矩阵可对角化的充要条件

定理 6.4.2. n 阶方阵 $\boldsymbol{A}$ 相似于对角阵的充要条件是 $\boldsymbol{A}$ 有 n 个线性无关的特征向量.

证明. (1) 设 $\boldsymbol{A}$ 可以对角化, 即存在可逆矩阵 $\boldsymbol{T}$ 使得 $\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T} = \boldsymbol{D}$ 为对角阵. 设 $\boldsymbol{D} = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$, 而 $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ 为 $\boldsymbol{T}$ 的各个列向量. 此时 $\boldsymbol{A}\boldsymbol{T} = \boldsymbol{T}\boldsymbol{D}$, 即有

$$\boldsymbol{A}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}.$$

这说明 $\boldsymbol{A}\boldsymbol{x}_i = \lambda_i \boldsymbol{x}_i$, $i = 1, 2, \ldots, n$. 由于 $\boldsymbol{T}$ 可逆, 向量组 $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ 线性无关. 特别地, 每个 $\boldsymbol{x}_i$ 都非零, 从而是 $\boldsymbol{A}$ 的属于特征值 λ_i 的特征向量.

(2) 反过来, 设 $\boldsymbol{A}\boldsymbol{x}_i = \lambda_i \boldsymbol{x}_i$, 满足 $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ 线性无关. 此时, 用矩阵表示即有

$$\boldsymbol{A}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}.$$

将矩阵 $(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$ 记作 $\boldsymbol{T}$. 由于 $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ 线性无关, 方阵 $\boldsymbol{T}$ 列满秩, 从而可逆. 此时, 对于上式左乘矩阵 $\boldsymbol{T}^{-1}$, 即有 $\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T} = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$. 这说明方阵 $\boldsymbol{A}$ 可以对角化. □

注 6.4.3. (1) 前面的证明说明, 若方阵 $\boldsymbol{A}$ 可对角化, 则将其对角化的矩阵 $\boldsymbol{T}$ 的列向量为 $\boldsymbol{A}$ 的特征向量, 而对角化后的对角阵上的主对角线上的元素为 $\boldsymbol{A}$ 的特征值.

(2) 前面的定理说明, 方阵 $\boldsymbol{A}$ 可对角化的充要条件是 $\boldsymbol{A}$ 有足够多的特征向量来形成 F^n 的基. 我们不妨称这样的基为**特征向量基**.

(3) 用来对角化的矩阵 $\boldsymbol{T}$ 若存在, 并不是唯一的. 例如, 将 $\boldsymbol{T}$ 的各列重新排列, 或者将它们分别乘以不同的非零标量, 这样得到的新矩阵仍然可以将 $\boldsymbol{A}$ 对角化.

接下来讨论矩阵特征向量之间的线性相关性.

命题 6.4.4. 设 $\boldsymbol{A}$ 是数域 F 上的 n 阶方阵, $\lambda_1, \ldots, \lambda_m$ 为 $\boldsymbol{A}$ 的两两互不相等的特征值, $\{\, \boldsymbol{x}_{i,j} \mid 1 \le j \le k_i \,\}$ 为属于 λ_i 的线性无关的特征向量, 则 $\{\, \boldsymbol{x}_{i,j} \mid 1 \le i \le m, 1 \le j \le k_i \,\}$ 线性无关.

证明. 设 $\mu_{i,j} \in F$ 使得

$$\sum_{i=1}^{m} \underbrace{\sum_{j=1}^{k_i} \mu_{i,j} \boldsymbol{x}_{i,j}}_{\text{记作 } \boldsymbol{x}_i} = \boldsymbol{0},$$

从而有 $\sum_i \boldsymbol{x}_i = \boldsymbol{0}$. 由于 $\boldsymbol{A}\boldsymbol{x}_{i,j} = \lambda_i \boldsymbol{x}_{i,j}$, 容易验证, $\boldsymbol{A}\boldsymbol{x}_i = \lambda_i \boldsymbol{x}_i$. 对于 $k = 1, 2, \ldots, m-1$, 左乘 $\boldsymbol{A}^k$, 我们有

$$\boldsymbol{0} = \boldsymbol{A}^k \sum_{i=1}^{m} \boldsymbol{x}_i = \sum_{i=1}^{m} \boldsymbol{A}^k \boldsymbol{x}_i = \sum_{i=1}^{m} \lambda_i^k \boldsymbol{x}_i.$$

写成矩阵形式后, 我们有

$$(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m) \begin{pmatrix} \lambda_1^k \\ \vdots \\ \lambda_m^k \end{pmatrix} = \boldsymbol{0},$$

从而进一步有

$$(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_m) \begin{pmatrix} 1 & \lambda_1 & \lambda_1^2 & \cdots & \lambda_1^{m-1} \\ 1 & \lambda_2 & \lambda_2^2 & \cdots & \lambda_2^{m-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \lambda_m & \lambda_m^2 & \cdots & \lambda_m^{m-1} \end{pmatrix} = \boldsymbol{O}.$$

由于 $\lambda_1, \ldots, \lambda_m$ 互不相等, 其中的 Vandermonde 矩阵可逆, 这说明

$$(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_m) = \boldsymbol{O}.$$

从而每个 $\boldsymbol{x}_i$ 都是零向量, 即 $\sum_j \mu_{ij} \boldsymbol{x}_{ij} = \boldsymbol{0}$. 对于这个 i, 由于向量组 $\{\, \boldsymbol{x}_{i,j} \mid 1 \le j \le k_i \,\}$ 线性无关, 故每个 $\mu_{i,j} = 0$. 这说明 $\{\, \boldsymbol{x}_{i,j} \mid 1 \le i \le m, 1 \le j \le k_i \,\}$ 线性无关. □

例 6.4.5. (1) *在考察 Fibonacci 数列时, 我们遇到了矩阵* $\boldsymbol{A} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$. *不难求得,* $\boldsymbol{A}$ *有特征值* $\lambda_{1,2} = \frac{1 \pm \sqrt{5}}{2}$, *相应地有特征向量* $\boldsymbol{x}_{1,2} = (\frac{1 \pm \sqrt{5}}{2}, 1)^{\mathsf{T}}$. *基于此, 若取* $\boldsymbol{T} = \begin{pmatrix} \frac{1+\sqrt{5}}{2} & \frac{1-\sqrt{5}}{2} \\ 1 & 1 \end{pmatrix}$, *则必有* $\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T} = \operatorname{diag}(\frac{1+\sqrt{5}}{2}, \frac{1-\sqrt{5}}{2})$.

(2) *在例 6.3.10 中, 对于矩阵* $\boldsymbol{A} = \begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{pmatrix}$, *我们求出了对于特征值* $\lambda_1 = 5$ *的一个特征向量* $\boldsymbol{x}_1 = (1, 1, 1)^{\mathsf{T}}$, *和对于特征值* $\lambda_2 = \lambda_3 = -1$ *的两个线性无关的特征向量* $\boldsymbol{x}_2 = (-1, 1, 0)^{\mathsf{T}}$ *和* $\boldsymbol{x}_3 = (-1, 0, 1)^{\mathsf{T}}$. *由命题 6.4.4 可知,* $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3$ *线*

性无关, 从而矩阵 $\boldsymbol{T}=(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{x}_3)$ 是可逆矩阵. 由前面的讨论可知, $\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}=\operatorname{diag}(\lambda_1,\lambda_2,\lambda_3)=\operatorname{diag}(5,-1,-1)$.

习题 6.4.6. 设 $\mathscr{A}$ 是 $\mathbb{R}^2$ 上的线性变换, 由 $\boldsymbol{x}\mapsto\boldsymbol{A}\boldsymbol{x}$ 给出, 其中 $\boldsymbol{A}=\begin{pmatrix}7&2\\-4&1\end{pmatrix}$. 求 $\mathbb{R}^2$ 的一组基, 使得 $\mathscr{A}$ 在这组基下的矩阵是对角阵.

习题 6.4.7. 对于 n 阶实方阵 $\boldsymbol{A}$, 我们定义

$$\sin(\boldsymbol{A}):=\boldsymbol{A}-\frac{1}{3!}\boldsymbol{A}^3+\frac{1}{5!}\boldsymbol{A}^5-\cdots.$$

对于

$$\boldsymbol{A}=\frac{\pi}{4}\begin{pmatrix}7&-3\\-3&7\end{pmatrix},$$

证明 $\sin(\boldsymbol{A})$ 收敛, 并给出计算结果.

推论 6.4.8. 设 $\boldsymbol{A}$ 是数域 F 上的 n 阶方阵, $\lambda_1,\dots,\lambda_m$ 为 $\boldsymbol{A}$ 的两两互不相等的特征值, $\boldsymbol{x}_i$ 为属于 λ_i 的特征向量, 则 $\boldsymbol{x}_1,\dots,\boldsymbol{x}_m$ 线性无关.

证明. 对于每个 i $(1\le i\le m)$, 由一个非零向量 $\boldsymbol{x}_i$ 构成的向量组显然是线性无关的, 从而可以使用命题 6.4.4 中的结果. □

另外, 如下的推论给出了一个常用的矩阵可对角化的充分条件.

推论 6.4.9. 如果 n 阶方阵 $\boldsymbol{A}$ 的 n 个特征值两两不同, 则 $\boldsymbol{A}$ 可以对角化.

证明. 对于 $\boldsymbol{A}$ 的每个特征值 λ_i, $i=1,2,\dots,n$, 由定义知至少存在一个属于 λ_i 的特征向量 $\boldsymbol{x}_i$. 推论 6.4.8 保证了 $\boldsymbol{x}_1,\dots,\boldsymbol{x}_n$ 是线性无关的. 再利用定理 6.4.2, 由此可以推出 $\boldsymbol{A}$ 可以对角化. □

注 6.4.10. 若 $\boldsymbol{A}$ 为 n 阶对角阵, 其对角线上的元素各不相等. 那么我们可以证明: 同阶方阵 $\boldsymbol{B}$ 与 $\boldsymbol{A}$ 乘法可交换的充要条件是 $\boldsymbol{B}$ 为对角阵. 教材习题 *P192#22* 给出了与之等价的形式, 留作课后作业.

例 6.4.11. 设 $\boldsymbol{A}$ 为数域 F 上的 n 阶方阵, 有 n 个互不相等的特征值 $\lambda_1,\dots,\lambda_n$, 对应分别有特征向量 $\boldsymbol{x}_1,\dots,\boldsymbol{x}_n$. 记 $\boldsymbol{T}=(\boldsymbol{x}_1,\dots,\boldsymbol{x}_n)$, 则 $\boldsymbol{T}$ 是一个可逆方阵, 并且 $\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}=\operatorname{diag}(\lambda_1,\dots,\lambda_n)\xlongequal{\text{记作}}\boldsymbol{B}$.

与教材 P157 的作业题 #48 题相关

容易验证 $V=\{\,\boldsymbol{C}\in F^{n\times n}\mid \boldsymbol{C}\boldsymbol{A}=\boldsymbol{A}\boldsymbol{C}\,\}$ 是 $F^{n\times n}$ 的一个子空间. 同样地, 易验证 $\boldsymbol{C}\in V$ 当且仅当 $\boldsymbol{T}\boldsymbol{B}\boldsymbol{T}^{-1}\boldsymbol{C}=\boldsymbol{C}\boldsymbol{T}\boldsymbol{B}\boldsymbol{T}^{-1}$, 当且仅当 $\boldsymbol{B}(\boldsymbol{T}^{-1}\boldsymbol{C}\boldsymbol{T})=(\boldsymbol{T}^{-1}\boldsymbol{C}\boldsymbol{T})\boldsymbol{B}$. 由于

$\boldsymbol{B}$ 是主对角线上元素互不相等的对角阵, 这等价于说 $\boldsymbol{T}^{-1}\boldsymbol{C}\boldsymbol{T}$ 是一个 F 上的对角阵. 由此, 我们看出 $\dim(V)=n$, 而 V 有一组基 $\{\,\boldsymbol{T}\boldsymbol{E}_{ii}\boldsymbol{T}^{-1} \mid 1\le i\le n\,\}$.

接下来, 我们给出 V 的另外一组基. 若记之前提到的对角阵 $\boldsymbol{T}^{-1}\boldsymbol{C}\boldsymbol{T}=\operatorname{diag}(\mu_1,\dots,\mu_n)$, 由 *Lagrange* 插值公式 (教材 *P115#32*) 可知, 存在 $f(x)\in F_{n-1}[x]$, 使得 $f(\lambda_i)=\mu_i$, $1\le i\le n$. 这说明矩阵的多项式 $f(\boldsymbol{B})=\boldsymbol{T}^{-1}\boldsymbol{C}\boldsymbol{T}$, 即

$$\boldsymbol{C}=\boldsymbol{T}f(\boldsymbol{B})\boldsymbol{T}^{-1}=f(\boldsymbol{T}\boldsymbol{B}\boldsymbol{T}^{-1})=f(\boldsymbol{A}).$$

这说明 $\boldsymbol{C}$ 可由 $\{\,\boldsymbol{I},\boldsymbol{A},\boldsymbol{A}^2,\dots,\boldsymbol{A}^{n-1}\,\}$ 生成. 由于这组向量显然在 V 中, 它们构成了 V 的一组生成元. 又由于已知 $\dim(V)=n$, 这足以说明 $\{\,\boldsymbol{I},\boldsymbol{A},\boldsymbol{A}^2,\dots,\boldsymbol{A}^{n-1}\,\}$ 也构成了 V 的一组基.

若更一般地, $\boldsymbol{A}$ 不可对角化, 我们也有相应的结果; 请参考习题 *6.5.3*.

特征值的代数重数与几何重数 (※) 对于复数域 $\mathbb{C}$ 上的 n 阶矩阵 $\boldsymbol{A}$, 设其特征多项式为 这是书上的打星号内容, 考试不作要求, 但是强烈建议掌握

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

其中 $\lambda_1,\dots,\lambda_s$ 是 $\boldsymbol{A}$ 的所有两两不同的特征值 (不计重数). $n_i\ge 1$ 称为 λ_i 的**代数重数** (algebraic multiplicity). 对于 λ_i, 考察特征子空间

$$V_{\boldsymbol{A}}(\lambda_i)=\{\,\boldsymbol{x}\in F^n \mid \boldsymbol{A}\boldsymbol{x}=\lambda_i\boldsymbol{x}\,\}.$$

则向量空间维数 $m_i=\dim(V_{\boldsymbol{A}}(\lambda_i))\ge 1$ 称为 λ_i 的**几何重数** (geometric multiplicity).

定理 6.4.12. 用上面的记号, 则对于每个 i, 我们总有 $1\le m_i\le n_i$. 而 $\boldsymbol{A}$ 可以对角化的充要条件是对于每个 i, 等号成立: $m_i=n_i$. 教材引理 6.4.3, 定理 6.4.2

证明. 我们先证明 $1\le m_i\le n_i$. 不妨设 $i=1$. 由于 λ_1 为 $\boldsymbol{A}$ 的特征值, 矩阵 $\lambda_1\boldsymbol{I}-\boldsymbol{A}$ 的行列式为 0, 从而该矩阵不可逆, 这意味着特征子空间 $V_{\boldsymbol{A}}(\lambda_1)$ 不是零空间, 从而 $m_1\ge 1$. 此时, 可以找到 $V_{\boldsymbol{A}}(\lambda_1)$ 的一组基 $\boldsymbol{x}_1,\dots,\boldsymbol{x}_{m_1}$. 我们可以将其扩充为 F^n 的一组基 $\boldsymbol{x}_1,\dots,\boldsymbol{x}_{m_1},\dots,\boldsymbol{x}_n$. 记 $\boldsymbol{T}=(\boldsymbol{x}_1,\dots,\boldsymbol{x}_n)$ 为由这些列向量按行排列构成的方阵. 由于 $\boldsymbol{A}\boldsymbol{x}_i=\lambda_1\boldsymbol{x}_i$ $(i=1,2,\dots,m_1)$, 我们有 $\boldsymbol{A}\boldsymbol{T}=(\lambda_1\boldsymbol{x}_1,\lambda_1\boldsymbol{x}_2,\dots,\lambda_1\boldsymbol{x}_{m_1},*,\dots,*)$, 这说明 $\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}=\begin{pmatrix}\lambda_1\boldsymbol{I}_{m_1} & \boldsymbol{B}_{12}\\ \boldsymbol{O} & \boldsymbol{B}_{22}\end{pmatrix}$. 我们将这个矩阵记作 $\boldsymbol{B}$, 于是 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 通过 $\boldsymbol{T}$ 相似, 从而有相等的特征多项式, 即 $p_{\boldsymbol{A}}(\lambda)=p_{\boldsymbol{B}}(\lambda)=(\lambda-\lambda_1)^{m_1}p_{\boldsymbol{B}_{22}}(\lambda)$. 这说明 λ_1 作为 $\boldsymbol{A}$ 的特征值的重数至少为 m_1, 从而 $m_1\le n_1$. 定理的证明与本节的例子, 留给学生课后阅读.

我们接下来证明: $\boldsymbol{A}$ 可以对角化的充要条件是对于每个 i, 都有 $m_i=n_i$. 先假设几何重数为代数重数这一条件得到满足, 于是, 我们在特征子空间 $V_{\boldsymbol{A}}(\lambda_i)$ 中我们可以找到

一组基 $\boldsymbol{x}_{i,1}, \boldsymbol{x}_{i,2}, \dots, \boldsymbol{x}_{i,n_i}$. 由于 $\sum_i n_i = n$, 若将这些向量组合在一起, 我们得到 F^n 的一组基 $\{\, \boldsymbol{x}_{i,j} : 1 \le i \le s, 1 \le j \le n_i \,\}$. 这说明 $\boldsymbol{A}$ 可以相似对角化.

反之, 设 $\boldsymbol{A}$ 可以相似对角化. 于是, 存在 $\boldsymbol{T}$ 使得 $\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T} = \operatorname{diag}(\mu_1, \dots, \mu_n)$ 为对角阵. 设 $\boldsymbol{T}$ 的列向量为 $\boldsymbol{x}_1, \dots, \boldsymbol{x}_n$. 由于 λ_1 的代数重数为 n_1, $\{\, i : \mu_i = \lambda_1 \,\} = \{i_1 < i_2 < \cdots < i_{n_1}\}$ 是一个 n_1 元集合. 此时, $\boldsymbol{x}_{i_1}, \boldsymbol{x}_{i_2}, \dots, \boldsymbol{x}_{i_{n_1}}$ 是属于 λ_1 的特征向量. 由于 $\boldsymbol{T}$ 可逆, 这是特征子空间 $V_{\boldsymbol{A}}(\lambda_1)$ 中线性无关的向量组, 从而 $n_1 \le m_1$. 但是在一开始的位置我们已经证明了 $m_1 \le n_1$, 从而 $m_1 = n_1$. 当然, 对于其它特征值的证明也是类似的. $\square$

注 6.4.13. 综合之前的讨论, 我们总结一下求相似对角阵的方法. 设 $\boldsymbol{A}$ 是给定的 n 阶方阵.

(1) 求 $\boldsymbol{A}$ 的特征值, 得到 $p_{\boldsymbol{A}}(\lambda) = \det(\lambda \boldsymbol{I}_n - \boldsymbol{A}) = \prod_{i=1}^{s} (\lambda - \lambda_i)^{n_i}$, 其中 $\lambda_1, \lambda_2, \dots, \lambda_s$ 是 $\boldsymbol{A}$ 的所有两两不同的特征值.

(2) 若对每个 i 都有 $n - \operatorname{rank}(\lambda_i \boldsymbol{I}_n - \boldsymbol{A}) = n_i$, 则 $\boldsymbol{A}$ 可对角化; 否则, $\boldsymbol{A}$ 不可对角化.

(3) 在可对角化的前提下, 对每个 λ_i, 求出方程组 $(\lambda_i \boldsymbol{I}_n - \boldsymbol{A})\boldsymbol{X} = \boldsymbol{0}$ 的一组基础解系 $\boldsymbol{x}_{i1}, \boldsymbol{x}_{i2}, \dots, \boldsymbol{x}_{in_i}$.

(4) 令 $\boldsymbol{T} = (\boldsymbol{x}_{11}, \boldsymbol{x}_{12}, \dots, \boldsymbol{x}_{1n_1}, \boldsymbol{x}_{21}, \boldsymbol{x}_{22}, \dots, \boldsymbol{x}_{in_2}, \dots, \boldsymbol{x}_{s1}, \boldsymbol{x}_{s2}, \dots, \boldsymbol{x}_{sn_s})$, 则 $\boldsymbol{T}$ 是可逆方阵, 并满足

$$\boldsymbol{T}^{-1} A \boldsymbol{T} = \operatorname{diag}(\underbrace{\lambda_1, \dots, \lambda_1}_{n_1\text{ 个}}, \underbrace{\lambda_2, \dots, \lambda_2}_{n_2\text{ 个}}, \dots, \underbrace{\lambda_s, \dots, \lambda_s}_{n_s\text{ 个}}).$$

例 6.4.14. 重新考察矩阵 $\boldsymbol{A} = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$. 则 $p_{\boldsymbol{A}}(\lambda) = (\lambda - 2)^3$, 故 $\boldsymbol{A}$ 的特征值 2 的代数重数为 3. 另一方面, 相应的特征子空间为

$$V_{\boldsymbol{A}}(2) = \left\{\, \boldsymbol{x} \in F^3 \mid \boldsymbol{A}\boldsymbol{x} = 2\boldsymbol{x} \,\right\},$$

即线性方程组 $(2\boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{0}$ 的解空间. 由于

$$\operatorname{rank}(2\boldsymbol{I} - \boldsymbol{A}) = \operatorname{rank} \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix} = 2,$$

故解空间的维数为 $3 - 2 = 1$, 即特征值 2 的几何重数为 1. 由于 $1 < 3$, 故原矩阵 $\boldsymbol{A}$ 不可对角化.

例 6.4.15. *对于下三角方阵* $\boldsymbol{A}=\begin{pmatrix}1&0&0&0\\a&1&0&0\\2&b&2&0\\3&2&c&2\end{pmatrix}$*, 问* a,b,c *各为何值时,* $\boldsymbol{A}$ *可以相似对角化?*

解. 容易看出, 特征多项式 $p_{\boldsymbol{A}}(\lambda)=(\lambda-1)^2(\lambda-2)^2$, 从而特征值为 $\lambda_1=1$, $\lambda_2=2$, 皆为两重. 此时, $\boldsymbol{A}$ 可以相似对角化的充要条件是 $\lambda_1=1$ 与 $\lambda_2=2$ 的几何重数皆为 2.

(1) 当 $\lambda=1$, $\lambda\boldsymbol{I}-\boldsymbol{A}=\begin{pmatrix}0&0&0&0\\-a&0&0&0\\-2&-b&\boxed{\text{-1}}&0\\-3&-2&-c&\boxed{\text{-1}}\end{pmatrix}$. 则 $\lambda_1=1$ 的几何重数为 2, 当且仅当 $2=4-\operatorname{rank}(\boldsymbol{I}-\boldsymbol{A})$, 当且仅当 $\operatorname{rank}(\boldsymbol{A})=2$. 运用初等列变换, 我们可以用标记的 -1 把矩阵的第三行和第四行的其它数消去. 由于初等变化不改变矩阵的秩, 我们可以由此推出 $a=0$, 此时对于 b 和 c 没有限制.

(2) 类似地, 考虑 $\lambda=2$. 可以推出 $c=0$, 对于 a 和 b 没有限制.

综上, $a=c=0$, 而 b 任意. □

例 6.4.16. *已知* $\boldsymbol{A}_t=\begin{pmatrix}t&t-2&4-2t\\3&-1&0\\1+t&t-2&3-2t\end{pmatrix}$*. 若* $\boldsymbol{A}_t$ *可对角化, 描述此时的* t*, 并求出* $\boldsymbol{P}$*, 使得* $\boldsymbol{P}^{-1}\boldsymbol{A}_t\boldsymbol{P}$ *是对角阵.*

解. 特征多项式 $p_{\boldsymbol{A}_t}(\lambda)=(\lambda-2)(\lambda+1)(\lambda-1+t)$, 故特征值为 $2,-1,1-t$.

(1) 若 $t\neq-1,2$, 则 $\boldsymbol{A}_t$ 有三个不同的特征值, 可以对角化. 容易求得 $\boldsymbol{P}=\begin{pmatrix}1&0&t-2\\1&2&-3\\1&1&t-2\end{pmatrix}$, 使得 $\boldsymbol{P}^{-1}\boldsymbol{A}_t\boldsymbol{P}=\operatorname{diag}(2,-1,1-t)$.

(2) 若 $t=2$, 容易求得 $\boldsymbol{P}=\begin{pmatrix}1&0&0\\1&2&-3\\1&1&0\end{pmatrix}$, 使得 $\boldsymbol{P}^{-1}\boldsymbol{A}_t\boldsymbol{P}=\operatorname{diag}(2,-1,-1)$.

(3) 若 $t=-1$, 则特征值 2 的几何重数为 1, 小于代数重数, 从而矩阵不可对角化. □

例 6.4.17. *设方阵* $\boldsymbol{A}$ *为***幂等矩阵***:* $\boldsymbol{A}^2=\boldsymbol{A}$*. 证明:*

(1) $\boldsymbol{A}$ 的特征值只有 0 和 1;

(2) $\boldsymbol{A}$ 相似于其相抵标准形 $\begin{pmatrix} \boldsymbol{I}_r & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} \end{pmatrix}$, 其中 $r = \operatorname{rank}(\boldsymbol{A})$;

(3) $\operatorname{tr}(\boldsymbol{A}) = \operatorname{rank}(\boldsymbol{A})$. (教材第六章作业题 #27)

证明. (1) 设 $\boldsymbol{x}$ 是 $\boldsymbol{A}$ 的属于特征值 λ 的特征向量: $\boldsymbol{A}\boldsymbol{x} = \lambda\boldsymbol{x}$. 同时左乘 $\boldsymbol{A}$, 我们有 $\boldsymbol{A}^2\boldsymbol{x} = \boldsymbol{A}(\lambda\boldsymbol{x}) = \lambda\boldsymbol{A}\boldsymbol{x} = \lambda^2\boldsymbol{x}$. 另一方面, 利用幂等性质, $\boldsymbol{A}^2\boldsymbol{x} = \boldsymbol{A}\boldsymbol{x} = \lambda\boldsymbol{x}$. 由于 $\boldsymbol{x}$ 不是零向量, 这说明 $\lambda^2 = \lambda$, 从而 $\lambda = 0$ 或 1.

(2) 对于 (可能的) 特征值 $\lambda_1 = 0$, 其几何重数为

$$m_1 = \dim \{ \boldsymbol{x} \mid (0\boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{0} \} = n - \operatorname{rank}(\boldsymbol{A}).$$

类似地, 对于 (可能的) 特征值 $\lambda_2 = 1$, 其几何重数为 $m_2 = n - \operatorname{rank}(\boldsymbol{I} - \boldsymbol{A})$. 以前, 我们已经推导过, 对于幂等矩阵 $\boldsymbol{A}$, 有

$$\operatorname{rank}(\boldsymbol{A}) + \operatorname{rank}(\boldsymbol{I} - \boldsymbol{A}) = n.$$

这说明 $m_1 + m_2 = n$. 但是, 我们知道特征值的代数重数不小于几何重数: $n_1 \geq m_1$, $n_2 \geq m_2$. 另一方面, 我们总有 $n_1 + n_2 = n$. 这说明 $n_1 = m_1$, $n_2 = m_2$. 从而对于 $\boldsymbol{A}$ 的所有特征值, 其相应的代数重数总是等于几何重数. 从而 $\boldsymbol{A}$ 可以相似对角化. 由定理 6.4.2 的证明, 我们可以假设这些 1 都在对角形的左上角. (注: 若 λ 不是真正的特征值, 则相应的 "代数重数" 与 "几何重数" 皆为零, 不实质影响这儿的讨论.) 从而此时, 有 $\boldsymbol{A} \sim \boldsymbol{D} := \begin{pmatrix} \boldsymbol{I}_r & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} \end{pmatrix}$. 由于矩阵的秩是相似不变量, 我们有 $r = \operatorname{rank}(\boldsymbol{D}) = \operatorname{rank}(\boldsymbol{A})$.

(3) 从前一问的讨论来看, 这是显然的. 另外, 在例 4.5.19 中, 我们给出了另外一个证明. 除此之外, 也可以看看教材上用若尔当标准形的方法重新考虑这道例题 (P188 例 6.5.6). □