矩阵可对角化的充要条件

董晟渤

2021 年 8 月 29 日

摘要

本文从特征值与特征向量、特征子空间、代数重数与几何重数、初等因子等多个角度, 整理了矩阵可对角化的充要条件, 并借助特征值与特征向量, 给出了将矩阵对角化的方法. 同时, 本文根据矩阵与线性变换的对应关系, 也得到了线性变换可对角化的充要条件. 最后, 本文举例说明了矩阵对角化的应用, 并说明了后续的研究和学习的方向.

关键词: 线性代数, 矩阵, 线性变换, 对角化.

目录

1 前言

在矩阵理论中, 形如

$$\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}_{n\times n}$$

的矩阵叫做对角矩阵 (diagonal matrix), 通常也可以简记为 $\operatorname{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_n\}$.

对于一般的 n 阶方阵 $\boldsymbol{A} = [a_{ij}]_{n\times n}$, 为了简化问题, 我们通常需要将其化成对角矩阵; 同时, 对于 n 维空间 V 上的线性变换 $\mathscr{A}$ 而言, 我们任意取一组基, 都可以得到 $\mathscr{A}$ 的矩阵表示, 我们也希望可以找到一组基, 使得在这组基下, $\mathscr{A}$ 的矩阵表示是最简洁的. 结合矩阵的相似的概念, 我们给出如下定义.

定义 1.1. *设 $\mathscr{A}$ 是 n 维空间 V 上的线性变换, 若存在 V 中的一组基, 使得 $\mathscr{A}$ 在这组基下的矩阵是对角矩阵, 则称 $\mathscr{A}$ 是可对角化的 (diagonalizable).*

定义 1.2. *设 $\boldsymbol{A}$ 是 n 阶方阵, 若其与对角矩阵相似, 也即存在可逆矩阵 $\boldsymbol{P}$, 使得*

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \operatorname{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_n\},$$

则称矩阵 $\boldsymbol{A}$ 是可对角化的 (diagonalizable).

根据线性变换 $\mathscr{A}$ 和矩阵 $\boldsymbol{A}$ 的对应关系, 在接下来的讨论中, 不区分这两个概念.

2 特征值与特征向量

我们从研究矩阵的特征值与特征向量出发, 给出矩阵可对角化的第一个充要条件.

定理 2.1. *n 阶方阵 $\boldsymbol{A}$ 可对角化的充要条件是 $\boldsymbol{A}$ 有 n 个线性无关的特征向量.*

证明. 必要性. 设 $\boldsymbol{A}$ 可对角化, 则存在可逆矩阵 $\boldsymbol{P} = [\boldsymbol{P}_1, \boldsymbol{P}_2, \cdots, \boldsymbol{P}_n]$, 使得

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \operatorname{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_n\},$$

从而

$$\begin{aligned}\boldsymbol{A}[\boldsymbol{P}_1,\boldsymbol{P}_2,\cdots,\boldsymbol{P}_n]&=\operatorname{diag}\{\lambda_1,\lambda_2,\cdots,\lambda_n\}[\boldsymbol{P}_1,\boldsymbol{P}_2,\cdots,\boldsymbol{P}_n]\\&=[\lambda_1\boldsymbol{P}_1,\lambda_2\boldsymbol{P}_2,\cdots,\lambda_n\boldsymbol{P}_n].\end{aligned}$$

从而对任意的 $1\le i\le n$, 都有 $\boldsymbol{A}\boldsymbol{P}_i=\lambda_i\boldsymbol{P}_i$. 该式说明了 $\boldsymbol{P}$ 的每一列 $\boldsymbol{P}_i$ 所构成的列向量都是 $\boldsymbol{A}$ 的特征向量. 又由 $\boldsymbol{P}$ 可逆知 $\boldsymbol{P}_1,\boldsymbol{P}_2,\cdots,\boldsymbol{P}_n$ 线性无关, 因此 $\boldsymbol{A}$ 有 n 个线性无关的特征向量.

充分性. 设 $\boldsymbol{A}$ 有 n 个线性无关的特征向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_2$, 并且对任意的 $1\le i\le n$, 都有 $\boldsymbol{A}\boldsymbol{\alpha}_i=\lambda_i\boldsymbol{\alpha_i}$, 从而

$$\begin{aligned}\boldsymbol{A}[\boldsymbol{\alpha_1},\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n]&=[\lambda_1\boldsymbol{\alpha_1},\lambda_2\boldsymbol{\alpha}_2,\cdots,\lambda_n\boldsymbol{\alpha}_n]\\&=\operatorname{diag}\{\lambda_1,\lambda_2,\cdots,\lambda_n\}[\boldsymbol{\alpha_1},\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n].\end{aligned}$$

作矩阵 $\boldsymbol{P}=[\boldsymbol{\alpha_1},\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n]$, 由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_2$ 线性无关, 知矩阵 $\boldsymbol{P}$ 可逆, 并且

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\operatorname{diag}\{\lambda_1,\lambda_2,\cdots,\lambda_n\},$$

从而矩阵 $\boldsymbol{A}$ 可对角化. □

在这里, 我们指出, 根据线性变换和矩阵的对应关系, 我们可以得到类似的结论.

推论 2.2. *n 维线性空间 V 上的线性变换 $\mathscr{A}$ 可对角化的充要条件是 $\mathscr{A}$ 有 n 个线性无关的特征向量.*

同时, 注意到不同的特征值所对应的特征向量必定是线性无关的, 我们可以得到一个矩阵可对角化的充分不必要条件.

推论 2.3. *若 n 阶方阵 $\boldsymbol{A}$ 有 n 个不同的特征向量, 则 $\boldsymbol{A}$ 可对角化.*

上面的过程, 实质上是找到了一种将给定矩阵 $\boldsymbol{A}$ 对角化的方式, 并且对角矩阵的元素都是矩阵 $\boldsymbol{A}$ 的特征值. 为了体现这种方法的应用, 以下是一个简单的例子.

例 2.4. *将矩阵*

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

对角化.

解答 首先, 我们求出矩阵 $\boldsymbol{A}$ 的特征值和特征向量. 计算得

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

因此矩阵 $\boldsymbol{A}$ 的特征值分别为 $5,2,1$. 再代入方程

$$(\lambda\boldsymbol{I}-\boldsymbol{A})\boldsymbol{\alpha}_\lambda = \boldsymbol{0},$$

可以得到上述特征值所对的特征向量分别为 $\boldsymbol{\alpha}_5 = [2,-1,0]^T, \boldsymbol{\alpha}_2 = [1,1,0]^T, \boldsymbol{\alpha}_1 = [0,0,1]^T$. 记矩阵

$$\boldsymbol{P} = [\boldsymbol{\alpha}_5, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_1] = \begin{bmatrix} 2 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

计算得

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \mathrm{diag}\{5,2,1\}$$

为对角矩阵.

3 特征子空间

考虑矩阵 $\boldsymbol{A}$ 对应某个特征值 λ_i 的特征子空间 V_{λ_i}, 其具有许多重要的性质. 下面的定理便给出了矩阵可对角化的另一充要条件.

定理 3.1. 给定 n 阶方阵 $\boldsymbol{A}$, 设其的互异特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_r$, 对应的特征子空间为 $V_{\lambda_1}, V_{\lambda_2}, \cdots, V_{\lambda_r}$, 则 $\boldsymbol{A}$ 可对角化的充要条件是

$$V = V_{\lambda_1} \oplus V_{\lambda_2} \oplus \cdots \oplus V_{\lambda_r}.$$

证明. 必要性. 设 $\boldsymbol{A}$ 可对角化, 则由定理2.1知其有 n 个线性无关的特征向量. 对任意的

$1 \le i \le r$, 在 V_{λ_i} 中取一组基 $\boldsymbol{\alpha}_1^{(i)}, \boldsymbol{\alpha}_2^{(i)}, \cdots, \boldsymbol{\alpha}_{k_i}^{(i)}$, 则

$$\{\boldsymbol{\alpha}_j^{(i)} : 1 \le i \le r, 1 \le j \le k_i\}$$

是 $\boldsymbol{A}$ 的 n 个线性无关的特征向量, 从而其也是 V 的一组基. 并且对应不同特征值的特征子空间中的向量是线性无关的, 从而

$$V = V_{\lambda_1} \oplus V_{\lambda_2} \oplus \cdots \oplus V_{\lambda_r}.$$

充分性. 假设 V 是 $V_{\lambda_1}, V_{\lambda_2}, \cdots, V_{\lambda_r}$ 的直和, 对任意的 $1 \le i \le r$, 在 V_{λ_i} 中取一组基 $\boldsymbol{\alpha}_1^{(i)}, \boldsymbol{\alpha}_2^{(i)}, \cdots, \boldsymbol{\alpha}_{k_i}^{(i)}$, 则

$$\{\boldsymbol{\alpha}_j^{(i)} : 1 \le i \le r, 1 \le j \le k_i\}$$

是 V 的一组基, 从而其为 n 个线性无关的向量, 也即 $\boldsymbol{A}$ 有 n 个线性无关的特征向量, 从而由定理2.1知 $\boldsymbol{A}$ 可对角化. □

同样地, 考虑线性变换 $\mathscr{A}$ 的特征子空间, 也有类似的结论.

推论 3.2. 给定 n 维线性空间 V 上的线性变换 $\mathscr{A}$, 设其的互异特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_r$, 对应的特征子空间为 $V_{\lambda_1}, V_{\lambda_2}, \cdots, V_{\lambda_r}$, 则 $\mathscr{A}$ 可对角化的充要条件是

$$V = V_{\lambda_1} \oplus V_{\lambda_2} \oplus \cdots \oplus V_{\lambda_r}.$$

4 代数重数与几何重数

对于矩阵 $\boldsymbol{A}$ 的特征值而言, 还有两个重要的概念, 叫作代数重数和几何重数. 其中, 代数重数表示特征多项式中特征值所对应的根的重数, 几何重数表示特征子空间的维数. 并且我们知道, 某个特征值所对应的几何重数一定是小于等于代数重数的.

定理 4.1. n 阶方阵 $\boldsymbol{A}$ 可对角化的充要条件是对于 $\boldsymbol{A}$ 的每一个特征值, 其所对的代数重数和几何重数相等.

证明. 设矩阵 $\boldsymbol{A}$ 的特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_r$, 其中对任意的 $1 \le i \le r$, λ_i 的代数重数为

m_{λ_i}, 几何重数为 k_{λ_i}, 则 $k_{\lambda_i} \leq m_{\lambda_i}$, 且

$$m_{\lambda_1} + m_{\lambda_2} + \cdots + m_{\lambda_r} = n.$$

必要性. 设 $\boldsymbol{A}$ 可对角化, 则由定理3.1知

$$k_{\lambda_1} + k_{\lambda_2} + \cdots + k_{\lambda_r} = n.$$

从而

$$\begin{aligned} n &= k_{\lambda_1} + k_{\lambda_2} + \cdots + k_{\lambda_r} \\ &\leq m_{\lambda_1} + m_{\lambda_2} + \cdots + m_{\lambda_r} \\ &= n. \end{aligned}$$

取等时当且仅当对任意的 $1 \leq i \leq n$, 都有 $k_{\lambda_i} = m_{\lambda_i}$, 也即每一个代数重数和几何重数都相等.

充分性. 如果对任意的 $1 \leq i \leq n$, 都有 $k_{\lambda_i} = m_{\lambda_i}$, 那么必然有

$$k_{\lambda_1} + k_{\lambda_2} + \cdots + k_{\lambda_r} = m_{\lambda_1} + m_{\lambda_2} + \cdots + m_{\lambda_r} = n.$$

因此由定理3.1知 $\boldsymbol{A}$ 可对角化. □

另外，对于线性变换而言也有类似的结论.

推论 4.2. n 维线性空间 V 上的线性变换 $\mathscr{A}$ 可对角化的充要条件是对于 $\mathscr{A}$ 的每一个特征值, 其所对的代数重数和几何重数相等.

5 初等因子

接下来的这个定理, 是建立在 Jordan 标准型的理论的基础上的. 形如

$$\boldsymbol{J_i} = \begin{bmatrix} \lambda_i & 1 & 0 & \cdots & 0 \\ 0 & \lambda_i & 1 & \cdots & 0 \\ 0 & 0 & \lambda_i & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_i \end{bmatrix}_{n_i \times n_i}$$

的矩阵称为Jordan 块 (Jordan block), 而由若干个 Jordan 块形成的准对角矩阵

$$\boldsymbol{J} = \mathrm{diag}\{\boldsymbol{J_1}, \boldsymbol{J_2}, \cdots, \boldsymbol{J_n}\}$$

称为Jordan 形矩阵 (Jordan matrix). 所有的 n 阶方阵, 都可以被化为 Jordan 型矩阵. 同时, 根据上面的定义可以看出, 对角矩阵其实是一种特殊的 Jordan 标准型.

定理 5.1. *n 阶方阵 $\boldsymbol{A}$ 可对角化的充要条件是 $\boldsymbol{A}$ 的初等因子都是一次的.*

证明. 必要性. 设矩阵 $\boldsymbol{A}$ 可对角化, 则其的 Jordan 标准型矩阵为对角矩阵, 也即其的 Jordan 块都是一阶矩阵. 从而知矩阵 $\boldsymbol{A}$ 的初等因子都是一次的.

充分性. 设 $\boldsymbol{A}$ 的初等因子都是一次的, 设 $\boldsymbol{A}$ 的初等因子为

$$(\lambda - \lambda_1), (\lambda - \lambda_2), \cdots, (\lambda - \lambda_n),$$

其中 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 有可能重复, 则其对应的 n 个 Jordan 块都是一次的, 从而矩阵 $\boldsymbol{A}$ 的 Jordan 标准型为对角矩阵, 也即 $\boldsymbol{A}$ 可对角化. □

同样地, 对于线性变换 $\mathscr{A}$ 而言, 也有对应的 Jordan 标准型的理论, 这是因为可以找到一组基, 使得线性变换 $\mathscr{A}$ 在这组基下的矩阵为 Jordan 矩阵.

推论 5.2. *n 维线性空间 V 上的线性变换 $\mathscr{A}$ 可对角化的充要条件是 $\mathscr{A}$ 的初等因子都是一次的.*

最后, 我们将我们的目光移到矩阵 $\boldsymbol{A}$ 的最小多项式上. 在 λ 矩阵的理论当中, 我们得知, n 阶方阵 $\boldsymbol{A}$ 的最小多项式就是第 n 个不变因子, 从而我们可以得到如下推论.

推论 5.3. *n 阶方阵 $\boldsymbol{A}$ 可对角化的充要条件是 $\boldsymbol{A}$ 的最小多项式 $m(\lambda)$ 没有重根.*

事实上, 设设矩阵 $\boldsymbol{A}$ 的特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_r$, 则此时的最小多项式

$$m(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_r).$$

同样地, 考虑线性变换 $\mathscr{A}$, 类似的结论如下.

推论 5.4. *n 维线性空间 V 上的线性变换 $\mathscr{A}$ 可对角化的充要条件是 $\mathscr{A}$ 的最小多项式 $m(\lambda)$ 没有重根.*

至此, 我们给出了五个矩阵可对角化的充要条件. 根据线性变换和矩阵的对应关系, 我们也写出了在线性变换中对应的结论.

6 后记

矩阵的对角化, 或者更一般地, 研究矩阵的简化表示形式, 是具有重要的意义的. 在笔者的知识范围内, 可以举出以下的例子.

例 6.1. *在考虑齐次微分方程组 $\boldsymbol{x}' = \boldsymbol{A}\boldsymbol{x}$ 的解的时候, 作变换 $\boldsymbol{x} = \boldsymbol{P}\boldsymbol{y}$, 则有 $\boldsymbol{y}' = \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}\boldsymbol{y}$. 如果此时矩阵 $\boldsymbol{A}$ 可对角化, 或者可以简化表达, 那么此时的 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ 可以充分地简化, 方程组也更容易求解.*

例 6.2. *对于方阵 $\boldsymbol{A}$, 若需要求出 $\boldsymbol{A}^n$, 设 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{B}$, 则 $\boldsymbol{A} = \boldsymbol{P}\boldsymbol{B}\boldsymbol{P}^{-1}$, 从而*

$$\boldsymbol{A}^n = \underbrace{(\boldsymbol{P}\boldsymbol{B}\boldsymbol{P}^{-1})\cdot\cdots\cdot(\boldsymbol{P}\boldsymbol{B}\boldsymbol{P}^{-1})}_{n\text{个}} = \boldsymbol{P}\boldsymbol{B}^n\boldsymbol{P}^{-1}.$$

如果此时矩阵 $\boldsymbol{A}$ 可对角化, 或者可以简化表达, 那么此时的 $\boldsymbol{B}^n$ 可以更容易求解.

例 6.3. *将上面的例子作推广, 对于方阵 $\boldsymbol{A}$ 和解析函数 f, 若需要求出 $f(\boldsymbol{A})$, 设 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \mathrm{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_n\}$, 则可以证明*

$$f(\boldsymbol{A}) = \boldsymbol{P}\mathrm{diag}\{f(\lambda_1), f(\lambda_2), \cdots, f(\lambda_n)\}\boldsymbol{P}^{-1}.$$

本文只是研究了矩阵可对角化的一些充要条件, 同时也给出了将可对角化的矩阵化为对角矩阵的基本方法. 在接下来, 还需要寻求更好的变换方式 (例如正交变换), 以保

证矩阵在对角化的过程中不发现较大的改变; 同时, 也要对一般的矩阵, 找到其的最简表示形式 (例如 Jordan 标准型). 这些都是非常重要的.

参考文献

[1] 北京大学数学系前代数小组. 高等代数 *(第五版)*[M]. 北京: 高等教育出版社,2019.5.

[2] 戴华. 矩阵论 [M]. 北京: 科学出版社,2001.8.