymbol{I})^2\boldsymbol{x}_3=\boldsymbol{0}.\end{aligned}$$

但 $(\boldsymbol{A}-\lambda_0\boldsymbol{I})^2\boldsymbol{x}_3\neq\boldsymbol{0}$，故 $k_3=0$，于是(2.1-7)式成为

$$k_1\boldsymbol{x}_1+k_2\boldsymbol{x}_2=\boldsymbol{0}.$$

用 $(\boldsymbol{A}-\lambda_0\boldsymbol{I})$ 左乘上式两边得

$$(\boldsymbol{A}-\lambda_0\boldsymbol{I})(k_1\boldsymbol{x}_1+k_2\boldsymbol{x}_2)=k_1(\boldsymbol{A}-\lambda_0\boldsymbol{I})\boldsymbol{x}_1+k_2(\boldsymbol{A}-\lambda_0\boldsymbol{I})\boldsymbol{x}_2=k_2(\boldsymbol{A}-\lambda_0\boldsymbol{I})\boldsymbol{x}_2=\boldsymbol{0}.$$

但 $(\boldsymbol{A}-\lambda_0\boldsymbol{I})\boldsymbol{x}_2\neq\boldsymbol{0}$. 故 $k_2=0$，从而(2.1-7)式又成为

$$k_1\boldsymbol{x}_1=\boldsymbol{0}.$$

因 $\boldsymbol{x}_1\neq\boldsymbol{0}$，故 $k_1=0$. 因此 $k_1=k_2=k_3=0$，故 $\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{x}_3$ 线性无关.

这样一来，将 $\boldsymbol{A}$ 相似化简为 $\boldsymbol{J}$ 的相似变换矩阵 $\boldsymbol{P}$ 可以通过求解线性方程组(2.1-2)得到.

2.1 Jordan标准型

从(2.1-2)式看出，它所包含的三个线性方程组，其系数矩阵均为 $\boldsymbol{A}-\lambda_0\boldsymbol{I}$，只是右端向量不同，因此可以把它们统一考虑为

$$(\boldsymbol{A}-\lambda_0\boldsymbol{I})\boldsymbol{x}=\boldsymbol{y}, \tag{2.1-8}$$

取 $\boldsymbol{y}=\boldsymbol{0}$ 即为第一个线性方程组，取 $\boldsymbol{y}=\boldsymbol{x}_1$ 即为第二个线性方程组，取 $\boldsymbol{y}=\boldsymbol{x}_2$ 便是第三个线性方程组.

由于 $\boldsymbol{y}$ 可能为非零向量，所以一般要对线性方程组(2.1-8)的增广矩阵 $[\boldsymbol{A}-\lambda_0\boldsymbol{I} \mid \boldsymbol{y}]$ 施以行初等变换，找出 $\boldsymbol{y}$ 能使 $r([\boldsymbol{A}-\lambda_0\boldsymbol{I} \mid \boldsymbol{y}])=r(\boldsymbol{A}-\lambda_0\boldsymbol{I})$ 所应满足的条件，选取这样的 $\boldsymbol{y}$ 才能求解 $\boldsymbol{x}$.

例 2 设 $\boldsymbol{A}=\begin{bmatrix} 3 & 1 & 1 \\ 0 & 2 & 0 \\ -1 & 0 & 1 \end{bmatrix}$，求可逆矩阵 $\boldsymbol{P}$ 使 $\boldsymbol{P}^{-1}\boldsymbol{AP}$ 为 Jordan 矩阵.

解 由于

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

故 $\boldsymbol{A}$ 有三重特征值 2. 又因 $\boldsymbol{A}-2\boldsymbol{I}=\begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & -1 \end{bmatrix}$，

2.1 Jordan标准型

故 $r(\boldsymbol{A}-2\boldsymbol{I})=2$，所以特征值 2 的几何重数是 $3-2=1$，即对应于 2 只有一个线性无关的特征向量. 因此 $\boldsymbol{A}$ 不可相似对角化.

对 $\lambda_1=\lambda_2=\lambda_3=2$，考虑线性方程组(2.1-8)(这时其中的 λ_0 等于 2)，即

$$(\boldsymbol{A}-2\boldsymbol{I})\boldsymbol{x}=\boldsymbol{y},$$

亦即

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & -1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}. \tag{2.1-9}$$

对此线性方程组的增广矩阵施以行初等交换：

$$\left[\begin{array}{ccc:c} 1 & 1 & 1 & y_1 \\ 0 & 0 & 0 & y_2 \\ -1 & 0 & -1 & y_3 \end{array}\right] \to \left[\begin{array}{ccc:c} 1 & 1 & 1 & y_1 \\ 0 & 0 & 0 & y_2 \\ 0 & 1 & 0 & y_3+y_1 \end{array}\right],$$

由此看出，当 $\boldsymbol{y}\neq\boldsymbol{0}$ 时，只有在它的第 2 个分量 $y_2=0$ 的条件下，非齐次线性方程组(2.1-8)才有解.

令 $\boldsymbol{y}=\boldsymbol{0}$，即 $y_1=y_2=y_3=0$，得齐次线性方程组 $(\boldsymbol{A}-2\boldsymbol{I})\boldsymbol{x}=\boldsymbol{0}$ 的一个基础解系 $\boldsymbol{x}_1=[1,0,-1]^{\mathrm{T}}$，$\boldsymbol{x}_1$ 即是 $\boldsymbol{A}$ 的属于特征值 2 的特征向量.

由于 $\boldsymbol{x}_1$ 的第 2 个分量是 0，所以令 $\boldsymbol{y}=\boldsymbol{x}_1$，即 $y_1=1$，$y_2=0$，$y_3=-1$，代入(2.1-9)式解之，任取其一个解 $\boldsymbol{x}_2=[0,0,1]^{\mathrm{T}}$.

2.1 Jordan标准型

又因 $\boldsymbol{x}_2$ 的第 2 个分量是 0，所以令 $\boldsymbol{y}=\boldsymbol{x}_2$，即 $y_1=0, y_2=0, y_3=1$，代入(2.1-9)式解之，任取其一个解 $\boldsymbol{x}_3=[-1,1,0]^{\mathrm{T}}$.

由于 $\boldsymbol{x}_3$ 的第 2 个分量是 1，所以不能再通过令 $\boldsymbol{y}=\boldsymbol{x}_3$ 代入(2.1-9)式求解.

这样便得到 3 个向量 $\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{x}_3$，按上所述它们线性无关. 于是令

$$\boldsymbol{P}=[\boldsymbol{x}_1 \quad \boldsymbol{x}_2 \quad \boldsymbol{x}_3]=\begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 1 \\ -1 & 1 & 0 \end{bmatrix},$$

则有 $$\boldsymbol{AP}=[\boldsymbol{Ax}_1 \mid \boldsymbol{Ax}_2 \mid \boldsymbol{Ax}_3]=[2\boldsymbol{x}_1 \mid 2\boldsymbol{x}_2+\boldsymbol{x}_1 \mid 2\boldsymbol{x}_3+\boldsymbol{x}_2]=\boldsymbol{P}\begin{bmatrix} 2 & 1 & \\ & 2 & 1 \\ & & 2 \end{bmatrix}.$$

由于 $\boldsymbol{P}$ 可逆，事实上 $$\boldsymbol{P}^{-1}=\begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix},$$

所以 $$\boldsymbol{P}^{-1}\boldsymbol{AP}=\begin{bmatrix} 2 & 1 & \\ & 2 & 1 \\ & & 2 \end{bmatrix}.$$ ∎

2.1 Jordan标准型

值得指出的是：①求解线性方程组(2.1-8)与求解对其施以行初等变换所得的线性方程组是等价的，一般来说，后者较易求解；②在满足增广矩阵的秩等于系数矩阵的秩的情形下，只需求得线性方程组(2.1-8)的一个解，一般选取尽可能多的分量为零的(非零)解，以便求 $\boldsymbol{P}^{-1}$ 较容易；③尽管选取的线性方程组(2.1-8)的一个解可以不同，从而构成不同的可逆矩阵 $\boldsymbol{P}$，但 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ 是一致的.

对于一般的 n 阶方阵 $\boldsymbol{A}$，求相似变换矩阵 $\boldsymbol{P}$ 使 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ 为 Jordan 矩阵的方法，基本上与上述过程类似. 首先求出 $\boldsymbol{A}$ 的特征多项式，确定其不同的特征值；其次对于单重特征值只需求出对应的一个特征向量；然后对每一多重特征值 λ_0，考虑形如(2.1-8)的线性方程组，得到与 λ_0 的代数重数一样多的线性无关解；最后将属于单重特征值的特征向量和对应多重特征值的那些线性无关解合成矩阵 $\boldsymbol{P}$. 后面将说明 $\boldsymbol{P}$ 是可逆的.

2.1 Jordan标准型

例 3 求可逆矩阵 P 使 $P^{-1}AP$ 为 Jordan 矩阵，其中

$$A=\begin{bmatrix}4&3&0&1\\0&2&0&0\\1&3&2&1\\0&0&0&2\end{bmatrix}.$$

解 A 的特征多项式为

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

$$=(2-\lambda)^2\begin{vmatrix}4-\lambda&3\\0&2-\lambda\end{vmatrix}=(4-\lambda)(2-\lambda)^3,$$

故 A 有两个不同的特征值：单重特征值 4 和三重特征值 2.

对于单重特征值 4，求解齐次线性方程组

$$(A-4I)x=\begin{bmatrix}0&3&0&1\\0&-2&0&0\\1&3&-2&1\\0&0&0&-2\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\\x_4\end{bmatrix}=\mathbf{0},$$

得其一个基础解系 $x_1=[2,0,1,0]^T$，即 x_1 是 A 的属于特征值 4 的特征向量.

2.1 Jordan标准型

对于三重特征值 2，考虑线性方程组 $(\boldsymbol{A}-2\boldsymbol{I})\boldsymbol{x}=\boldsymbol{y}$，即

$$\begin{bmatrix} 2 & 3 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 3 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}=\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix}. \qquad (2.1\text{-}10)$$

由于 $r(\boldsymbol{A}-2\boldsymbol{I})=2$，所以特征值 2 的几何重数是 $4-2=2$，即属于特征值 2 的线性无关的特征向量只有两个，故还需在方程组(2.1-10)中令 $\boldsymbol{y}$ 是属于特征值 2 的一个适当的特征向量，然后解此方程组. 为此，对此方程组的增广矩阵施以行初等变换

$$\left[\begin{array}{cccc:c} 2 & 3 & 0 & 1 & y_1 \\ 0 & 0 & 0 & 0 & y_2 \\ 1 & 3 & 0 & 1 & y_3 \\ 0 & 0 & 0 & 0 & y_4 \end{array}\right] \to \left[\begin{array}{cccc:c} 1 & 0 & 0 & 0 & y_1-y_3 \\ 0 & 0 & 0 & 0 & y_2 \\ 0 & 1 & 0 & \frac{1}{3} & \frac{1}{3}(-y_1+2y_3) \\ 0 & 0 & 0 & 0 & y_4 \end{array}\right],$$

容易看出，当且仅当 $y_2=y_4=0$ 时，线性方程组(2.1-10) 才有解.

2.1 Jordan标准型

令 $y_1=y_2=y_3=y_4=0$，得到关于特征值 2 的两个线性无关的特征向量

$$x_2=[0,-1,0,3]^T,\quad x_3=[0,0,1,0]^T,$$

且 x_3 的第二和第四个分量均为零. 从而令 $y=x_3$，即 $y_1=y_2=y_4=0$，$y_3=1$. 代入(2.1-10)式解之，任取其一个解 $x_4=[-1,0,0,2]^T$（这个解实际上由求解经行初等变换后的同解方程组

$$\begin{cases}1\cdot x_1+0\cdot x_2+0\cdot x_3+0\cdot x_4=y_1-y_3,\\ 0\cdot x_1+0\cdot x_2+0\cdot x_3+0\cdot x_4=y_2,\\ 0\cdot x_1+1\cdot x_2+0\cdot x_3+\dfrac{1}{3}\cdot x_4=\dfrac{1}{3}(-y_1+2y_3),\\ 0\cdot x_1+0\cdot x_2+0\cdot x_3+0\cdot x_4=0\end{cases}$$

较容易得到).

取
$$P=[x_1\quad x_2\quad x_3\quad x_4]=\begin{bmatrix}2&0&0&-1\\0&-1&0&0\\1&0&1&0\\0&3&0&2\end{bmatrix},$$

则有
$$AP=[Ax_1\mid Ax_2\mid Ax_3\mid Ax_4]=[4x_1\mid 2x_2\mid 2x_3\mid 2x_4+x_3]$$

$$=[x_1\quad x_2\quad x_3\quad x_4]\begin{bmatrix}4&&&\\&2&&\\&&2&1\\&&&2\end{bmatrix}=P\begin{bmatrix}4&&&\\&2&&\\&&2&1\\&&&2\end{bmatrix}.$$

2.1 Jordan标准型

由于 $|\boldsymbol{P}|=-4\neq 0$，所以 $\boldsymbol{P}$ 可逆，且

$$\boldsymbol{P}^{-1}=\frac{1}{4}\begin{bmatrix} 2 & 3 & 0 & 1 \\ 0 & -4 & 0 & 0 \\ -2 & -3 & 4 & -1 \\ 0 & 6 & 0 & 2 \end{bmatrix}.$$

于是

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\begin{bmatrix} 4 & & & \\ & 2 & & \\ & & 2 & 1 \\ & & & 2 \end{bmatrix}. \quad \blacksquare$$

定义 2.1-2 设 λ_0 是方阵 $\boldsymbol{A}$ 的特征值，如果对于向量 $\boldsymbol{x}$，存在一个正整数 k，使

$$(\boldsymbol{A}-\lambda_0\boldsymbol{I})^{k-1}\boldsymbol{x}\neq\boldsymbol{0} \quad 但 \quad (\boldsymbol{A}-\lambda_0\boldsymbol{I})^{k}\boldsymbol{x}=\boldsymbol{0}, \qquad (2.1\text{-}11)$$

则称 $\boldsymbol{x}$ 为 $\boldsymbol{A}$ 的属于 λ_0 的 **k 级根向量**. 在不会发生混淆时，也称 $\boldsymbol{x}$ 为关于 λ_0 的 k 级根向量.

按根向量的定义可知，1 级根向量即是特征向量，反之，特征向量必为 1 级根向量. 又，例 2 中的 $\boldsymbol{x}_2$ 与 $\boldsymbol{x}_3$ 分别是 2 级与 3 级根向量. 例 3 中的 $\boldsymbol{x}_4$ 是关于特征值 2 的 2 级根向量.

2.1 Jordan标准型

定理 2.1-2 设 λ_0 是方阵 $\boldsymbol{A}$ 的特征值，则 $\boldsymbol{A}$ 的属于 λ_0 的不同级的根向量是线性无关的.

证 设 $\boldsymbol{x}_i(i=1,2,\cdots,p)$ 是 $\boldsymbol{A}$ 关于 λ_0 的 i 级根向量，要证明向量组 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_p$ 线性无关.

设有等式

$$k_1\boldsymbol{x}_1+\cdots+k_{p-1}\boldsymbol{x}_{p-1}+k_p\boldsymbol{x}_p=\boldsymbol{0}, \tag{2.1-12}$$

用 $(\boldsymbol{A}-\lambda_0\boldsymbol{I})^{p-1}$ 左乘上式两边，得

$$k_1(\boldsymbol{A}-\lambda_0\boldsymbol{I})^{p-1}\boldsymbol{x}_1+\cdots+k_{p-1}(\boldsymbol{A}-\lambda_0\boldsymbol{I})^{p-1}\boldsymbol{x}_{p-1} + k_p(\boldsymbol{A}-\lambda_0\boldsymbol{I})^{p-1}\boldsymbol{x}_p=\boldsymbol{0}. \tag{2.1-13}$$

由根向量的定义知，

$$(\boldsymbol{A}-\lambda_0\boldsymbol{I})^{p-1}\boldsymbol{x}_1=(\boldsymbol{A}-\lambda_0\boldsymbol{I})^{p-2}(\boldsymbol{A}-\lambda_0\boldsymbol{I})\boldsymbol{x}_1=(\boldsymbol{A}-\lambda_0\boldsymbol{I})^{p-2}\boldsymbol{0}=\boldsymbol{0},$$
$$\vdots$$
$$(\boldsymbol{A}-\lambda_0\boldsymbol{I})^{p-1}\boldsymbol{x}_{p-1}=\boldsymbol{0},$$

故(2.1-13)式成为

$$k_p(\boldsymbol{A}-\lambda_0\boldsymbol{I})^{p-1}\boldsymbol{x}_p=\boldsymbol{0}.$$

但 $(\boldsymbol{A}-\lambda_0\boldsymbol{I})^{p-1}\boldsymbol{x}_p\neq\boldsymbol{0}$，故 $k_p=0$. 类似地可证

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

因此，仅在 $k_1=\cdots=k_{p-1}=k_p=0$ 时才有等式(2.1-12)，从而向量组 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_p$ 线性无关. ∎

2.1 Jordan标准型

定理 2.1-3 方阵 $\boldsymbol{A}$ 的属于不同特征值的根向量是线性无关的.

证 设 $\lambda_1,\lambda_2,\cdots,\lambda_s$ 是 $\boldsymbol{A}$ 的所有不同特征值，$\boldsymbol{x}_i(i=1,2,\cdots,s)$ 是 $\boldsymbol{A}$ 关于 λ_i 的 n_i 级根向量，要证明向量组 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_s$ 线性无关.

设有等式

$$k_1\boldsymbol{x}_1+k_2\boldsymbol{x}_2+\cdots+k_s\boldsymbol{x}_s=\boldsymbol{0}, \tag{2.1-14}$$

用 $\boldsymbol{L}=(\boldsymbol{A}-\lambda_1\boldsymbol{I})^{n_1-1}(\boldsymbol{A}-\lambda_2\boldsymbol{I})^{n_2}\cdots(\boldsymbol{A}-\lambda_s\boldsymbol{I})^{n_s}$（它是 $\boldsymbol{A}$ 的多项式，$(\boldsymbol{A}-\lambda_i\boldsymbol{I})^{n_i}(i=1,2,\cdots,s)$ 是该多项式的因式，由于同一个变量的两个多项式乘法满足交换律，所以这些因式的乘积与其排列顺序无关.）左乘(2.1-14)式两边，得

$$\begin{aligned}&k_1\boldsymbol{L}\boldsymbol{x}_1+k_2\boldsymbol{L}\boldsymbol{x}_2+\cdots+k_s\boldsymbol{L}\boldsymbol{x}_s\\&=k_1(\boldsymbol{A}-\lambda_2\boldsymbol{I})^{n_2}\cdots(\boldsymbol{A}-\lambda_s\boldsymbol{I})^{n_s}(\boldsymbol{A}-\lambda_1\boldsymbol{I})^{n_1-1}\boldsymbol{x}_1+k_2(\boldsymbol{A}-\lambda_1\boldsymbol{I})^{n_1-1}(\boldsymbol{A}-\lambda_3\boldsymbol{I})^{n_3}\cdots\\&\quad\cdot(\boldsymbol{A}-\lambda_s\boldsymbol{I})^{n_s}(\boldsymbol{A}-\lambda_2\boldsymbol{I})^{n_2}\boldsymbol{x}_2+\cdots+k_s(\boldsymbol{A}-\lambda_1\boldsymbol{I})^{n_1-1}(\boldsymbol{A}-\lambda_2\boldsymbol{I})^{n_2}\cdots(\boldsymbol{A}-\lambda_s\boldsymbol{I})^{n_s}\boldsymbol{x}_s\\&=\boldsymbol{0}.\end{aligned} \tag{2.1-15}$$

由根向量的定义知 $(\boldsymbol{A}-\lambda_i\boldsymbol{I})^{n_i}\boldsymbol{x}_i=\boldsymbol{0}(i=2,3,\cdots,s)$，故(2.1-15)式成为

$$k_1(\boldsymbol{A}-\lambda_2\boldsymbol{I})^{n_2}\cdots(\boldsymbol{A}-\lambda_s\boldsymbol{I}_s)^{n_s}(\boldsymbol{A}-\lambda_1\boldsymbol{I})^{n_1-1}\boldsymbol{x}_1=\boldsymbol{0}. \tag{2.1-16}$$

令 $\boldsymbol{y}=(\boldsymbol{A}-\lambda_1\boldsymbol{I})^{n_1-1}\boldsymbol{x}_1$，则

$$\boldsymbol{y}\neq\boldsymbol{0},\quad 且\quad (\boldsymbol{A}-\lambda_1\boldsymbol{I})\boldsymbol{y}=(\boldsymbol{A}-\lambda_1\boldsymbol{I})^{n_1}\boldsymbol{x}_1=\boldsymbol{0},$$

故 $\boldsymbol{y}$ 是 $\boldsymbol{A}$ 的属于特征值 λ_1 的特征向量. 又因 $(\boldsymbol{A}-\lambda_i\boldsymbol{I})\boldsymbol{y}=\boldsymbol{A}\boldsymbol{y}-\lambda_i\boldsymbol{y}=\lambda_1\boldsymbol{y}-\lambda_i\boldsymbol{y}=(\lambda_1-\lambda_i)\boldsymbol{y}$，

2.1 Jordan标准型

从而对任意正整数 m 有

$$(A-\lambda_i I)^m y=(\lambda_1-\lambda_i)^m y \quad (i=2,3,\cdots,s).$$

于是,(2.1-16)式成为 $\quad k_1(\lambda_1-\lambda_2)^{n_2}\cdots(\lambda_1-\lambda_s)^{n_s}y=\mathbf{0}.$

由于 $\lambda_1,\lambda_2,\cdots,\lambda_s$ 互不相同,又 $y\neq\mathbf{0}$,故 $k_1=0$.

同理可证 $k_2=\cdots=k_s=0$,因而向量组 $x_1,x_2,\cdots,x_s$ 线性无关. ∎

定义 2.1-3 设 λ_0 是方阵 A 的 k 重特征值,即特征值 λ_0 的代数重数为 k,齐次线性方程组 $(A-\lambda_0 I)^k x=\mathbf{0}$ 的解空间 $\mathcal{N}((A-\lambda_0 I)^k)$ 称为 A 关于特征值 λ_0 的**根空间**,记作 $\mathcal{N}_{\lambda_0}$. $\mathcal{N}_{\lambda_0}$ 中根向量的最高级数 r 称为 λ_0 的**指标**.

值得指出的是,λ_0 的指标 r 可能小于 k. 例如,对于例 3 来说,特征值 2 的代数重数是 3,但由于

$$(A-2I)^2=\begin{bmatrix}2&3&0&1\\0&0&0&0\\1&3&0&1\\0&0&0&0\end{bmatrix}\begin{bmatrix}2&3&0&1\\0&0&0&0\\1&3&0&1\\0&0&0&0\end{bmatrix}=\begin{bmatrix}4&6&0&2\\0&0&0&0\\2&3&0&1\\0&0&0&0\end{bmatrix},$$

$$(A-2I)^3=(A-2I)^2(A-2I)=\begin{bmatrix}4&6&0&2\\0&0&0&0\\2&3&0&1\\0&0&0&0\end{bmatrix}\begin{bmatrix}2&3&0&1\\0&0&0&0\\1&3&0&1\\0&0&0&0\end{bmatrix}=\begin{bmatrix}8&12&0&4\\0&0&0&0\\4&6&0&2\\0&0&0&0\end{bmatrix},$$

2.1 Jordan标准型

所以 $(\boldsymbol{A}-2\boldsymbol{I})^2\boldsymbol{x}=\boldsymbol{0}$ 与 $(\boldsymbol{A}-2\boldsymbol{I})^3\boldsymbol{x}=\boldsymbol{0}$ 是等价的齐次线性方程组，实际上与 $2x_1+3x_2+0\cdot x_3+1\cdot x_4=0$ 等价，故它们的解空间相同，即

$$\mathcal{N}((\boldsymbol{A}-2\boldsymbol{I})^3)=\mathcal{N}((\boldsymbol{A}-2\boldsymbol{I})^2).$$

而 $\mathcal{N}((\boldsymbol{A}-2\boldsymbol{I})^2)$ 中根向量的最高级数是 2，且已找出一个 2 级根向量 $\boldsymbol{x}_4$，故特征值 2 的指标为 2，小于代数重数 3.

容易验证，如果方阵的特征值 λ_0 的指标是 r，则有

$$\mathcal{N}_{\lambda_0}=\mathcal{N}((\boldsymbol{A}-\lambda_0\boldsymbol{I})^r).$$

这是因为，对于任意两个正整数 m_1、m_2，当 $m_1<m_2$ 时必有

$$\mathcal{N}((\boldsymbol{A}-\lambda_0\boldsymbol{I})^{m_1})\subseteq\mathcal{N}((\boldsymbol{A}-\lambda_0\boldsymbol{I})^{m_2}).$$

我们将会知道，在把 $\boldsymbol{A}$ 相似化简成 Jordan 矩阵时，$\boldsymbol{A}$ 的特征值 λ_0 的指标 r 即是该 Jordan 矩阵的关于 λ_0 子 Jordan 矩阵之中 Jordan 块的最大阶数.

可以证明①，若 λ_0 是方阵 $\boldsymbol{A}$ 的 k 重特征值，则 $\boldsymbol{A}$ 关于 λ_0 的根空间的维数是 k，即

$$\dim\mathcal{N}_{\lambda_0}=k.$$

2.1 Jordan标准型

即
$$\begin{bmatrix}1 & -1 & 1 & 1 & 0 & 0\\ 1 & -1 & -1 & -1 & 0 & 0\\ 0 & 0 & 0 & 0 & 1 & 1\\ 0 & 0 & 0 & 0 & -1 & -1\\ 0 & 0 & 0 & 0 & -1 & 1\\ 0 & 0 & 0 & 0 & 1 & -1\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\\ x_4\\ x_5\\ x_6\end{bmatrix}=\begin{bmatrix}y_1\\ y_2\\ y_3\\ y_4\\ y_5\\ y_6\end{bmatrix}, \qquad (2.1\text{-}17)$$

对其增广矩阵施以行初等变换：

$$\left[\begin{array}{cccccc:c}1 & -1 & 1 & 1 & 0 & 0 & y_1\\ 1 & -1 & -1 & -1 & 0 & 0 & y_2\\ 0 & 0 & 0 & 0 & 1 & 1 & y_3\\ 0 & 0 & 0 & 0 & -1 & -1 & y_4\\ 0 & 0 & 0 & 0 & -1 & 1 & y_5\\ 0 & 0 & 0 & 0 & 1 & -1 & y_6\end{array}\right] \rightarrow \left[\begin{array}{cccccc:c}1 & -1 & 0 & 0 & 0 & 0 & \frac{1}{2}(y_1+y_2)\\ 0 & 0 & 1 & 1 & 0 & 0 & \frac{1}{2}(y_1-y_2)\\ 0 & 0 & 0 & 0 & 1 & 0 & \frac{1}{2}(y_3-y_5)\\ 0 & 0 & 0 & 0 & 0 & 0 & y_3+y_4\\ 0 & 0 & 0 & 0 & 0 & 1 & \frac{1}{2}(y_3+y_5)\\ 0 & 0 & 0 & 0 & 0 & 0 & y_5+y_6\end{array}\right].$$

因此，当且仅当 $\boldsymbol{y}$ 的第三与第四个分量之和及第五与第六个分量之和均为零时，该线性方程组才有解.

2.1 Jordan标准型

令 $\boldsymbol{y}=\boldsymbol{0}$，即 $y_1=y_2=\cdots=y_6=0$，则解线性方程组(2.1-17)得到关于特征值 2 的两个线性无关的 1 级根向量

$$\boldsymbol{x}_2^{(1)}=[0,0,1,-1,0,0]^{\mathrm{T}},\quad \boldsymbol{x}_3^{(1)}=[1,1,0,0,0,0]^{\mathrm{T}}.$$

由于这两个向量的第三与第四个分量之和及第五与第六个分量之和都为零，所以分别令 $\boldsymbol{y}=\boldsymbol{x}_2^{(1)},\boldsymbol{x}_3^{(1)}$，可相应地求出关于特征值 2 的 2 级根向量.

令 $\boldsymbol{y}=\boldsymbol{x}_2^{(1)}$，由线性方程组(2.1-17)解得一个 2 级根向量.

$$\boldsymbol{x}_2^{(2)}=\left[0,0,0,0,\frac{1}{2},\frac{1}{2}\right]^{\mathrm{T}}.$$

因其第五与第六个分量之和等于 1 而不是零，故不能由此 2 级根向量求得更高级根向量.

又令 $\boldsymbol{y}=\boldsymbol{x}_3^{(1)}$，由线性方程组(2.1-17)解得一个 2 级根向量.

$$\boldsymbol{x}_3^{(2)}=[1,0,0,0,0,0]^{\mathrm{T}}.$$

因其第三与第四个分量之和及第五与第六个分量之和均为零，故再令 $\boldsymbol{y}=\boldsymbol{x}_3^{(2)}$，解线性方程组(2.1-17)得一个 3 级根向量

$$\boldsymbol{x}_3^{(3)}=\left[\frac{1}{2},0,\frac{1}{2},0,0,0\right]^{\mathrm{T}}.$$

由于关于同一特征值的线性无关的同级根向量所求得的高一级根向量必线性无关(为什么?)，所以 $\boldsymbol{x}_2^{(2)},\boldsymbol{x}_3^{(2)}$ 线性无关，从而向量组 $\boldsymbol{x}_1,\boldsymbol{x}_2^{(1)},\boldsymbol{x}_2^{(2)},\boldsymbol{x}_3^{(1)},\boldsymbol{x}_3^{(2)},\boldsymbol{x}_3^{(3)}$ 线性无关.

2.1 Jordan标准型

于是，令 $P=[x_1 \quad x_2^{(1)} \quad x_2^{(2)} \quad x_3^{(1)} \quad x_3^{(2)} \quad x_3^{(3)}]$，即

$$P=\begin{bmatrix} 0 & 0 & 0 & 1 & 1 & \frac{1}{2} \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & \frac{1}{2} \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 1 & 0 & \frac{1}{2} & 0 & 0 & 0 \\ -1 & 0 & \frac{1}{2} & 0 & 0 & 0 \end{bmatrix},$$

则有

$$\begin{aligned} AP &= [Ax_1 \vdots Ax_2^{(1)} \vdots Ax_2^{(2)} \vdots Ax_3^{(1)} \vdots Ax_3^{(2)} \vdots Ax_3^{(3)}] \\ &= [0\cdot x_1 \vdots 2x_2^{(1)} \vdots 2x_2^{(2)}+x_2^{(1)} \vdots 2x_3^{(1)} \vdots 2x_3^{(2)}+x_3^{(1)} \vdots 2x_3^{(3)}+x_3^{(2)}] \\ &= [x_1 \vdots x_2^{(1)} \vdots x_2^{(2)} \vdots x_3^{(1)} \vdots x_3^{(2)} \vdots x_3^{(3)}]\begin{bmatrix} 0 & & & & & \\ & 2 & 1 & & & \\ & & 2 & & & \\ & & & 2 & 1 & \\ & & & & 2 & 1 \\ & & & & & 2 \end{bmatrix}. \end{aligned}$$

由于 P 可逆，故 $$P^{-1}AP=\operatorname{diag}\left(0,\begin{bmatrix} 2 & 1 \\ & 2 \end{bmatrix},\begin{bmatrix} 2 & 1 & \\ & 2 & 1 \\ & & 2 \end{bmatrix}\right).$$ ∎

2.1 Jordan标准型

在本例中，$\boldsymbol{A}$ 的特征值 2 的代数重数是 5，但特征值 2 的指标是 3，即 Jordan 矩阵的关于特征值 2 的子 Jordan 矩阵 $\mathrm{diag}\left(\begin{bmatrix}2 & 1\\ & 2\end{bmatrix}, \begin{bmatrix}2 & 1 & \\ & 2 & 1\\ & & 2\end{bmatrix}\right)$ 中，Jordan 块的最大阶数是 3.

通过上面几个例子，我们说明了任一方阵 $\boldsymbol{A}$ 是与一个 Jordan 矩阵相似的，称这个 Jordan 矩阵为 $\boldsymbol{A}$ 的 **Jordan 标准形**. 对于 $\boldsymbol{A}$ 的 Jordan 标准形的结构有下列几点结论：

(1) $\boldsymbol{A}$ 的 Jordan 标准形中子 Jordan 矩阵的个数等于 $\boldsymbol{A}$ 的不同特征值的个数；

(2) 每个子 Jordan 矩阵的阶数等于对应特征值的代数重数；

(3) 每个子 Jordan 矩阵中 Jordan 块的个数等于对应特征值的线性无关特征向量的个数，即对应特征值的几何重数；

(4) 每个子 Jordan 矩阵中 Jordan 块的最大阶数等于对应特征值的指标.

方阵的 Jordan 标准形，由于形式简单、揭示了方阵相似的本质关系，又可利用它进一步讨论方阵幂级数和方阵函数，所以在众多学科中都有广泛的应用.

2.1 Jordan标准型

例 5 设 $\boldsymbol{A}=\begin{bmatrix}-3 & 3 & -2\\ -7 & 6 & -3\\ 1 & -1 & 2\end{bmatrix}$，求 $\boldsymbol{A}^k$（k 是正整数）.

解 由例 1 知，$\det(\boldsymbol{A}-\lambda\boldsymbol{I})=(1-\lambda)(2-\lambda)^2$.

属于特征值 1 的一个特征向量取为

$$\boldsymbol{x}_1=[1,2,1]^{\mathrm{T}}.$$

对二重特征值 2，考虑线性方程组

$$(\boldsymbol{A}-2\boldsymbol{I})\boldsymbol{x}=\boldsymbol{y},$$

即

$$\begin{bmatrix}-5 & 3 & -2\\ -7 & 4 & -3\\ 1 & -1 & 0\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}=\begin{bmatrix}y_1\\ y_2\\ y_3\end{bmatrix}. \tag{2.1-18}$$

对其增广矩阵施以行初等变换：

$$\left[\begin{array}{ccc:c}-5 & 3 & -2 & y_1\\ -7 & 4 & -3 & y_2\\ 1 & -1 & 0 & y_3\end{array}\right]\to\left[\begin{array}{ccc:c}1 & -1 & 0 & y_3\\ -7 & 4 & -3 & y_2\\ -5 & 3 & -2 & y_1\end{array}\right]$$

$$\to\left[\begin{array}{ccc:c}1 & -1 & 0 & y_3\\ 0 & -3 & -3 & y_2+7y_3\\ 0 & -2 & -2 & y_1+5y_3\end{array}\right]\to\left[\begin{array}{ccc:c}1 & -1 & 0 & y_3\\ 0 & 1 & 1 & -\dfrac{1}{3}(y_2+7y_3)\\ 0 & 0 & 0 & 3y_1-2y_2+y_3\end{array}\right].$$

2.1 Jordan标准型

令 $\boldsymbol{y}=\boldsymbol{0}$，即 $y_1=y_2=y_3=0$，得齐次线性方程组 $(\boldsymbol{A}-2\boldsymbol{I})\boldsymbol{x}=\boldsymbol{0}$ 的一个基础解系

$$\boldsymbol{x}_2=[-1,-1,1]^{\mathrm{T}}.$$

由于 $3\cdot(-1)-2\cdot(-1)+1\cdot1=0$，所以令 $\boldsymbol{y}=\boldsymbol{x}_2$ 可得非齐次线性方程组(2.1-18)的一个解

$$\boldsymbol{x}_3=[-1,-2,0]^{\mathrm{T}}.$$

事实上，$\boldsymbol{x}_2$ 是属于特征值 2 的 1 级根向量(即特征向量)，$\boldsymbol{x}_3$ 是由 $\boldsymbol{x}_2$ 求得的关于特征值 2 的 2 级根向量.

令

$$\boldsymbol{P}=[\boldsymbol{x}_1\quad \boldsymbol{x}_2\quad \boldsymbol{x}_3]=\begin{bmatrix}1 & -1 & -1\\ 2 & -1 & -2\\ 1 & 1 & 0\end{bmatrix},$$

则 $\boldsymbol{P}$ 可逆，且

$$\boldsymbol{P}^{-1}=\begin{bmatrix}2 & -1 & 1\\ -2 & 1 & 0\\ 3 & -2 & 1\end{bmatrix},$$

2.1 Jordan标准型

得 $$P^{-1}AP=\begin{bmatrix}1 & & \\ & 2 & 1 \\ & & 2\end{bmatrix},\quad A=P\begin{bmatrix}1 & & \\ & 2 & 1 \\ & & 2\end{bmatrix}P^{-1}.$$

于是，
$$A^k=P\begin{bmatrix}1 & & \\ & 2 & 1 \\ & & 2\end{bmatrix}P^{-1}\cdot P\begin{bmatrix}1 & & \\ & 2 & 1 \\ & & 2\end{bmatrix}P^{-1}\cdots P\begin{bmatrix}1 & & \\ & 2 & 1 \\ & & 2\end{bmatrix}P^{-1}$$

$$=P\begin{bmatrix}1 & & \\ & 2 & 1 \\ & & 2\end{bmatrix}^k P^{-1}=P\begin{bmatrix}1^k & & \\ & 2^k & k2^{k-1} \\ & & 2^k\end{bmatrix}P^{-1}$$

$$=\begin{bmatrix}1 & -1 & -1 \\ 2 & -1 & -2 \\ 1 & 1 & 0\end{bmatrix}\begin{bmatrix}1 & & \\ & 2^k & k2^{k-1} \\ & & 2^k\end{bmatrix}\begin{bmatrix}2 & -1 & 1 \\ -2 & 1 & 0 \\ 3 & -2 & 1\end{bmatrix}$$

$$=\begin{bmatrix}2+2^{k+1}-3(k+2)2^{k-1} & -1-2^k+2(k+2)2^{k-1} & 1-(k+2)2^{k-1} \\ 4+2^{k+1}-3(k+4)2^{k-1} & -2-2^k+2(k+4)2^{k-1} & 2-(k+4)2^{k-1} \\ 2-2^{k+1}+3k2^{k-1} & -1+2^k-2k2^{k-1} & 1+k2^{k-1}\end{bmatrix}. \quad \blacksquare$$

作业（第1部分）

习题 2.1

2. 求可逆矩阵 $\boldsymbol{P}$ 使 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ 为 Jordan 矩阵，其中

$$(2)\ \boldsymbol{A}=\begin{bmatrix} 5 & 3 & 0 & 1 \\ -1 & 1 & 0 & 0 \\ 1 & 3 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{bmatrix}.$$

3. 求 $\boldsymbol{A}^k$（k 是正整数），其中

$$\boldsymbol{A}=\begin{bmatrix} 3 & 1 & -1 \\ 2 & 2 & -1 \\ 2 & 2 & 0 \end{bmatrix}.$$

CAYLEY-HAMILTON定理

矩阵论（7）

厦門大學
XIAMEN UNIVERSITY
信息学院
（国家示范性软件学院）
School of Informatics
黄 炜
博士·副教授
Dr. Wei Huang

2.2 Cayley-Hamilton定理

本节应用方阵的 Jordan 标准形来证明在矩阵论中占有重要地位的 Cayley-Hamilton 定理，并举例说明该定理的应用.

在 1.5 节中曾引入方阵多项式的概念.

定义方阵 $\boldsymbol{A}$ 的多项式为以方阵 $\boldsymbol{A}$ 作为变量的多项式，例如

$$a_0\boldsymbol{A}^m + a_1\boldsymbol{A}^{m-1} + \cdots + a_m\boldsymbol{I},$$

并把它看作是用 $\boldsymbol{A}$ 替换多项式

$$g(\lambda) = a_0\lambda^m + a_1\lambda^{m-1} + \cdots + a_m$$

中 λ，并用单位矩阵 $\boldsymbol{I}$ 替换 $\lambda^0 = 1$ 的结果，记为 $g(\boldsymbol{A})$. 显然 $g(\boldsymbol{A})$ 是与 $\boldsymbol{A}$ 同阶的方阵.

因此，对于方阵 $\boldsymbol{A}$ 来说，方阵多项式与通常所说的多项式是一一对应的.

如何计算 $g(\boldsymbol{A})$ 呢？由于 Jordan 矩阵是分块对角矩阵，且每个子块都是 Jordan 块，Jordan 块的幂次是容易计算的，所以自然会利用 $\boldsymbol{A}$ 的 Jordan 标准形来计算其多项式 $g(\boldsymbol{A})$.

因为对任一方阵 $\boldsymbol{A}$ 都有可逆矩阵 $\boldsymbol{P}$ 使

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \mathrm{diag}(\boldsymbol{J}_1, \boldsymbol{J}_2, \cdots, \boldsymbol{J}_s),$$

即

$$\boldsymbol{A} = \boldsymbol{P}\mathrm{diag}(\boldsymbol{J}_1, \boldsymbol{J}_2, \cdots, \boldsymbol{J}_s)\boldsymbol{P}^{-1},$$

2.2 Cayley-Hamilton定理

其中 $J_i(i=1,2,\cdots,s)$都是 Jordan 块，所以

$$
\begin{aligned}
g(A) &= \sum_{l=0}^{m} a_l A^{m-l} = \sum_{l=0}^{m} a_l (P\mathrm{diag}(J_1, J_2, \cdots, J_s)P^{-1})^{m-l} \\
&= \sum_{l=0}^{m} a_l P\mathrm{diag}(J_1^{m-l}, J_2^{m-l}, \cdots, J_s^{m-l})P^{-1} \\
&= P\left(\sum_{l=0}^{m} a_l \mathrm{diag}(J_1^{m-l}, J_2^{m-l}, \cdots, J_s^{m-l})\right)P^{-1} \\
&= P\mathrm{diag}\left(\sum_{l=0}^{m} a_l J_1^{m-l}, \sum_{l=0}^{m} a_l J_2^{m-l}, \cdots, \sum_{l=0}^{m} a_l J_s^{m-l}\right)P^{-1} \\
&= P\mathrm{diag}(g(J_1), g(J_2), \cdots, g(J_s))P^{-1}.
\end{aligned}
\tag{2.2-1}
$$

因而只需对 Jordan 块

$$
J_i = \begin{bmatrix} \lambda_i & 1 & & & \\ & \lambda_i & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda_i \end{bmatrix} \quad (i = 1, \cdots, s)
$$

计算 $g(J_i)$.

2.2 Cayley-Hamilton定理

不妨设 $\boldsymbol{J}_i$ 是 r_i 阶矩阵，由于 $\boldsymbol{J}_i=(\lambda\boldsymbol{I}_{r_i}+\boldsymbol{U})_{\lambda=\lambda_i}$，故对任一正整数 k 有

$$\boldsymbol{J}_i^k=(\lambda\boldsymbol{I}_{r_i}+\boldsymbol{U})_{\lambda=\lambda_i}^k=\left(\lambda^k\boldsymbol{I}_{r_i}+k\lambda^{k-1}\boldsymbol{U}+\frac{1}{2!}k(k-1)\lambda^{k-2}\boldsymbol{U}^2+\cdots+\boldsymbol{U}^k\right)_{\lambda=\lambda_i}$$

$$=(\lambda^k)_{\lambda=\lambda_i}\boldsymbol{I}_{r_i}+(\lambda^k)'_{\lambda=\lambda_i}\boldsymbol{U}+\frac{1}{2!}(\lambda^k)''_{\lambda=\lambda_i}\boldsymbol{U}^2+\cdots+\boldsymbol{U}^k, \tag{2.2-2}$$

其中 $\boldsymbol{U}$ 是 r_i 阶幂零矩阵：

$$\boldsymbol{U}=\begin{bmatrix}0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0\end{bmatrix}_{r_i\times r_i},$$

$\boldsymbol{U}$ 的逐次升幂是将其 1 所在的次对角线逐次向右上角移动，因此

$$\boldsymbol{J}_i^k=\begin{bmatrix}\lambda^k & (\lambda^k)' & \frac{1}{2!}(\lambda^k)'' & \cdots & \frac{1}{(r_i-1)!}(\lambda^k)^{(r_i-1)} \\ & \lambda^k & (\lambda^k)' & \ddots & \vdots \\ & & \lambda^k & \ddots & \frac{1}{2!}(\lambda^k)'' \\ & & & \ddots & (\lambda^k)' \\ & & & & \lambda^k\end{bmatrix}_{\lambda=\lambda_i}. \tag{2.2-2'}$$

2.2 Cayley-Hamilton定理

于是

$$g(\boldsymbol{J}_i)=\sum_{l=0}^{m}a_l J_i^{m-l}=\sum_{l=0}^{m}a_l\begin{bmatrix}\lambda^{m-l} & (\lambda^{m-l})' & \cdots & \frac{1}{(r-1)!}(\lambda^{m-l})^{(r_i-1)}\\ & \lambda^{m-l} & \ddots & \vdots\\ & & \ddots & (\lambda^{m-l})'\\ & & & \lambda^{m-l}\end{bmatrix}_{\lambda=\lambda_i}$$

$$=\begin{bmatrix}\sum\limits_{l=0}^{m}a_l\lambda^{m-l} & \left(\sum\limits_{l=0}^{m}a_l\lambda^{m-l}\right)' & \cdots & \frac{1}{(r_i-1)!}\left(\sum\limits_{l=0}^{m}a_l\lambda^{m-l}\right)^{(r_i-1)}\\ & \sum\limits_{l=0}^{m}a_l\lambda^{m-l} & \ddots & \vdots\\ & & \ddots & \left(\sum\limits_{l=0}^{m}a_l\lambda^{m-l}\right)'\\ & & & \sum\limits_{l=0}^{m}a_l\lambda^{m-l}\end{bmatrix}_{\lambda=\lambda_i}$$

$$=\begin{bmatrix}g(\lambda) & g'(\lambda) & \cdots & \frac{1}{(r_i-1)!}g^{(r_i-1)}_{(\lambda)}\\ & g(\lambda) & \ddots & \vdots\\ & & \ddots & g'(\lambda)\\ & & & g(\lambda)\end{bmatrix}_{\lambda=\lambda_i}. \qquad (2.2\text{-}3)$$

2.2 Cayley-Hamilton定理

例 1 求 $g(\boldsymbol{A})=2\boldsymbol{A}^5-3\boldsymbol{A}^4-\boldsymbol{A}^3+2\boldsymbol{A}-\boldsymbol{I}$，其中

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

解 由

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

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

知，$\boldsymbol{A}$ 有两个不同的特征值：2 和 1，且 1 是 2 重特征值.

对特征值 2，由

$$(\boldsymbol{A}-2\boldsymbol{I})\boldsymbol{x}=\boldsymbol{0}$$

解得它的一个基础解系 $\boldsymbol{x}_1=[1,0,-1]^{\mathrm{T}}$.

2.2 Cayley-Hamilton定理

对 2 重特征值 1，考虑线性方程组

$$(\boldsymbol{A}-\boldsymbol{I})\boldsymbol{x}=\boldsymbol{y},$$

即

$$\begin{bmatrix}0 & 1 & -1\\ 1 & 0 & 1\\ 0 & -1 & 1\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}=\begin{bmatrix}y_1\\ y_2\\ y_3\end{bmatrix}, \tag{2.2-4}$$

对其增广矩阵施以初等变换

$$\left[\begin{array}{ccc:c}0 & 1 & -1 & y_1\\ 1 & 0 & 1 & y_2\\ 0 & -1 & 1 & y_3\end{array}\right]\rightarrow\left[\begin{array}{ccc:c}0 & 1 & -1 & y_1\\ 1 & 0 & 1 & y_2\\ 0 & 0 & 0 & y_3+y_1\end{array}\right].$$

令 $\boldsymbol{y}=\boldsymbol{0}$，得到属于特征值 1 的一个特征向量（即 1 级根向量）$\boldsymbol{x}_2=[-1,1,1]^{\mathrm{T}}$，且它的第一与第三个分量之和为零．因此，令 $\boldsymbol{y}=\boldsymbol{x}_2$，即 $y_1=-1, y_2=y_3=1$，由方程组（2.2-4）解得一个解为 $\boldsymbol{x}_3=[0,0,1]^{\mathrm{T}}$，它是关于特征值 1 的 2 级根向量．令

$$\boldsymbol{P}=[\boldsymbol{x}_1 \quad \boldsymbol{x}_2 \quad \boldsymbol{x}_3]=\begin{bmatrix}1 & -1 & 0\\ 0 & 1 & 0\\ -1 & 1 & 1\end{bmatrix},$$

则 $\boldsymbol{P}$ 可逆，且

$$\boldsymbol{P}^{-1}=\begin{bmatrix}1 & 1 & 0\\ 0 & 1 & 0\\ 1 & 0 & 1\end{bmatrix}, \quad \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\begin{bmatrix}\boxed{2} & & \\ & 1 & 1\\ & & 1\end{bmatrix}, \quad \boldsymbol{A}=\boldsymbol{P}\begin{bmatrix}\boxed{2} & & \\ & 1 & 1\\ & & 1\end{bmatrix}\boldsymbol{P}^{-1}.$$

2.2 Cayley-Hamilton定理

$\boldsymbol{A}$ 的多项式 $g(\boldsymbol{A})=2\boldsymbol{A}^5-3\boldsymbol{A}^4-\boldsymbol{A}^3+2\boldsymbol{A}-\boldsymbol{I}$ 所对应的通常多项式是

$$g(\lambda)=2\lambda^5-3\lambda^4-\lambda^3+2\lambda-1.$$

由于 $\boldsymbol{J}_1=2$ 是 1 阶 Jordan 块，故由(2.2-3)式得

$$g(\boldsymbol{J}_1)=2\cdot 2^5-3\cdot 2^4-2^3+2\cdot 2-1=11.$$

$\boldsymbol{J}_2=\begin{bmatrix}1 & 1\\ & 1\end{bmatrix}$是 2 阶 Jordan 块，由(2.2-3)式得

$$g(\boldsymbol{J}_2)=\begin{bmatrix}g(1) & g'(1)\\ & g(1)\end{bmatrix}=\begin{bmatrix}-1 & -3\\ & -1\end{bmatrix}.$$

因此，由(2.2-1)式得

$$g(\boldsymbol{A})=\boldsymbol{P}\begin{bmatrix}g(\boldsymbol{J}_1) & \\ & g(\boldsymbol{J}_2)\end{bmatrix}\boldsymbol{P}^{-1}=\begin{bmatrix}1 & -1 & 0\\ 0 & 1 & 0\\ -1 & 1 & 1\end{bmatrix}\begin{bmatrix}11 & & \\ & -1 & -3\\ & & -1\end{bmatrix}\begin{bmatrix}1 & 1 & 0\\ 0 & 1 & 0\\ 1 & 0 & 1\end{bmatrix}$$

$$=\begin{bmatrix}14 & 12 & 3\\ -3 & -1 & -3\\ -15 & -12 & -4\end{bmatrix}.$$

▮

2.2 Cayley-Hamilton定理

对于给定的方阵 $\boldsymbol{A}$ 来说，有一类特殊的多项式具有重要作用，这就是 $\boldsymbol{A}$ 的零化多项式. 例如，对于 $\boldsymbol{A}=\begin{bmatrix}1 & 1\\0 & 2\end{bmatrix}$，多项式 $g(\lambda)=\lambda^2-3\lambda+2$ 能使

$$g(\boldsymbol{A})=\boldsymbol{A}^2-3\boldsymbol{A}+2\boldsymbol{I}=\begin{bmatrix}1 & 3\\0 & 4\end{bmatrix}-3\begin{bmatrix}1 & 1\\0 & 2\end{bmatrix}+2\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}=\begin{bmatrix}0 & 0\\0 & 0\end{bmatrix},$$

多项式 $g(\lambda)=\lambda^2-3\lambda+2$ 就称为 $\boldsymbol{A}$ 的零化多项式.

定义 2.2-1 设 $\boldsymbol{A}$ 是 n 阶方阵，$g(\lambda)$是一多项式，如果有$g(\boldsymbol{A})=\boldsymbol{O}$，则称 $g(\lambda)$是 $\boldsymbol{A}$ 的**零化多项式.**

如果 $g(\lambda)$是 $\boldsymbol{A}$ 的一个零化多项式，$h(\lambda)$是任一多项式，因$h(\boldsymbol{A})g(\boldsymbol{A})=\boldsymbol{O}$，故$h(\lambda)g(\lambda)$也是 $\boldsymbol{A}$ 的零化多项式. 因此方阵的零化多项式不是唯一的. 零化多项式是否存在呢？下述定理说明，任何方阵的零化多项式必存在.

定理 2.2-1 (Cayley-Hamilton) 方阵 $\boldsymbol{A}$ 的特征多项式一定是 $\boldsymbol{A}$ 的一个零化多项式.

2.2 Cayley-Hamilton定理

证 设 n 阶方阵 $\boldsymbol{A}$ 的特征多项式为

$$f(\lambda)=\det(\boldsymbol{A}-\lambda\boldsymbol{I})=(\lambda_1-\lambda)^{n_1}(\lambda_2-\lambda)^{n_2}\cdots(\lambda_l-\lambda)^{n_l},$$

其中$\lambda_1,\cdots,\lambda_l$是$\boldsymbol{A}$的所有不同特征值，$n_i(i=1,\cdots,l)$是特征值$\lambda_i$的代数重数，且$\sum_{i=1}^{l}n_i=n$，则由高等数学知

$$f(\lambda_i)=f'(\lambda_i)=\cdots=f^{(n_i-1)}(\lambda_i)=0,\quad i=1,2,\cdots,l. \tag{2.2-5}$$

而由 $\boldsymbol{A}$ 的 Jordan 标准形得

$$\boldsymbol{A}=\boldsymbol{P}\mathrm{diag}(\boldsymbol{J}_1,\boldsymbol{J}_2,\cdots,\boldsymbol{J}_l)\boldsymbol{P}^{-1},$$

其中 $\boldsymbol{J}_i(i=1,2,\cdots,l)$是主对角线均为 λ_i 的若干个 Jordan 块所构成的子 Jordan 矩阵，即

$$\boldsymbol{J}_i=\mathrm{diag}(\boldsymbol{J}_{i1},\boldsymbol{J}_{i2},\cdots,\boldsymbol{J}_{ik_i}),$$

且

$$\boldsymbol{J}_{ij}=\begin{bmatrix}\lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i\end{bmatrix}_{r_{ij}\times r_{ij}}\quad (j=1,\cdots,k_i).$$

按(2.2-1)式有 $f(\boldsymbol{A})=\boldsymbol{P}\mathrm{diag}(f(\boldsymbol{J}_1),f(\boldsymbol{J}_2),\cdots,f(\boldsymbol{J}_l))\boldsymbol{P}^{-1}$，

且

$$f(\boldsymbol{J}_i)=\mathrm{diag}(f(\boldsymbol{J}_{i1}),f(\boldsymbol{J}_{i2}),\cdots,f(\boldsymbol{J}_{ik_i})),\quad i=1,2,\cdots,l,$$

2.2 Cayley-Hamilton定理

并且根据(2.2-3)式知

$$f(\boldsymbol{J}_{ij})=\begin{bmatrix} f(\lambda_i) & f'(\lambda_i) & \cdots & \frac{1}{(r_{ij}-1)!}g^{(r_{ij}-1)}(\lambda_i) \\ & f(\lambda_i) & \ddots & \vdots \\ & & \ddots & f'(\lambda_i) \\ & & & f(\lambda_i) \end{bmatrix}.$$

由于 $r_{ij}\leqslant n_i(i=1,2,\cdots,l,j=1,2,\cdots,k_i)$，所以由(2.2-5)式知 $f(\boldsymbol{J}_{ij})=\boldsymbol{O}(i=1,2,\cdots,l,j=1,2,\cdots,k_i)$，从而 $f(\boldsymbol{J}_i)=\boldsymbol{O}(i=1,2,\cdots,l)$. 于是有 $f(\boldsymbol{A})=\boldsymbol{O}$，即 $f(\lambda)$ 是 $\boldsymbol{A}$ 的一个零化多项式. ■

Cayley-Hamilton 定理表明，n 阶方阵 $\boldsymbol{A}$ 的高次幂 $\boldsymbol{A}^k(k\geqslant n)$ 可表示为 $\boldsymbol{I},\boldsymbol{A},\cdots,\boldsymbol{A}^{n-1}$ 的线性组合.

这里举两个例子说明 Cayley-Hamilton 定理在计算方阵多项式和逆矩阵方面的应用.

2.2 Cayley-Hamilton定理

例 2 已知 $A=\begin{bmatrix}1 & 1 & -1\\ 1 & 1 & 1\\ 0 & -1 & 2\end{bmatrix}$，求 $g(A)=2A^5-3A^4-A^3+2A-I$.

解 例 6 是用 A 的 Jordan 标准形计算 $g(A)$ 的，现在应用 Cayley-Hamilton 定理计算 $g(A)$.

A 的特征多项式为

$$f(\lambda)=|A-\lambda I|=(2-\lambda)(1-\lambda)^2,$$

且 $g(A)$ 所对应的多项式是 $g(\lambda)=2\lambda^5-3\lambda^4-\lambda^3+2\lambda-1$. 用 $f(\lambda)$ 除 $g(\lambda)$ 一般得到下述表达式

$$g(\lambda)=q(\lambda)f(\lambda)+r(\lambda), \qquad (2.2\text{-}6)$$

其中 $g(\lambda)$、$r(\lambda)$ 都是多项式，且余式 $r(\lambda)$ 的次数至少比 $f(\lambda)$ 的次数低一次. 因而可以设

$$r(\lambda)=a_0\lambda^2+a_1\lambda+a_2.$$

将其代入(2.2-6)式得

$$g(\lambda)=q(\lambda)f(\lambda)+a_0\lambda^2+a_1\lambda+a_2. \qquad (2.2\text{-}7)$$

将特征值 2、1 分别代入(2.2-7)式，因 $f(2)=f(1)=0$，故有

$$4a_0+2a_1+a_2=g(2)=11, \quad a_0+a_1+a_2=g(1)=-1.$$

(2.2-6)式两边对 λ 求导得

$$g'(\lambda)=q'(\lambda)f(\lambda)+q(\lambda)f'(\lambda)+r'(\lambda).$$

2.2 Cayley-Hamilton定理

将特征值 1 代入上式，因 $f(1)=f'(1)=0$，故有 $r'(1)=g'(1)$，即

$$2a_0+a_1=g'(1)=-3.$$

这样就得到关于待定系数 a_0, a_1, a_2 的线性方程组

$$\begin{cases} 4a_0+2a_1+a_2=11, \\ a_0+a_1+a_2=-1, \\ 2a_0+a_1=-3. \end{cases}$$

解之得 $a_0=15, a_1=-33, a_2=17$.

另一方面，由(2.2-6)式和 $f(\lambda)$ 是 $\boldsymbol{A}$ 的零化多项式得

$$g(\boldsymbol{A})=q(\boldsymbol{A})f(\boldsymbol{A})+r(\boldsymbol{A})=r(\boldsymbol{A}),$$

于是，
$$g(\boldsymbol{A})=a_0\boldsymbol{A}^2+a_1\boldsymbol{A}+a_2\boldsymbol{I}=15\boldsymbol{A}^2-33\boldsymbol{A}+17\boldsymbol{I}$$

$$=15\begin{bmatrix} 2 & 3 & -2 \\ 2 & 1 & 2 \\ -1 & -3 & 3 \end{bmatrix}-33\begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & 1 \\ 0 & -1 & 2 \end{bmatrix}+17\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$=\begin{bmatrix} 14 & 12 & 3 \\ -3 & -1 & -3 \\ -15 & -12 & -4 \end{bmatrix}. \blacksquare$$

2.2 Cayley-Hamilton定理

例 3 就例 2 中的 $\boldsymbol{A}$，证明 $\boldsymbol{A}$ 可逆，并求 $\boldsymbol{A}^{-1}$.

解 因 $\boldsymbol{A}$ 的特征值是 2，1，1，故 $|\boldsymbol{A}|=2\times1\times1=2\neq0$，从而 $\boldsymbol{A}$ 可逆.

由于 $\boldsymbol{A}$ 的特征多项式 $f(\lambda)=(2-\lambda)(1-\lambda)^2=-\lambda^3+4\lambda^2-5\lambda+2$ 是 $\boldsymbol{A}$ 的零化多项式，故

$$f(\boldsymbol{A})=-\boldsymbol{A}^3+4\boldsymbol{A}^2-5\boldsymbol{A}+2\boldsymbol{I}=\boldsymbol{O}.$$

用 $\boldsymbol{A}^{-1}$ 乘之得

$$-\boldsymbol{A}^2+4\boldsymbol{A}-5\boldsymbol{I}+2\boldsymbol{A}^{-1}=\boldsymbol{O},$$

由此得到

$$\boldsymbol{A}^{-1}=\frac{1}{2}(\boldsymbol{A}^2-4\boldsymbol{A}+5\boldsymbol{I})=\frac{1}{2}\begin{bmatrix}3&-1&2\\-2&2&-2\\-1&1&0\end{bmatrix}.$$

∎

2.2 Cayley-Hamilton定理

定义 2.2-2 在 $\boldsymbol{A}$ 的零化多项式中，次数最低的首一(即最高次项的系数为 1)多项式称为 $\boldsymbol{A}$ 的**最小多项式**，记作 $m_A(\lambda)$.

定理 2.2-2 $\boldsymbol{A}$ 的最小多项式 $m_A(\lambda)$ 可整除 $\boldsymbol{A}$ 的任何零化多项式，且 $m_A(\lambda)$ 是唯一的.

证 设 $g(\lambda)$ 是 $\boldsymbol{A}$ 的任一零化多项式，若 $m_A(\lambda)$ 不能整除 $g(\lambda)$，则有

$$g(\lambda) = q(\lambda)m_A(\lambda) + r(\lambda),$$

且余式 $r(\lambda)$ 的次数至少比 $m_A(\lambda)$ 的次数低 1 次.

由于 $g(\boldsymbol{A})=\boldsymbol{O}$ 和 $m_A(\boldsymbol{A})=\boldsymbol{O}$，故由 $g(\boldsymbol{A})=q(\boldsymbol{A})m_A(\boldsymbol{A})+r(\boldsymbol{A})$ 得 $r(\boldsymbol{A})=\boldsymbol{O}$，故 $r(\lambda)$ 是 $\boldsymbol{A}$ 的一个零化多项式，且其次数低于 $m_A(\boldsymbol{A})$ 的次数. 这与 $m_A(\lambda)$ 的定义矛盾. 因此 $m_A(\lambda)$ 必能整除 $g(\lambda)$，记为 $m_A(\lambda)\mid g(\lambda)$.

再证唯一性. 设 $m_1(\lambda)$、$m_2(\lambda)$ 都是 $\boldsymbol{A}$ 的最小多项式，则有 $m_1(\lambda)\mid m_2(\lambda)$ 和 $m_2(\lambda)\mid m_1(\lambda)$. 因此 $m_1(\lambda)=bm_2(\lambda)$ 且 $b\neq 0$ 是常数. 但 $m_1(\lambda)$、$m_2(\lambda)$ 都是首一多项式，故 $b=1$，从而 $m_1(\lambda)=m_2(\lambda)$. ∎

推论 $\boldsymbol{A}$ 的最小多项式必能整除 $\boldsymbol{A}$ 的特征多项式.

2.2 Cayley-Hamilton定理

定理 2.2-3 λ_0 是方阵 $\boldsymbol{A}$ 的特征值的充要条件是，λ_0 是 $\boldsymbol{A}$ 的最小多项式 $m_{\boldsymbol{A}}(\lambda)$ 的根.

证 设 λ_0 是 $\boldsymbol{A}$ 的特征值，$\boldsymbol{x}_0 \neq \boldsymbol{0}$ 是属于 λ_0 的特征向量，则由 1.5 节的定理1.5-3知，

$$m_{\boldsymbol{A}}(\boldsymbol{A})\boldsymbol{x}_0 = m_{\boldsymbol{A}}(\lambda_0)\boldsymbol{x}_0.$$

但 $m_{\boldsymbol{A}}(\boldsymbol{A}) = \boldsymbol{O}$，故由上式可得 $\qquad m_{\boldsymbol{A}}(\lambda_0)\boldsymbol{x}_0 = \boldsymbol{0}.$

又因 $\boldsymbol{x}_0 \neq \boldsymbol{0}$，故 $m_{\boldsymbol{A}}(\lambda_0) = 0$，即 λ_0 是 $m_{\boldsymbol{A}}(\lambda)$ 的根.

反之，若 λ_0 是 $m_{\boldsymbol{A}}(\lambda)$ 的根，那么由于 $m_{\boldsymbol{A}}(\lambda)$ 可整除 $\boldsymbol{A}$ 的特征多项式，故 λ_0 也是 $\boldsymbol{A}$ 的特征多项式的根，即 λ_0 是 $\boldsymbol{A}$ 的特征值. ∎

这个定理给出了用特征多项式检验最小多项式的方法. 事实上，设 n 阶方阵 $\boldsymbol{A}$ 的所有不同特征值是 $\lambda_1, \lambda_2, \cdots, \lambda_s$，且 $\boldsymbol{A}$ 的特征多项式为

$$f(\lambda) = \det(\boldsymbol{A} - \lambda \boldsymbol{I}) = (\lambda_1 - \lambda)^{n_1} \cdots (\lambda_s - \lambda)^{n_s},$$

其中 $n_i (i = 1, 2, \cdots, s)$ 是特征值 λ_i 的代数重数，$\sum_{i=1}^{s} n_i = n$，那么由定理 2.2-2 的推论和定理 2.2-3 知，$\boldsymbol{A}$ 的最小多项式一定有如下形式：

$$m_{\boldsymbol{A}}(\lambda) = (\lambda - \lambda_1)^{r_1} \cdots (\lambda - \lambda_s)^{r_s}, \tag{2.2-8}$$

且 $$1 \leqslant r_i \leqslant n_i \quad (i = 1, 2, \cdots, s).$$

特别地，若 $\boldsymbol{A}$ 的每个特征值的代数重数都是 1，则有

$$m_{\boldsymbol{A}}(\lambda) = (-1)^n \det(\boldsymbol{A} - \lambda \boldsymbol{I}).$$

2.2 Cayley-Hamilton定理

例 4 求 Jordan 矩阵 $\boldsymbol{J}$ 的最小多项式，其中

$$\boldsymbol{J}=\begin{bmatrix} -2 & 1 & & & & & & \\ & -2 & & & & & & \\ & & -2 & 1 & & & & \\ & & & -2 & 1 & & & \\ & & & & -2 & & & \\ & & & & & 3 & & \\ & & & & & & 3 & 1 \\ & & & & & & & 3 \end{bmatrix}.$$

解 $\boldsymbol{J}$ 的特征多项式为

$$\det(\boldsymbol{J}-\lambda\boldsymbol{I})=(-2-\lambda)^5(3-\lambda)^3,$$

故 $\boldsymbol{J}$ 的最小多项式的形式是$(\lambda+2)^{r_1}(\lambda-3)^{r_2}$.

由于

$$\boldsymbol{J}+2\boldsymbol{I}=\begin{bmatrix} 0 & 1 & & & & & & \\ 0 & 0 & & & & & & \\ & & 0 & 1 & & & & \\ & & & 0 & 1 & & & \\ & & & & 0 & & & \\ & & & & & 5 & & \\ & & & & & & 5 & 1 \\ & & & & & & & 5 \end{bmatrix},$$

2.2 Cayley-Hamilton定理

$$\boldsymbol{J}-3\boldsymbol{I}=\begin{bmatrix} -5 & 1 & & & & & & \\ & -5 & & & & & & \\ & & -5 & 1 & & & & \\ & & & -5 & 1 & & & \\ & & & & -5 & & & \\ & & & & & 0 & & \\ & & & & & & 0 & 1 \\ & & & & & & & 0 \end{bmatrix},$$

所以由分块矩阵乘法知，要使$(\boldsymbol{J}+2\boldsymbol{I})^{r_1}(\boldsymbol{J}-3\boldsymbol{I})^{r_2}=\boldsymbol{O}$，只需注意使 $\boldsymbol{J}+2\boldsymbol{I}$，$\boldsymbol{J}-3\boldsymbol{I}$ 中的幂零子块分别变为零所要求的幂次，它们依次即是 r_1、r_2.

对于 $\boldsymbol{J}+2\boldsymbol{I}$ 中的幂零子块$\begin{bmatrix} 0 & 1 \\ & 0 \end{bmatrix}$，$\begin{bmatrix} 0 & 1 & \\ & 0 & 1 \\ & & 0 \end{bmatrix}$，要使它们变为零矩阵所需的幂次分别为 2 与 3，故要使它们同时变为零矩阵所需的幂次是 $\max\{2,3\}=3$，即 $r_1=3$.

对于 $\boldsymbol{J}-3\boldsymbol{I}$ 中的幂零子块$\begin{bmatrix} 0 & 1 \\ & 0 \end{bmatrix}$，要使它变零矩阵所需的幂次是 2，即 $r_2=2$($\boldsymbol{J}-3\boldsymbol{I}$ 中的 1 阶零子块按定义不是幂零子块).

因此，$\boldsymbol{J}$ 的最小多项式是

$$m_J(\lambda)=(\lambda+2)^3(\lambda-3)^2. \qquad ■$$

2.2 Cayley-Hamilton定理

从例 4 得知，方阵 $\boldsymbol{A}$ 的最小多项式(2.2-8)中的 $r_i(i=1,2,\cdots,s)$应为特征值 λ_i 的指标.

若方阵 $\boldsymbol{A}=\operatorname{diag}(\boldsymbol{A}_1,\boldsymbol{A}_2,\cdots,\boldsymbol{A}_l)$，其中 $\boldsymbol{A}_i(i=1,2,\cdots,l)$也都是方阵，那么只要知道每个子块 $\boldsymbol{A}_i$ 的最小多项式 $m_{\boldsymbol{A}_i}(\lambda)$，就可得到 $\boldsymbol{A}$ 的最小多项式 $m_{\boldsymbol{A}}(\lambda)$. 事实上，$m_{\boldsymbol{A}}(\lambda)$是 $m_{\boldsymbol{A}_1}(\lambda),\cdots,m_{\boldsymbol{A}_l}(\lambda)$的最小公倍式. 因为 $\boldsymbol{A}$ 的最小多项式是 $\boldsymbol{A}$ 的次数最低的零化多项式，而$\boldsymbol{A}_i(i=1,2,\cdots,l)$的最小多项式必可整除 $\boldsymbol{A}$ 的最小多项式.

例如，对于分块对角矩阵

$$\boldsymbol{A}=\left[\begin{array}{cc|cc} 2 & 1 & & \\ 0 & 2 & & \\ \hline & & 1 & 1 \\ & & -2 & 4 \end{array}\right]$$

来说，$\boldsymbol{A}_1=\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}$的最小多项式是$(\lambda-2)^2$，$\boldsymbol{A}_2=\begin{bmatrix} 1 & 1 \\ -2 & 4 \end{bmatrix}$的最小多项式是$(\lambda-3)(\lambda-2)$，它们的最小公倍式是$(\lambda-3)(\lambda-2)^2$. 容易验证，这确是 $\boldsymbol{A}$ 的最小多项式.

2.2 Cayley-Hamilton定理

定理 2.2-4 方阵 $\boldsymbol{A}$ 可相似对角化的充要条件是，$\boldsymbol{A}$ 的最小多项式没有重根.

证 **必要性：**由于对角矩阵可以看作 1 阶子块所构成的分块对角矩阵，而 1 阶方阵的最小多项式是一次多项式，所以 $\boldsymbol{A}$ 的最小多项式没有重根.

充分性：若 $\boldsymbol{A}$ 的最小多项式没有重根，则 $\boldsymbol{A}$ 的每个特征值的指标是 1，从而 $\boldsymbol{A}$ 的 Jordan 标准形中每个 Jordan 块的阶数为 1，故 $\boldsymbol{A}$ 可相似对角化. ■

例 5 设方阵 $\boldsymbol{A}$ 满足关系式

$$\boldsymbol{A}^3-3\boldsymbol{A}^2-\boldsymbol{A}+3\boldsymbol{I}=\boldsymbol{O},$$

证明 $\boldsymbol{A}$ 必可相似对角化.

证 由题设可知

$$g(\lambda)=\lambda^3-3\lambda^2-\lambda+3$$

是 $\boldsymbol{A}$ 的一个零化多项式，而 $g(\lambda)=(\lambda+1)(\lambda-1)(\lambda-3)$ 没有重根，且 $\boldsymbol{A}$ 的最小多项式 $m_A(\lambda)$ 应能整除 $g(\lambda)$，故 $m_A(\lambda)$ 不会有重根. 因此 $\boldsymbol{A}$ 必可相似对角化. ■

作业（第2部分）

习题 2.2

1. 求 $g(\boldsymbol{A})=\boldsymbol{A}^8-9\boldsymbol{A}^6+\boldsymbol{A}^4-3\boldsymbol{A}^3+4\boldsymbol{A}^2+\boldsymbol{I}$，其中

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

2. 求下列方阵的最小多项式：

$$(2)\ \left[\begin{array}{cc:cc} 3 & 1 & 1 & 1 \\ -4 & -1 & -1 & 1 \\ \hdashline & & 2 & 1 \\ & & -1 & 0 \end{array}\right].$$

謝謝觀看！

厦門大學
XIAMEN UNIVERSITY

信息学院
（国家示范性软件学院）
School of Informatics

黃 煒
博士，副教授
Dr. Wei Huang