# 23-24 计算物理第一次习题课

## 线性代数知识补充

2023 年 10 月 17 日

# 目录

# 矩阵的本征值与本征矢量

## 本征值与本征矢量定义

对于方阵 $\mathbf{A}$，如果存在一个标量 $\lambda$ 和一个非零矢量 $\mathbf{v}$，使得以下关系成立：

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$$

其中，$\lambda$ 称为矩阵 $\mathbf{A}$ 的本征值，$\mathbf{v}$ 称为对应于本征值 $\lambda$ 的本征向量。

## 本征值的求解方法

要找到矩阵 $\mathbf{A}$ 的本征值和本征向量，我们需要解以下方程：

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$$

这是一个特征方程，可以通过解特征多项式来求解 $\lambda$，特征多项式为：

$$\det(\mathbf{A} - \lambda\mathbf{I}) = 0$$

其中，$\mathbf{I}$ 是单位矩阵。

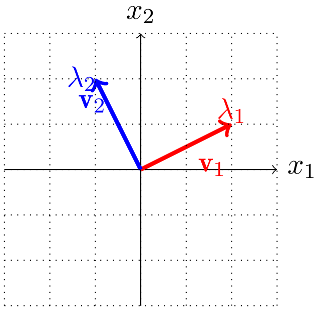
# 本征值与本征矢量

## 本征矢量的含义和求解方法

本征矢量是线性变换中的特殊方向，它们保持在变换后方向不变。本征矢量可以通过求解特征方程得到：

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$$

其中，$\lambda$ 是本征值，$\mathbf{v}$ 是对应的本征矢量。

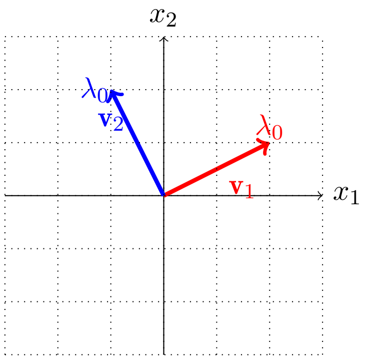



## 本征矢量张成本征空间

本征矢量 $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ 对应于矩阵 $\mathbf{A}$ 的不同本征值 $\lambda_1, \lambda_2, \ldots, \lambda_n$，它们张成了矩阵 $\mathbf{A}$ 的本征空间。

- 矩阵 $\mathbf{A}$ 的每个本征值都对应一个唯一的本征空间。

# 特征多项式与多重度




## 特征多项式

特征多项式为：

$$(\lambda - \lambda_0)^2 \times ... = 0$$

## 代数多重度与几何多重度

- **代数多重度：** 代表特征根在特征多项式中出现的次数。2。
- **几何多重度：** 几何多重度表示为 $\dim(\ker(\mathbf{A} - \lambda\mathbf{I}))$，它表示与特征值 $\lambda_0$ 相关联的本征矢量的数量

# 相似变换与本征值

## 相似变换的定义

对于两个 $n \times n$ 矩阵 $\mathbf{A}$ 和 $\mathbf{B}$，如果存在一个可逆矩阵 $\mathbf{P}$，使得以下关系成立：

$$\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$$

则称矩阵 $\mathbf{A}$ 与 $\mathbf{B}$ 相似，矩阵 $\mathbf{P}$ 称为相似变换矩阵。

## 相似变换保持本征值

证明：假设 $\mathbf{A}$ 和 $\mathbf{B}$ 相似，即 $\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$，其中 $\mathbf{P}$ 是可逆矩阵。
假设 $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$， $\mathbf{w} = \mathbf{P}^{-1}\mathbf{v}$。
那么 $\mathbf{B}\mathbf{w} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}\mathbf{w} = \mathbf{P}^{-1}\lambda\mathbf{v}$。
因此 $\mathbf{B}\mathbf{w} = \lambda\mathbf{w}$

## 幺正变换

可逆矩阵 $\mathbf{P}$ 为幺正矩阵 $\mathbf{U}$ 满足 $\mathbf{U}^{\dagger}\mathbf{U} = \mathbf{U}\mathbf{U}^{\dagger} = \mathbf{I}$

# Schur 分解

- 相似变换可以通过将矩阵对角化或者准对角化来求解矩阵的特征值。

**Schur 分解**

对于任意一个方阵 $\mathbf{A}$，存在一个幺正矩阵（unitary matrix）$\mathbf{U}$，使得 $\mathbf{U}^{\dagger}\mathbf{A}\mathbf{U}$ 是一个上三角矩阵（upper triangular matrix）。
换句话说，$\mathbf{A}$ 可以被相似变换成一个上三角矩阵。
具体来说，上三角矩阵的对角线上的元素是 $\mathbf{A}$ 的特征值，而非对角线上的元素可能为零。

- Schur 分解定理提供了一种将任意方阵转化为上三角形式的方法，从而更容易求解矩阵的特征值。

# 正规矩阵和厄米矩阵的对角化

## 正规矩阵的对角化

正规矩阵 $AA^* = A^*A$ 可以通过幺正矩阵 $U$ 的相似变换对角化，即：

$$A = UDU^*$$

特征根为复数

## 厄米矩阵的对角化

**厄米矩阵** $H = H^*$

- 厄米矩阵总是可以进行**实对角化**，即可以通过实正交矩阵对角化。即：

$$A = Q\Sigma Q^*$$

# Jordan 标准型

## 对角化的条件

当代数多重度（一个特征值的重复次数）大于几何多重度（与特征值相关联的线性独立的本征矢量的数量）时，矩阵无法实现对角化，但可以进行块对角化。

## Jordan 块示例

$$\begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix}$$

这是一个 $3 \times 3$ 的 Jordan 块，$1$ 是块的上三角部分。特征值 Jordan 块的数量是几何多重度。

## Jordan 标准型

Jordan 标准型是一种矩阵的特殊形式，用于块对角化。它由多个 Jordan 块组成，每个块对应一个特征值。

# Jordan 标准型

## Jordan 标准型

任意 $n \times n$ 的复矩阵 $\mathbf{A}$ 可以通过相似变换转化为 Jordan 标准型：

$$\mathbf{X}^{-1}\mathbf{A}\mathbf{X} = \mathbf{J} = \mathrm{Diag}(\mathbf{J}_1, \mathbf{J}_2, \ldots, \mathbf{J}_n)$$

其中，$\mathbf{X}$ 是可逆矩阵，$\mathbf{J}_i$ 是第 $i$ 个 Jordan 块。

## Jordan 标准型的稳定性

需要注意的是，Jordan 标准型并不是稳定的。这意味着即使微小的扰动也可能改变矩阵的 Jordan 标准型。

## Jordan 块的数量和对角化条件

Jordan 块的数量等于线性独立的本征矢的数量。矩阵可以对角化的充要条件是每个 Jordan 块都是 $1 \times 1$ 的，即只包含一个特征值。

# 本征值分布与 Gershgorin 圆盘定理

## Gershgorin 圆盘定理

对于任意 $n \times n$ 矩阵 $\mathbf{A}$ 的每个本征值 $\lambda_i$，存在至少一个 Gershgorin 圆盘，其中心是矩阵的对角元素 $a_{ii}$，半径是其他同一行元素的绝对值之和：

$$|\lambda_i - a_{ii}| \leq \sum_{j \neq i} |a_{ij}|$$

## Gershgorin 圆盘定理的应用

对于厄米矩阵（Hermitian matrix），Gershgorin 圆盘定理为求解本征值提供了有用的信息。在搜索本征值时，可以使用这些圆盘的上下限来缩小搜索范围，从而更有效地找到本征值。

# 奇异值分解定理

## 奇异值分解定理（Singular Value Decomposition，SVD）

对于任意复数域上的 $m \times n$ 矩阵 $\mathbf{A}$，存在以下奇异值分解：

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^*$$

其中，$\mathbf{U}$ 和 $\mathbf{V}$ 分别是复数域上的幺正矩阵（unitary matrices），$\mathbf{\Sigma}$ 是一个大小为 $m \times n$ 的对角矩阵，其对角线上的非负元素称为奇异值 (singular values)，通常按降序排列。

## 奇异值分解的应用

奇异值分解是矩阵分析和线性代数中的重要工具，广泛应用于数据降维、矩阵逼近、图像压缩、信号处理、推荐系统等领域。它能够揭示数据的主要特征、去除噪声、简化计算，并为复杂问题提供了有效的数值解法。

# 手算奇异值分解示例

假设我们有一个矩阵 $\mathbf{A}$：

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$$

# 手算奇异值分解示例

假设我们有一个矩阵 $\mathbf{A}$：

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$$

**步骤 1：计算 $\mathbf{A}\mathbf{A}^T$**

$$\mathbf{A}\mathbf{A}^T = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 5 & 8 \\ 8 & 13 \end{bmatrix}$$

# 手算奇异值分解示例

假设我们有一个矩阵 $\mathbf{A}$：

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$$

**步骤 1：计算 $\mathbf{A}\mathbf{A}^T$**

$$\mathbf{A}\mathbf{A}^T = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 5 & 8 \\ 8 & 13 \end{bmatrix}$$

**步骤 2：计算 $\mathbf{A}^T\mathbf{A}$**

$$\mathbf{A}^T\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 5 & 8 \\ 8 & 13 \end{bmatrix}$$

# 手算奇异值分解示例 (续)

**步骤 3：计算 $\mathbf{AA}^T$ 的特征值和特征向量**

特征方程：$\det(\mathbf{AA}^T - \lambda\mathbf{I}) = 0$

计算得到特征值：$\lambda_1 = 18$，$\lambda_2 = 0$

对于 $\lambda_1 = 18$，求解特征向量：

$$\begin{bmatrix} 5-18 & 8 \\ 8 & 13-18 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -13 \\ -5 \end{bmatrix}$$

解得特征向量 $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

对于 $\lambda_2 = 0$，求解特征向量：

$$\begin{bmatrix} 5 & 8 \\ 8 & 13 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

解得特征向量 $\mathbf{v}_2 = \begin{bmatrix} -8 \\ 5 \end{bmatrix}$

# 手算奇异值分解示例 (续)

**步骤 4：计算 $\mathbf{A}^T\mathbf{A}$ 的特征值和特征向量**

特征方程：$\det(\mathbf{A}^T\mathbf{A} - \lambda\mathbf{I}) = 0$

计算得到特征值：$\lambda_1 = 18$，$\lambda_2 = 0$

对于 $\lambda_1 = 18$，求解特征向量：

$$\begin{bmatrix} 5-18 & 8 \\ 8 & 13-18 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -13 \\ -5 \end{bmatrix}$$

解得特征向量 $\mathbf{u}_1 = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$

对于 $\lambda_2 = 0$，求解特征向量：

$$\begin{bmatrix} 5 & 8 \\ 8 & 13 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

解得特征向量 $\mathbf{u}_2 = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$

# 手算奇异值分解示例 (续)

**步骤 5: 构建奇异值分解**

我们已经得到特征值和特征向量:

$\lambda_1 = 18$, $\mathbf{u}_1 = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$, $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

$\lambda_2 = 0$, $\mathbf{u}_2 = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} -8 \\ 5 \end{bmatrix}$

奇异值 $\sigma_i$ 是特征值的平方根，按降序排列:

$\sigma_1 = \sqrt{\lambda_1} = \sqrt{18}$, $\sigma_2 = \sqrt{\lambda_2} = 0$

构建奇异值矩阵 $\Sigma$ 和左右奇异向量矩阵 $\mathbf{U}$、$\mathbf{V}$:

$$\Sigma = \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix} = \begin{bmatrix} \sqrt{18} & 0 \\ 0 & 0 \end{bmatrix}$$

$$\mathbf{U} = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 \end{bmatrix} = \begin{bmatrix} -2 & 2 \\ 1 & -1 \end{bmatrix}$$

$$\mathbf{V} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{bmatrix} = \begin{bmatrix} 1 & -8 \\ 2 & 5 \end{bmatrix}$$

# 手算奇异值分解示例 (续)

**步骤 6：计算 $\mathbf{A}$ 的奇异值分解**
现在我们可以计算 $\mathbf{A}$ 的奇异值分解了：

$$\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}^T = \begin{bmatrix} -2 & 2 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} \sqrt{18} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ -8 & 5 \end{bmatrix}^T$$

计算得到：

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$$

这就是矩阵 $\mathbf{A}$ 的奇异值分解结果。

# 矩阵模

## 数值求解线性方程组

在计算物理中，我们经常需要求解线性方程组，通常采用数值方法。其中，一个常见的任务是解决形如 $\min \|\mathbf{A}\mathbf{x} - \mathbf{b}\|$ 的问题，其中 $\mathbf{A}$ 是系数矩阵，$\mathbf{x}$ 是未知变量，$\mathbf{b}$ 是右侧向量。

## 诱导矩阵模

对于一个矩阵 $\mathbf{A}$，其模被定义为：

$$\|\mathbf{A}\| = \sup_{\|\mathbf{x}\| \neq 0} \frac{\|\mathbf{A}\mathbf{x}\|}{\|\mathbf{x}\|}$$

其中，$\|\mathbf{x}\|$ 表示向量 $\mathbf{x}$ 的模，$\sup$ 表示上确界。

- 非负性：$\|\mathbf{A}\|$ 总是非负的。
- 均匀性：$\|\alpha\mathbf{A}\| = |\alpha|\|\mathbf{A}\|$ 对于所有标量 $\alpha$ 成立。
- 三角不等式：$\|\mathbf{A}+\mathbf{B}\| \leq \|\mathbf{A}\| + \|\mathbf{B}\|$ 对于所有矩阵 $\mathbf{A}$ 和 $\mathbf{B}$ 成立。

# 矢量 $p$ 模和矩阵 $p$ 模

## 矢量 $p$ 模

给定一个 $n$ 维实数矢量 $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ 和一个正整数 $p$，矢量 $p$ 模定义如下：

$$\|\mathbf{x}\|_p = \left( \sum_{i=1}^{n} |x_i|^p \right)^{1/p}$$

这里，$|x_i|$ 表示 $x_i$ 的绝对值。

## 诱导的矩阵 $p$ 模

对于一个 $n \times m$ 实数矩阵 $\mathbf{A}$，其诱导的矩阵 $p$ 模定义如下：

$$\|\mathbf{A}\|_p = \sup_{\|\mathbf{x}\|_p \neq 0} \frac{\|\mathbf{A}\mathbf{x}\|_p}{\|\mathbf{x}\|_p}$$

其中，$\|\mathbf{x}\|_p$ 表示矢量 $\mathbf{x}$ 的 $p$ 模。

# $p = 1$ 情况

## 矩阵 1 模

矩阵 1 模表示为:

$$\|\mathbf{A}\|_1 = \max_{\|\mathbf{x}\|_1=1} \|a_j\|_1$$

## 推导过程

$$\|\mathbf{A}\|_1 = \|\Sigma_j x_j a_j\|_1 \leq \Sigma_j |x_j| \|a_j\|_1$$

因为 $\Sigma |x_i| = 1$, 因此

$$\Sigma_j |x_j| \|a_j\|_1 \leq \max_{\|\mathbf{x}\|_1=1} \|a_j\|_1$$

- 矩阵 1 模是矩阵列和的最大值

# $p = 2$ 情况

## 矩阵 2 模

矩阵 2 模表示为：

$$\|\mathbf{B}\|_2 = \sqrt{\lambda_{\mathsf{max}}(\mathbf{B}^\top \mathbf{B})}$$

## 推导过程

$$\frac{\|\mathbf{B}\vec{x}\|_2}{\|\vec{x}\|_2} = \frac{(\vec{x}^\top \mathbf{B}^\top \mathbf{B}\vec{x})^{\frac{1}{2}}}{(\vec{x}^\top \vec{x})^{\frac{1}{2}}} = \left(\frac{\Sigma_{i=0}^n \|c_i\|^2 v_i}{\Sigma_{j=0}^n \|c_j\|^2}\right)^{\frac{1}{2}}$$

其中，$\vec{x} = \Sigma_{j=1}^n c_i \vec{\alpha}_i$，$\vec{\alpha}_i$ 是 $\mathbf{B}^\top \mathbf{B}$ 的本征矢

$$\frac{\|\mathbf{B}\vec{x}\|_2}{\|\vec{x}\|_2} \leq \sqrt{v_n} = \sqrt{\lambda_{\mathsf{max}}(\mathbf{B}^\top \mathbf{B})}$$

若 $\mathbf{B}$ 是厄米矩阵 $\|\mathbf{B}\|_2^2 = \lambda_{max}(B)$