

MIT 18.06 线性代数

第二十八讲：相似矩阵和若尔当型

整理自 MIT 18.06 课程及相关公开笔记

引言

本讲在上一讲正定矩阵的基础上，补充了关于正定矩阵的若干重要性质。随后引入 **相似矩阵** (Similar Matrices) 的概念，这是描述矩阵“家族”的重要工具。当矩阵无法对角化时，可以借助 **若尔当标准型** (Jordan Form) 获得一种近似的、结构规整的标准形式，从而继续分析矩阵的本质。

1 正定矩阵

1.1 逆矩阵的正定性

若矩阵 A 是正定矩阵，则它的逆矩阵 A^{-1} 也是正定矩阵。**理由**：正定矩阵 A 的全部特征值 λ_i 均大于 0，而逆矩阵的特征值为 $1/\lambda_i$ ，由于 $1/\lambda_i$ 同样是正数，因此 A^{-1} 也是正定矩阵。

1.2 正定矩阵的和

若矩阵 A 和 B 都是正定矩阵，则它们的和 $A + B$ 也是正定矩阵。**证明**：根据正定矩阵的定义，对于任意非零向量 $\mathbf{x}$ ，有 $\mathbf{x}^T A \mathbf{x} > 0$ 且 $\mathbf{x}^T B \mathbf{x} > 0$ 。因此：

$$\mathbf{x}^T (A + B) \mathbf{x} = \mathbf{x}^T A \mathbf{x} + \mathbf{x}^T B \mathbf{x} > 0$$

1.3 $A^T A$ 的正定性

设 A 是一个 $m \times n$ 的长方形矩阵。 A 通常不对称，但 $A^T A$ 总是对称方阵，其正定性可以通过二次型进行判断：

$$\mathbf{x}^T (A^T A) \mathbf{x} = (A \mathbf{x})^T (A \mathbf{x}) = \|A \mathbf{x}\|^2 \geq 0$$

此表达式等号成立，当且仅当 $A \mathbf{x} = \mathbf{0}$ 。若矩阵 A 是列满秩的（即各列线性无关），则只有零向量 $\mathbf{x} = \mathbf{0}$ 满足该条件，此时 $A^T A$ 是正定矩阵。

2 相似矩阵

2.1 相似矩阵的定义

设 A 和 B 是 n 阶方阵。若存在一个可逆矩阵 M ，使得： $B = M^{-1} A M$ ，则称 A 与 B 相似。**几何意义**：相似矩阵表示的是同一个线性变换在不同基下的矩阵描述。

2.2 相似矩阵的基本性质

1. 自反性: $A \sim A$ (取 $M = I$)。
2. 对称性: $A \sim B$ 则 $B \sim A$ (取 M^{-1})。
3. 传递性: $A \sim B$ 且 $B \sim C$ 则 $A \sim C$ 。

2.3 相似矩阵的必要条件

如果 A 与 B 相似, 则它们必然满足以下条件:

1. 特征多项式相同: $|\lambda I - A| = |\lambda I - B|$ 。
2. 特征值相同 (包含重数)。
3. 迹相同: $\sum_{i=1}^n a_{ii} = \sum_{i=1}^n b_{ii}$ 。
4. 行列式相同: $|A| = |B|$ 。
5. 秩相同: $r(A) = r(B)$ 。

2.4 相似矩阵的特征向量关系

相似矩阵的特征向量不同, 但存在明确转换关系。假设 $A\mathbf{x} = \lambda\mathbf{x}$, 由 $B = M^{-1}AM$ 可以推导:

$$\begin{aligned} AMM^{-1}\mathbf{x} &= \lambda\mathbf{x} \\ M^{-1}AM(M^{-1}\mathbf{x}) &= \lambda(M^{-1}\mathbf{x}) \\ B(M^{-1}\mathbf{x}) &= \lambda(M^{-1}\mathbf{x}) \end{aligned}$$

因此, 若 $\mathbf{x}$ 是 A 的特征向量, 则 $M^{-1}\mathbf{x}$ 是 B 的特征向量。

3 若尔当标准型

3.1 重特征值与不可对角化

当矩阵具有重特征值时, 可能无法找到足够的线性无关特征向量, 从而无法对角化。以二阶矩阵为例, 若特征值为重根 $\lambda_1 = \lambda_2 = 4$, 根据相似关系, 可以将所有可能的矩阵分为两类:

- 第一类: 仅包含矩阵 $\begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}$ 。因为对于任意可逆矩阵 M , 均有 $M^{-1} \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix} M = \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}$ 。
- 第二类: 包含所有其他具有特征值 4 的矩阵, 其中形式最简单且最具代表性的矩阵为 $\begin{bmatrix} 4 & 1 \\ 0 & 4 \end{bmatrix}$ 。

再如 $\begin{bmatrix} 5 & 1 \\ 3 & -1 \end{bmatrix}$ 、 $\begin{bmatrix} 4 & 0 \\ 17 & 4 \end{bmatrix}$, 这类矩阵都不能被对角化, 因为特征向量数量不足。

3.2 若尔当块与若尔当矩阵

若尔当块 (Jordan Block) J_i 是一个方阵，其主对角线上是同一个重特征值 λ_i ，主对角线正上方的元素为 1，其余位置均为 0，其形式如下：

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

若干若尔当块排列在对角线上，可以构成**若尔当矩阵** J ：

$$J = \begin{bmatrix} J_1 & 0 & \cdots & 0 \\ 0 & J_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_d \end{bmatrix}$$

3.3 若尔当定理与相似判定

任意 n 阶方阵 A 都相似于一个若尔当矩阵 J 。关于若尔当矩阵的结构，有以下重要结论：

- 若尔当块的个数 d 等于矩阵 A 线性无关特征向量的个数，每个块对应一个特征向量。
- 若矩阵 A 拥有 n 个互不相同的特征值，则它拥有 n 个线性无关特征向量 ($d = n$)，此时若尔当矩阵退化为对角矩阵 Λ (主对角线上方的元素均为 0)。
- 如果矩阵存在重特征值且特征向量缺失，则若尔当矩阵上方将出现 $n - d$ 个 1，这正是不可对角化矩阵的特征。
- 两个矩阵即使具有相同的特征值和相同数量的特征向量，若若尔当块的大小不同，则二者仍不是相似矩阵。

3.4 相似判定示例 (四重根情形)

考虑两个四阶矩阵，它们的特征值均为 0： $A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ ， $C = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ 。

这两个矩阵的秩均为 2，零空间维数为 2，因此都只有 2 个线性无关特征向量。然而， A 的若尔当块尺寸为 3 和 1，而 C 的若尔当块尺寸为 2 和 2，由于块大小不同， A 与 C 并不相似。

总结

- **正定矩阵的性质**：逆矩阵仍正定、和仍正定、列满秩下 $A^T A$ 正定、消元无需换行。
- **相似矩阵**：通过 $B = M^{-1} A M$ 联系，共享特征值、迹、行列式与秩。
- **若尔当标准型**：用若尔当块描述不可对角化矩阵的标准形式。若尔当块的个数等于线性无关特征向量的个数，块的大小决定了相似性的类别。