

4.3. 可对角化的矩阵

矩阵 $\mathbf{A} \in M_n(\mathbb{R})$ 可对角化当且仅当 $\mathbf{A}$ 存在 n 个线性无关的特征向量.

要判断 $\mathbf{A}$ 是否可对角化, 可以

- 找出 $\mathbf{A}$ 的所有特征值, 通过解方程 $f_{\mathbf{A}}(t) = 0$ 实现.
- 计算出每个特征值对应的特征子空间的一组基.
- 从中寻找线性无关的特征向量.

问题: 属于不同特征值的特征向量组成的向量组是否线性无关?

引理 (4.9)

假设 $\lambda_1, \lambda_2, \dots, \lambda_r \in \mathbb{R}$ 是 $\mathbf{A} \in M_n(\mathbb{R})$ 的不同特征值.

假设 $\mathcal{B}_i$ 是特征子空间 V_{λ_i} 中线性无关的向量组, $i = 1, \dots, r$

则这些向量组拼成的向量组

$$\mathcal{B}_1, \mathcal{B}_2, \dots, \mathcal{B}_r$$

线性无关.

推论 (4.10)

假设 $\eta_1, \eta_2, \dots, \eta_r$ 是矩阵 $\mathbf{A}$ 的对应不同特征值的特征向量, 那么 $\eta_1, \eta_2, \dots, \eta_r$ 线性无关.

定理 (4.11)

假设 $\mathbf{A} \in M_n(\mathbb{R})$, 而 $\lambda_1, \lambda_2, \dots, \lambda_r \in \mathbb{R}$ 是 $\mathbf{A}$ 所有互不相同的特征值. 则 $\mathbf{A}$ 可对角化当且仅当

$$\dim V_{\lambda_1} + \dim V_{\lambda_2} + \dots + \dim V_{\lambda_r} = n,$$

当且仅当

$$\text{rank}(\lambda_1 \mathbf{I} - \mathbf{A}) + \text{rank}(\lambda_2 \mathbf{I} - \mathbf{A}) + \dots + \text{rank}(\lambda_r \mathbf{I} - \mathbf{A}) = (r-1)n.$$

例

判断下面的实矩阵是否可对角化.

$$\mathbf{A} = \begin{pmatrix} 3 & 2 & -1 \\ -2 & -2 & 2 \\ 3 & 6 & -1 \end{pmatrix}.$$

解:先计算 $\mathbf{A}$ 的特征多项式

$$f_{\mathbf{A}}(t) = \begin{vmatrix} t-3 & -2 & 1 \\ 2 & t+2 & -2 \\ -3 & -6 & t+1 \end{vmatrix} = (t-2)^2(t+4).$$

不同的特征值分别是 $\lambda_1 = 2, \lambda_2 = -4$.

再计算特征子空间维数

$$\dim V_{\lambda_1} = 3 - \text{rank}(2\mathbf{I} - \mathbf{A}) = 2,$$

$$\dim V_{\lambda_2} = 3 - \text{rank}(-4\mathbf{I} - \mathbf{A}) = 1.$$

由于两个特征子空间维数之和为 3, 由定理 4.11, $\mathbf{A}$ 可对角化.

实际上, 可以求出一个可逆矩阵 $\mathbf{T}$, 使得 $\mathbf{T}^{-1}\mathbf{A}\mathbf{T}$ 是对角矩阵.

先求解齐次线性方程组 $(\mathbf{A} - 2\mathbf{I})\mathbf{x} = \mathbf{0}$ 可得基础解系

$$\boldsymbol{\eta}_1 = (-2, 1, 0)^T, \quad \boldsymbol{\eta}_2 = (1, 0, 1)^T.$$

它们组成 $\mathbf{A}$ 的特征子空间 V_{λ_1} 的一组基

再求解齐次线性方程组 $(\mathbf{A} + 4\mathbf{I})\mathbf{x} = \mathbf{0}$ 的基础解系

$$\boldsymbol{\eta}_3 = (1, -2, 3)^T.$$

这就找到了 3 个线性无关特征向量 $\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \boldsymbol{\eta}_3$. 将它们拼成矩阵

$$\mathbf{T} = (\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \boldsymbol{\eta}_3) = \begin{pmatrix} -2 & 1 & 1 \\ 1 & 0 & -2 \\ 0 & 1 & 3 \end{pmatrix}.$$

则 $\mathbf{T}$ 可逆且

$$\mathbf{AT} = (\mathbf{A}\boldsymbol{\eta}_1, \mathbf{A}\boldsymbol{\eta}_2, \mathbf{A}\boldsymbol{\eta}_3) = (\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \boldsymbol{\eta}_3) \begin{pmatrix} 2 & & \\ & 2 & \\ & & -4 \end{pmatrix} = \mathbf{TD}$$

所以 $\mathbf{T}^{-1}\mathbf{AT} = \mathbf{D} = \text{diag}[2, 2, -4]$ 是对角矩阵.

下面我们对特征子空间的维数进行估计.

假设 λ 是 $\mathbf{A}$ 的特征值, 对应的特征子空间 V_λ 的维数是 m .

取 V_λ 的一组基 $\beta_1, \beta_2, \dots, \beta_m$, 再扩展成 $\mathbb{R}^n$ 的一组基.

那么 $\mathbf{A}$ 作为线性变换在这组基下的矩阵形如

$$\mathbf{B} := \begin{pmatrix} \lambda \mathbf{I}_m & \mathbf{C} \\ \mathbf{O} & \mathbf{D} \end{pmatrix}.$$

由于线性变换在不同基下的矩阵相似, 因此 $\mathbf{B}$ 与 $\mathbf{A}$ 相似, 从而

$$f_{\mathbf{A}}(t) = f_{\mathbf{B}}(t) = \begin{vmatrix} (t - \lambda)\mathbf{I}_m & -\mathbf{C} \\ \mathbf{O} & t\mathbf{I} - \mathbf{D} \end{vmatrix} = (t - \lambda)^m g(t),$$

其中 $g(t) = |t\mathbf{I} - \mathbf{D}|$.

把 $\mathbf{A}$ 的特征多项式分解为

$$f_{\mathbf{A}}(t) = (t - \lambda)^{n_\lambda} h(t), \quad h(\lambda) \neq 0$$

那么有 $\dim V_\lambda = m \leq n_\lambda$.

V_λ 的维数称为特征值 λ 的几何重数(geometric multiplicity).

而 n_λ 称为 λ 的代数重数(algebraic multiplicity).

由上面的讨论可知, 一个特征值的几何重数不超过其代数重数.

A 可对角化当且仅当

$$\sum_{i=1}^r \dim V_{\lambda_i} = \dim V,$$

其中 $\lambda_1, \lambda_2, \dots, \lambda_r$ 为 **A** 在 $\mathbb{R}$ 中的全部互不相同的特征值, 即 $f_{\mathbf{A}}(t)$ 在 $\mathbb{R}$ 中的所有互不相同的根.

假设它们的重数 (即代数重数) 分别为 $n_1, n_2, \dots, n_r$, 则

$$f_{\mathbf{A}}(t) = (t - \lambda_1)^{n_1} (t - \lambda_2)^{n_2} \cdots (t - \lambda_r)^{n_r} h(t),$$

其中 $h(t)$ 在 $\mathbb{R}$ 中没有根.

由于特征值的几何重数总是不超过代数重数, 有

$$1 \leq \dim V_{\lambda_i} \leq n_i, \quad i = 1, 2, \dots, r.$$

从而有

$$\begin{aligned} \sum_{i=1}^r \dim V_{\lambda_i} &\leq \sum_{i=1}^r n_i \leq \deg h(t) + \sum_{i=1}^r n_i \\ &= \deg f_A(t) = \dim V, \end{aligned}$$

我们发现, 如果某个特征值的几何重数小于代数重数, 或者 $h(t)$ 的次数大于零, 都有

$$\sum_{i=1}^r \dim V_{\lambda_i} < \dim V$$

从而 $\mathbf{A}$ 不可对角化.

定理 (4.12)

方阵 $A \in M_n(\mathbb{R})$ 可对角化当且仅当下面两个条件成立.

- ① A 的特征多项式可以分解为 $\mathbb{R}$ 系数的一次因式的乘积;
- ② 对 A 的每个特征值, 其几何重数等于其代数重数.

推论 (4.13)

若 $A \in M_n(\mathbb{R})$ 在 $\mathbb{R}$ 中有 n 个互不相同特征值. 则 A 可对角化.

斐波那契 (Fibonacci) 数列的通项公式. 斐波那契数列是由数学家斐波那契以兔子繁殖为例引入的, 也称为黄金分割序列. 它的定义为 $F_0 = 0$, $F_1 = 1$, 对 $n \geq 2$, 有

$$F_n = F_{n-1} + F_{n-2}$$

如何求其通项公式? 可以将递推关系写成矩阵乘法形式

$$\begin{pmatrix} F_{n+1} \\ F_n \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} F_n \\ F_{n-1} \end{pmatrix}.$$

记

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

则有

$$\begin{pmatrix} F_{n+1} \\ F_n \end{pmatrix} = \mathbf{A} \begin{pmatrix} F_n \\ F_{n-1} \end{pmatrix} = \cdots = \mathbf{A}^n \begin{pmatrix} F_1 \\ F_0 \end{pmatrix} = \mathbf{A}^n \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \quad (1)$$

问题转化为求矩阵 $\mathbf{A}$ 的 n 次幂!

计算 $\mathbf{A}$ 的特征多项式

$$f_{\mathbf{A}}(t) = \begin{vmatrix} t-1 & -1 \\ -1 & t \end{vmatrix} = t^2 - t - 1 = \left(t - \frac{1 + \sqrt{5}}{2}\right) \left(t - \frac{1 - \sqrt{5}}{2}\right).$$

由推论 4.13, $\mathbf{A}$ 可对角化. 分别求出两个特征值

$$\lambda_1 = \frac{1 + \sqrt{5}}{2}, \quad \lambda_2 = \frac{1 - \sqrt{5}}{2}$$

和它们对应的特征向量

$$\boldsymbol{\eta}_1 = \begin{pmatrix} \lambda_1 \\ 1 \end{pmatrix}, \quad \boldsymbol{\eta}_2 = \begin{pmatrix} \lambda_2 \\ 1 \end{pmatrix}$$

将它们拼在一起可得

$$\mathbf{T} := (\boldsymbol{\eta}_1, \boldsymbol{\eta}_2) = \begin{pmatrix} \lambda_1 & \lambda_2 \\ 1 & 1 \end{pmatrix}, \quad \mathbf{T}^{-1} = \frac{1}{\lambda_1 - \lambda_2} \begin{pmatrix} 1 & -\lambda_2 \\ -1 & \lambda_1 \end{pmatrix}.$$

由 $\mathbf{A}\boldsymbol{\eta}_1 = \lambda_1\boldsymbol{\eta}_1$, $\mathbf{A}\boldsymbol{\eta}_2 = \lambda_2\boldsymbol{\eta}_2$, 得

$$\mathbf{A}\mathbf{T} = \mathbf{T} \begin{pmatrix} \lambda_1 & \\ & \lambda_2 \end{pmatrix}, \quad \mathbf{A} = \mathbf{T} \begin{pmatrix} \lambda_1 & \\ & \lambda_2 \end{pmatrix} \mathbf{T}^{-1}.$$

从而

$$\mathbf{A}^n = \mathbf{T} \begin{pmatrix} \lambda_1^n & \\ & \lambda_2^n \end{pmatrix} \mathbf{T}^{-1} = \frac{1}{\lambda_1 - \lambda_2} \begin{pmatrix} \lambda_1^{n+1} - \lambda_2^{n+1} & \lambda_1 \lambda_2^{n+1} - \lambda_2 \lambda_1^{n+1} \\ \lambda_1^n - \lambda_2^n & \lambda_1 \lambda_2^n - \lambda_2 \lambda_1^n \end{pmatrix}.$$

计算过程中需要注意利用根与系数关系 $\lambda_1 \lambda_2 = -1$.

由式 (1) 可得

$$F_n = \frac{\lambda_1^n - \lambda_2^n}{\lambda_1 - \lambda_2} = \frac{\lambda_1^n - (1 - \lambda_1)^n}{\sqrt{5}}, \quad \lambda_1 = \frac{1 + \sqrt{5}}{2}.$$