

线性代数期末复习

李谨硕

2025 年 12 月 31 日

目录

1 基本概念和记号	3
1.1 数域	3
1.2 线性方程组	3
1.3 矩阵	3
2 矩阵的运算, 变换和性质	5
2.1 数乘, 加法, 乘法	5
2.2 矩阵的变换	6
2.3 可逆矩阵 (逆矩阵)	10
2.3.1 定义与判别	10
2.3.2 逆矩阵的性质	10
2.3.3 3. 伴随矩阵 (Adjugate Matrix)	11
2.3.4 4. 逆矩阵的求法	11
3 向量与线性方程组解的结构	13
3.1 向量 (组) 的概念和定义	13
3.2 齐次方程组解的结构	15
3.3 非齐次方程组解的结构	17
4 线性空间与线性变换	18
4.1 线性空间的基本概念	18
4.2 线性变换的基本概念	20
4.3 欧氏空间	21
4.3.1 正交基与正交化	23
4.4 线性变换	25
4.5 相似对角化	28
4.5.1 实对称矩阵的相似对角化	34

5 实二次型	36
5.1 定义和性质	36
5.2 惯性定理和正定二次型	38
6 全书总结	40
6.1 线性公理	40
6.2 秩-零化度定理的多种表示形式	41
6.3 秩的多种含义	41

1 基本概念和记号

1.1 数域

定义 1.1: 数域

1. 设 F 是一个数集, 如果 F 中任意两个数作某种运算之后的结果仍然属于数域 F , 称数集 F 对这种运算封闭。
2. 设 F 是包括 0, 1 的数集, 若 F 对四则运算封闭, 则称 F 是一个数域。 $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ 都是常见的数域。我们一般在 \mathbb{R} 上讨论。

1.2 线性方程组

定义 1.2: 线性方程组

我们把形如下面这种表达式的方程 (组) 称为线性方程 (组):

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_n \end{cases}$$

对于一个线性方程组, 倘若存在一组数 $c_1, c_2, c_3, \dots, c_n \in F$ 满足上述所有方程, 我们称这组数为方程的一组解。所有解的集合称为解集。

如果两个方程组的解集相同, 我们称这两组方程为同解方程组。

1.3 矩阵

定义 1.3: 矩阵

由 $m \times n$ 个数 a_{ij} 排列成的数表称为矩阵, 形如:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

对于两个矩阵 A, B , 若他们的行数和列数相同, 则称 A 和 B 同型。若对两个同型矩阵, 对应位置上的每个元素的值都相同, 那我们称两个矩阵相等。

性质 1.1: 常见矩阵类型

- 零矩阵: 所有元素都是 0。
- 行矩阵/列矩阵: 也称为行/列向量, 仅有一行和一列的矩阵。
- 方阵: 行数等于列数的矩阵。
在方阵中, 从 $(1, 1)$ 到 (n, n) 位置的线称为主对角线, 从 $(1, n)$ 到 $(n, 1)$ 位置的线称为副对角线。
- 对角矩阵: 除了对角线上的元素都是 0 的方阵。
- 单位矩阵: 对角线元素都是 1 的对角矩阵, 简记为 E_n 和 I_n 。
- 数量矩阵: 主对角元都相等且非零的对角阵, 简记为 kE_n 。
- 三角矩阵: 上下三角矩阵分别对应对角线以下和以上元素都是 0。
- 对称矩阵: 若 $a_{ij} = a_{ji}$, 则为对称矩阵。
- 反对称矩阵: 若 $a_{i,j} = -a_{j,i}$, 则为反对称矩阵。

对于一个线性方程, 把系数写成数表的形式, 得到系数矩阵, 加入右侧常数项, 我们会得到增广矩阵。线性方程组和矩阵存在一一对应的关系。同理的, 对线性方程组的操作本质上也是对矩阵的操作。更深刻的理解是: 对未知数 x_1, x_2, \dots, x_n 的线性方程组本质上是一种变换, 被称作线性变换。而线性代数的核心目标就是在探究这种线性变换的各种性质。

2 矩阵的运算, 变换和性质

2.1 数乘, 加法, 乘法

定义 2.1: 矩阵的基本运算

1. **数乘:** 矩阵的数乘本质上就是对矩阵的每个元素的数乘。用 k 数乘 A 记作 kA , 特别的, 当 $k = -1$ 时, 得到矩阵的逆矩阵 $-A$ 。
2. **加法:** 矩阵的加法只能对 **同型矩阵** 进行, 对于不同型矩阵不存在加法运算。我们规定矩阵加法为对两个同型矩阵相同位置的元素进行加法运算。特别的, 我们定义同型矩阵的减法运算为 $A - B = A + (-B)$ 。
数乘和加法运算统称矩阵的线性运算。矩阵的两种线性运算具有类似数的所有运算律, 比如交换律, 结合律, 分配律。我们非常看重这种线性运算所带来的一系列性质。
3. **乘法:** 对于矩阵 A 和 B , 我们在计算其乘法时, 要求 A 的列数和 B 的行数相同, 然后才能进行矩阵的运算。得到的新矩阵继承了左边矩阵的行数和右边矩阵的列数。用形式化语言来说, 就是: 对于 $A_{m \times n}$ 和 $B_{n \times s}$, $A \times B = C_{m \times s}$ 更具体的, 对于 C 中的每个元素, 我们有:

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} = \sum_{k=1}^n a_{ik}b_{kj}$$

切记, 在矩阵和向量中, 我们并没有定义除法。即使是未来提及的逆矩阵也不能被认为是乘法的逆运算。

4. **幂运算:** $A^k = A \times A \times \cdots \times A$ 。

注意 (Note)

一定记住, 矩阵乘法的交换律对于随意的两个矩阵并不总是成立, 也就是说, 矩阵的乘法不具有交换律。如果我们发现一个矩阵乘以另一个矩阵总是等于他们交换位置后的乘法运算, 那么我们可以推知它一定是数量矩阵。

乘法满足结合律, 左交换律和右交换律。

性质 2.1: 幂运算的注意事项

对于同底数的幂运算, 矩阵的行为和数的行为是一致的。但是除此之外就必须考虑交换律的问题, 比如下面两个式子 (A, B 为方阵):

$$(A + B)^2 = A^2 + AB + BA + B^2 \neq A^2 + 2AB + B^2$$

$$(A + B)(A - B) = A^2 - AB + BA - B^2 \neq A^2 - B^2$$

2.2 矩阵的变换

定义 2.2: 转置

转置: 设 A , 那么 A^T 为:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \quad A^T = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{pmatrix}$$

性质 2.2: 矩阵转置的性质

- $(A^T)^T = A$
- $(B + C)^T = B^T + C^T$
- $(kA)^T = kA^T$
- $(AB)^T = B^T A^T$ (重点关注)
- $(A^m)^T = (A^T)^m$, A 为方阵

矩阵的性质

矩阵具有某些特殊的性质和量, 他们在特殊的一类变换下保持不变。

定义 2.3: 方阵的迹

设方阵 $A = (a_{ij})_{n \times n}$, 称 $\sum_{i=1}^n a_{ii}$ 为方阵的迹, 记为 $\text{tr}(A)$ 。矩阵的迹主要和特征值有关。

方阵的行列式: 矩阵的行列式具有多种定义方式, 他们是等价的。

定义 2.4: 行列式

设 $A = (a_{ij})$ 是一个 $n \times n$ 阶方阵, 其行列式定义为:

$$\det(A) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^n a_{i,\sigma(i)}$$

其中:

- S_n 是 $\{1, 2, \dots, n\}$ 上所有排列的集合 (共有 $n!$ 个元素)。
- σ 是 S_n 中的一个排列, $\sigma(i)$ 表示排列中第 i 个位置的元素。
- $\operatorname{sgn}(\sigma)$ 是排列 σ 的符号: 若 σ 是偶排列 (逆序数为偶数), 取值为 $+1$; 若为奇排列 (逆序数为奇数), 取值为 -1 。
- 乘积 $\prod_{i=1}^n a_{i,\sigma(i)}$ 表示从矩阵 A 的每一行取一个元素, 且这些元素位于不同列 (即取自不同行不同列的 n 个元素的乘积)。

行列式在解决方程的解的问题上具有相当重要的意义, 我们后面会进行介绍。

计算行列式的方式是多种多样的, 一种被称作 Laplace 展开:

$$\det(A) = \sum_{j=1}^n a_{ij} \cdot C_{ij} = \sum_{j=1}^n (-1)^{i+j} a_{ij} \cdot \det(M_{ij})$$

当然也可以按列展开。这里略而不谈。

另一种方式是通过初等变换法解决。

性质 2.3: 行列式初等变换规则

1. 交换矩阵 A 的 i, j 两行, 行列式反号
2. 矩阵 A 的某行乘以非零常数 k , 行列式乘 k
3. 把矩阵 A 的 i 行乘以 k 倍加到 j 行上去, 行列式不变

利用这些初等变换方法, 我们可以把行列式化为上三角矩阵, 然后行列式即等于对角线元素的乘积。这种方法在计算机中非常实用。它的时间复杂度为 $O(n^3)$ 。但是传统的 Laplace 展开法时间复杂度为 $O(n!)$, 因此在实际计算中并不实用。

行列式在计算中最重要的性质是:

$$\det(AB) = \det(A) \cdot \det(B)$$

我们还可以定义子式: 对于一个矩阵, 任取 k 行和 k 列, 他们的交叉元素组成的行列式被称为 k 阶子式。若这 k 行和 k 列的标号相同, 那么我们得到的 k 阶子式被称为 k 阶主子式。

定义 2.5: 余子式与代数余子式

- **余子式 Minor:** 在 n 阶方阵 A 中, 元素 a_{ij} 的余子式 M_{ij} 是指划去第 i 行和第 j 列后, 剩余的 $n-1$ 阶方阵的行列式。
- **代数余子式 Cofactor:** 元素 a_{ij} 的代数余子式 C_{ij} 是其余子式 M_{ij} 与符号系数 $(-1)^{i+j}$ 的乘积, 即 $C_{ij} = (-1)^{i+j} M_{ij}$ 。该符号系数由元素位置决定。

性质 2.4: 行列式的性质

- $|A| = |A^T|$
- 若行列式两行相等或者对应成比例, 则行列式为 0。
- 若行列式的每一行都是两个元素的和, 那么此行列式可以拆分成两个的和。

定义 2.6: 矩阵的秩

对于一个矩阵, 其对应的线性方程的“有效方程个数”被称作矩阵的秩。它的严格定义是:

矩阵的秩 r , 存在一个 r 阶子式不为零, 但是所有 $r+1$ 阶子式都为零, 那么 r 就被称为是矩阵的秩。

矩阵的秩在线性代数全书中具有非常鲜明的几何意义, 我们会在后面发现这一点。它可以是空间的维度, 是线性变换的维度等一众含义。

矩阵的初等变换

定义 2.7: 初等变换

对矩阵进行如下操作被称作矩阵的初等行变换。

1. 交换矩阵 A 的 i, j 两行
2. 矩阵 A 的某行乘以非零常数 k 。
3. 把矩阵 A 的 i 行乘以 k 倍加到 j 行上去。

对矩阵的列进行相同操作会被称作矩阵的初等列变换。二者合称矩阵的初等变换。

定理 2.1: 初等矩阵定理

设 A 是一个 $m \times n$ 的矩阵, 对 A 进行一次初等行变换, 相当于左乘一个经过同样初等行变换的单位矩阵 E_m 。而对 A 进行一次初等列变换, 相当于右乘一个经过同样初等列变换的单位矩阵 E_n 。这种经过一次初等变换得到的矩阵被称作 **初等矩阵**。

矩阵和线性方程组

我们在用加减消元法解方程的过程本质上就是在用矩阵做初等行变换。

- **高斯消元法 (Gaussian Elimination)**

这是求解线性方程组最通用、最基础的方法。其核心思想是通过 **初等行变换** 将增广矩阵 (A, b) 化为 **行阶梯形矩阵** 或 **行最简形矩阵**。

定理 2.2: 秩的判别准则

通过行阶梯形矩阵, 我们可以直观地判断方程组解的情况:

- 若 $r(A) < r(A, b)$: 方程组无解 (出现了 $0 = c, c \neq 0$ 的矛盾方程)。
- 若 $r(A) = r(A, b) = n$ (未知数个数): 方程组有唯一解。
- 若 $r(A) = r(A, b) < n$: 方程组有无穷多解, 此时即使化为行最简形, 也需要选定 $n - r(A)$ 个自由未知量。

- **克拉默法则 (Cramer's Rule)**

适用于 **方程个数等于未知数个数** (即系数矩阵 A 为方阵) 且 **系数行列式不为零** ($|A| \neq 0$) 的情形。

定理 2.3: 克拉默法则

如果线性方程组的系数矩阵 A 的行列式 $D = |A| \neq 0$, 那么方程组有唯一解, 且解为:

$$x_1 = \frac{D_1}{D}, \quad x_2 = \frac{D_2}{D}, \quad \dots, \quad x_n = \frac{D_n}{D}$$

其中 D_j 是把系数行列式 D 中第 j 列的元素用常数项 b_1, b_2, \dots, b_n 替换后得到的行列式。

注意 (Note)

注: 克拉默法则主要用于理论证明和某些特定性质的推导, 在实际计算高阶方程组时, 由于计算量过大 (需要计算 $n + 1$ 个行列式, 计算机计算时间复杂度为 $O(n!)$), 通常不如高斯消元法高效。

2.3 可逆矩阵 (逆矩阵)

矩阵的逆是线性代数中类比实数倒数的一个概念, 但它仅针对方阵定义, 且具有更复杂的性质。

2.3.1 定义与判别

定义 2.8: 可逆矩阵

- **定义:** 对于 n 阶方阵 A , 如果存在 n 阶方阵 B , 使得 $AB = BA = E$ (其中 E 为单位矩阵), 则称 A 是 **可逆的** (Invertible) 或 **非奇异的** (Non-singular), 并称 B 为 A 的 **逆矩阵**, 记作 A^{-1} 。
- **唯一性:** 若矩阵 A 可逆, 则其逆矩阵是唯一的。
- **奇异矩阵:** 若方阵 A 的行列式 $|A| = 0$, 则称 A 为 **奇异矩阵** (Singular Matrix), 此时 A 不可逆。

2.3.2 逆矩阵的性质

性质 2.5: 逆矩阵的性质

设 A, B 均为 n 阶可逆方阵, $\lambda \neq 0$ 为常数:

1. $(A^{-1})^{-1} = A$
2. $(\lambda A)^{-1} = \frac{1}{\lambda}A^{-1}$
3. **反序定律:** $(AB)^{-1} = B^{-1}A^{-1}$ (这一点非常重要, 类似于转置的性质)
4. $(A^T)^{-1} = (A^{-1})^T$ (转置和求逆可交换)
5. $|A^{-1}| = \frac{1}{|A|} = |A|^{-1}$ (行列式的性质)

2.3.3 3. 伴随矩阵 (Adjugate Matrix)

为了给出逆矩阵的公式解, 我们需要引入伴随矩阵的概念。

定义 2.9: 伴随矩阵

- 定义: 行列式 $|A|$ 的各个元素的 代数余子式 A_{ij} 所构成的矩阵的 转置矩阵 称为 A 的伴随矩阵, 记为 A^* 。

$$A^* = \begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix}$$

注意: 写伴随矩阵时, 切记要将代数余子式 A_{ij} 放在第 j 行第 i 列的位置 (即下标转置)。

- 重要性质:

$$AA^* = A^*A = |A|E$$

这个公式联系了原矩阵、伴随矩阵和行列式, 是推导逆矩阵公式的关键。

2.3.4 4. 逆矩阵的求法

求逆矩阵主要有两种方法, 分别适用于不同的场景。

方法一: 伴随矩阵法 (公式法) 利用公式:

$$A^{-1} = \frac{1}{|A|} A^*$$

- 适用场景: 二阶、三阶矩阵, 或者理论推导。对于高阶矩阵, 计算所有代数余子式 极其繁琐。

方法二: 初等变换法 (高斯-约当消元法) 利用分块矩阵的初等行变换。

算法步骤 2.1: 高斯-约当消元法求逆

- 原理: 对 (A, E) 施行初等行变换。当 A 被化简为单位矩阵 E 时, 右边的 E 就同步变成了 A^{-1} 。

$$(A \mid E) \xrightarrow{\text{初等行变换}} (E \mid A^{-1})$$

- 适用场景: 具体的数值计算, 特别是三阶及以上的矩阵。这也是计算机求解逆矩阵的常用算法思路。

定理 2.4: 核心概念大一统

线性代数中美妙的地方在于同一个性质可以从不同角度描述。对于 n 阶方阵 A , 以下命题是 等价的 (即要么同时成立, 要么同时不成立):

1. A 是可逆矩阵 (A^{-1} 存在)。
2. A 的行列式不为零 ($|A| \neq 0$, 即 A 非奇异)。
3. A 的秩等于 n ($r(A) = n$, 即满秩)。
4. 齐次线性方程组 $Ax = 0$ 只有零解。
5. 非齐次线性方程组 $Ax = b$ 对任意 b 有唯一解。
6. A 可以通过初等行变换化为单位矩阵 E 。
7. A 的行向量组 (或列向量组) 线性无关。
8. 0 不是 A 的特征值。

理解这些等价关系, 是掌握线性代数“概念联系”的关键。

3 向量与线性方程组解的结构

3.1 向量(组)的概念和定义

定义 3.1: 向量与向量组

1. 向量的定义: 在 n 维空间中, 我们把有向线段称为向量。向量具有大小和方向两个属性。我们用 \vec{a} 表示向量 a 。
2. 向量的表示: 在 n 维空间中, 我们可以用有序数组 (a_1, a_2, \dots, a_n) 来表示向量 \vec{a} 。这种表示方式被称作向量的坐标表示法。另外, 我们也可以用列矩阵的形式来表示向量:

$$\vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$$

3. 向量的线性运算: 向量的线性运算包括向量的加法和数乘, 其定义和矩阵的线性运算类似。
4. 向量的线性组合: 设 $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n$ 是 n 个向量, k_1, k_2, \dots, k_n 是 n 个数, 那么向量 $\vec{b} = k_1\vec{a}_1 + k_2\vec{a}_2 + \dots + k_n\vec{a}_n$ 被称作向量 $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n$ 的线性组合。
5. 向量的线性相关性: 设 $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n$ 是 n 个向量, 如果存在不全为零的数 k_1, k_2, \dots, k_n , 使得 $k_1\vec{a}_1 + k_2\vec{a}_2 + \dots + k_n\vec{a}_n = \vec{0}$, 那么我们称向量 $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n$ 线性相关。否则, 我们称他们线性无关。

运用向量的记法, 我们可以把线性方程组写成矩阵的形式:

$$Ax = b$$

线性方程组 $Ax = b$ 有解的充分必要条件是 b 在 A 的列空间中。用秩的语言来说, 就是 $r(A) = r(A, b)$ 。

6. 向量组: 设 $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n$ 是 n 个向量, 那么我们把这 n 个向量称作一个向量组, 记作 $\{\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n\}$ 。
7. 向量组的等价性: 设 $\{\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n\}$ 和 $\{\vec{b}_1, \vec{b}_2, \dots, \vec{b}_m\}$ 是两个向量组, 如果 $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n$ 都可以被表示成 $\vec{b}_1, \vec{b}_2, \dots, \vec{b}_m$ 的线性组合, 且 $\vec{b}_1, \vec{b}_2, \dots, \vec{b}_m$ 也可以被表示成 $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n$ 的线性组合, 那么我们称这两个向量组是等价的。

定理 3.1: 向量组等价定理

n 维向量组 $\{\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n\}$ 和 m 维向量组 $\{\vec{b}_1, \vec{b}_2, \dots, \vec{b}_m\}$ 等价的充分必要条件是 $r(A) = r(B) = r(A, B)$, 其中 A 和 B 分别是这两个向量组的矩阵表示, 而 $C = (A, B)$ 是把这两个矩阵拼接起来的矩阵表示。

定理 3.2: 线性相关判定

n 维向量组 a_1, a_2, \dots, a_n 线性相关的充分必要条件是齐次方程组: $(a_1, a_2, \dots, a_n)x = 0$ 有非零解。

推论 3.1

m 个 n 维向量线性相关的充分必要条件是 $r(A) < m$, 其中 A 是由这 m 个向量组成的矩阵, 而线性无关的充分必要条件是 $r(A) = m$ 。

推论 3.2

n 个 n 维向量线性相关的充要条件是 $\det(A) = 0$, 其中 A 是由这 n 个向量组成的矩阵。反之也成立。

推论 3.3

m 个 n 维向量线性相关的充分必要条件是 $m \geq n$ 。

定义 3.2: 极大无关组

极大线性无关组: 设向量组 $\{\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n, \dots\}$, 如果这个向量组中存在一个子组 $\{\vec{a}_1, \vec{a}_2, \dots, \vec{a}_k\}$, 他们线性无关, 且任意添加原来的组中的一个向量后这个向量组就线性相关, 那么我们称这个向量组是极大线性无关组。

定理 3.3: 极大无关组

一个向量组与其每一个极大线性无关组等价。每一个向量组的极大线性无关组的向量个数都相同, 这个数被称作这个向量组的秩。每一个极大线性无关组之间都相互等价。

定理 3.4

向量组 $\{\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n\}$ 线性无关的充分必要条件是其向量组的秩等于 n 。即其极大无关组是它自己。

3.2 齐次方程组解的结构

定义 3.3: 齐次方程组

1. 齐次方程组: 形如 $Ax = 0$ 的线性方程组被称作齐次线性方程组。
2. 齐次方程组的解: 齐次线性方程组 $Ax = 0$ 的解被称作齐次线性方程组的解。显然, 齐次线性方程组至少有一个解, 即零解 $x = 0$ 。

性质 3.1: 解的性质

设齐次方程组 $Ax = 0, \eta_1, \eta_2, \dots, \eta_s$ 是这个方程组的解, 那么 $k_1\eta_1 + k_2\eta_2 + \dots + k_s\eta_s$ 也是这个方程组的解, 其中 k_1, k_2, \dots, k_s 是任意数。

定理 3.5: 基础解系

设齐次方程组 $Ax = 0$ 的一组解向量 $\eta_1, \eta_2, \dots, \eta_s$ 满足:

- 向量组 $\{\eta_1, \eta_2, \dots, \eta_s\}$ 线性无关
- 任意齐次方程组 $Ax = 0$ 的解都可以被表示成向量组 $\{\eta_1, \eta_2, \dots, \eta_s\}$ 的线性组合

则称 $\eta_1, \eta_2, \dots, \eta_s$ 是齐次方程组 $Ax = 0$ 的一个基础解系。

定理 3.6

齐次方程组 $Ax = 0$ 的基础解系的向量个数等于 $n - r(A)$, 其中 n 是未知数的个数, $r(A)$ 是矩阵 A 的秩。

定义 3.4: 齐次方程组的通解

齐次线性方程组 $Ax = 0$ 的所有解的集合被称作齐次线性方程组 $Ax = 0$ 的通解。其结构如下:

定理 3.7: 齐次通解结构

设齐次线性方程组 $Ax = 0$ 的一个基础解系为 $\eta_1, \eta_2, \dots, \eta_s$, 那么齐次线性方程组 $Ax = 0$ 的通解可以表示成:

$$x = k_1\eta_1 + k_2\eta_2 + \dots + k_s\eta_s$$

其中 k_1, k_2, \dots, k_s 是任意数。

下面给出齐次方程组通解的求法:

算法步骤 3.1: 齐次方程组通解求法

步骤一: 把齐次线性方程组 $Ax = 0$ 的增广矩阵 $(A, 0)$ 化为行最简形矩阵。

步骤二: 设 $r(A) = r$, 那么前 r 个未知数 x_1, x_2, \dots, x_r 是主变量, 后 $n - r$ 个未知数 $x_{r+1}, x_{r+2}, \dots, x_n$ 是自由变量。把每一个自由变量依次设为 1, 其余自由变量设为 0, 求出对应的主变量的值, 就可以得到一个基础解系的向量。重复这个过程, 直到把所有自由变量都设为 1 为止。

最终得到的通解的形式应该为:

$$x = k_1 \begin{pmatrix} -a_{1,r+1} \\ -a_{2,r+1} \\ \vdots \\ -a_{r,r+1} \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + k_2 \begin{pmatrix} -a_{1,r+2} \\ -a_{2,r+2} \\ \vdots \\ -a_{r,r+2} \\ 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \cdots + k_s \begin{pmatrix} -a_{1,n} \\ -a_{2,n} \\ \vdots \\ -a_{r,n} \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

3.3 非齐次方程组解的结构

定理 3.8: 非齐次通解结构

设非齐次线性方程组 $Ax = b$ 有解, η 是其对应的齐次线性方程组 $Ax = 0$ 的一个基础解系, x_0 是非齐次线性方程组 $Ax = b$ 的一个特解, 那么非齐次线性方程组 $Ax = b$ 的通解可以表示成:

$$x = x_0 + k_1\eta_1 + k_2\eta_2 + \cdots + k_s\eta_s$$

其中 k_1, k_2, \dots, k_s 是任意数, $\eta_1, \eta_2, \dots, \eta_s$ 是齐次线性方程组 $Ax = 0$ 的一个基础解系。注意 x_0 是非齐次线性方程组 $Ax = b$ 的一个特解。其选取是任意的。

下面给出非齐次方程组通解的求法:

算法步骤 3.2: 非齐次方程组通解求法

步骤一: 把非齐次线性方程组 $Ax = b$ 的增广矩阵 (A, b) 化为行最简形矩阵。

步骤二: 设 $r(A) = r$, 那么前 r 个未知数 x_1, x_2, \dots, x_r 是主变量, 后 $n - r$ 个未知数 $x_{r+1}, x_{r+2}, \dots, x_n$ 是自由变量。把每一个自由变量依次设为 1, 其余自由变量设为 0, 求出对应的主变量的值, 就可以得到一个基础解系的向量。重复这个过程, 直到把所有自由变量都设为 1 为止。

最终得到的通解的形式应该为:

$$x = x_0 + k_1 \begin{pmatrix} -a_{1,r+1} \\ -a_{2,r+1} \\ \vdots \\ -a_{r,r+1} \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + k_2 \begin{pmatrix} -a_{1,r+2} \\ -a_{2,r+2} \\ \vdots \\ -a_{r,r+2} \\ 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \cdots + k_s \begin{pmatrix} -a_{1,n} \\ -a_{2,n} \\ \vdots \\ -a_{r,n} \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

其中 x_0 是非齐次线性方程组 $Ax = b$ 的一个特解。

4 线性空间与线性变换

在本部分，我们将介绍线性空间和线性变换的基本概念和性质。线性空间是线性代数的核心概念之一，它为我们提供了一个统一的框架来研究各种数学对象之间的线性关系。线性变换则是在线性空间之间建立联系的重要工具。从这章开始，我们将逐步引入更抽象的概念，以便更深入地理解线性代数的本质。

4.1 线性空间的基本概念

定义 4.1: 线性空间公理

设 V 是一个非空集合， F 是一个数域，如果在 V 中定义了两个运算：向量加法和数乘，并且这两个运算满足以下八条公理，那么我们称 V 为一个定义在 F 上的线性空间（或向量空间）。

- 向量加法封闭性：对于任意的 $\vec{u}, \vec{v} \in V$, $\vec{u} + \vec{v} \in V$ 。
- 向量加法交换律：对于任意的 $\vec{u}, \vec{v} \in V$, $\vec{u} + \vec{v} = \vec{v} + \vec{u}$ 。
- 向量加法结合律：对于任意的 $\vec{u}, \vec{v}, \vec{w} \in V$, $(\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w})$ 。
- 零向量存在性：存在一个零向量 $\vec{0} \in V$ ，使得对于任意的 $\vec{v} \in V$, $\vec{v} + \vec{0} = \vec{v}$ 。
- 负向量存在性：对于任意的 $\vec{v} \in V$ ，存在一个负向量 $-\vec{v} \in V$ 。
- 数乘封闭性：对于任意的 $\alpha \in F$ 和 $\vec{v} \in V$, $\alpha\vec{v} \in V$ 。
- 数乘分配律：对于任意的 $\alpha, \beta \in F$ 和 $\vec{v} \in V$, $(\alpha + \beta)\vec{v} = \alpha\vec{v} + \beta\vec{v}$ 。
- 数乘结合律：对于任意的 $\alpha, \beta \in F$ 和 $\vec{v} \in V$, $\alpha(\beta\vec{v}) = (\alpha\beta)\vec{v}$ 。

上述公理规定了线性空间具有一系列性质。

性质 4.1: 线性空间的简单性质

- 零向量的唯一性：线性空间中的零向量是唯一的。
- 负向量的唯一性：对于任意向量 $\vec{v} \in V$ ，其负向量 $-\vec{v}$ 也是唯一的。
- 数乘零：对于任意向量 $\vec{v} \in V$ ，有 $0 \cdot \vec{v} = \vec{0}$ 。
- 数乘负一：对于任意向量 $\vec{v} \in V$ ，有 $(-1) \cdot \vec{v} = -\vec{v}$ 。
- 数乘零向量：对于任意标量 $\alpha \in F$ ，有 $\alpha \cdot \vec{0} = \vec{0}$ 。
- 若 $\alpha\vec{v} = \vec{0}$ ，则 $\alpha = 0$ 或 $\vec{v} = \vec{0}$ 。

定义 4.2: 线性子空间

设 V 是定义在数域 F 上的线性空间, W 是 V 的一个非空子集, 如果 W 本身也是一个线性空间 (其加法和数乘运算与 V 中的相同), 那么我们称 W 为 V 的一个子空间。

定理 4.1: 子空间判定

设 V 是定义在数域 F 上的线性空间, W 是 V 的一个非空子集, 那么 W 是 V 的一个子空间的充分必要条件是:

- 对于任意的 $\vec{u}, \vec{v} \in W$, $\vec{u} + \vec{v} \in W$ 。
- 对于任意的 $\alpha \in F$ 和 $\vec{v} \in W$, $\alpha\vec{v} \in W$ 。

即 W 在加法和数乘下封闭。

定义 4.3: 基、维数与坐标

- **线性空间的基:** 设 V 是定义在数域 F 上的线性空间, 如果向量组 $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ 满足:
 - 向量组 $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ 线性无关
 - 向量组 $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ 生成 V 中的任意向量 (即 V 中的任意向量都可以被表示成向量组 $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ 的线性组合)

则称向量组 $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ 为线性空间 V 的一组基 (basis), 向量组中的向量被称作基向量 (basis vector)。

- **线性空间的维数:** 设 V 是定义在数域 F 上的线性空间, 如果 V 有一个有限基 $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$, 那么 V 中任意两个基的向量个数都相同, 这个数被称作线性空间 V 的维数 (dimension), 记作 $\dim V = n$ 。
- **线性空间中的坐标:** 设 V 是定义在数域 F 上的线性空间, $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ 是 V 的一组基, 那么对于 V 中的任意向量 \vec{v} , 都存在唯一的一组数 k_1, k_2, \dots, k_n , 使得:

$$\vec{v} = k_1\vec{e}_1 + k_2\vec{e}_2 + \cdots + k_n\vec{e}_n$$

这组数 (k_1, k_2, \dots, k_n) 被称作向量 \vec{v} 在基 $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ 下的坐标 (coordinates)。用向量表示为: $(k_1, k_2, \dots, k_n)^T$

定义 4.4: 线性空间的同构

设 V 和 W 是定义在数域 F 上的两个线性空间, 如果存在一个双射映射 $T : V \rightarrow W$, 使得对于任意的 $\vec{u}, \vec{v} \in V$ 和任意的 $\alpha \in F$, 有:

- $T(\vec{u} + \vec{v}) = T(\vec{u}) + T(\vec{v})$
- $T(\alpha\vec{v}) = \alpha T(\vec{v})$

则称 T 为线性空间 V 到线性空间 W 的一个同构 (isomorphism), 并称线性空间 V 和线性空间 W 是同构的。

定理 4.2

数域 F 上的任意一个 n 维线性空间都与 F^n 同构。

4.2 线性变换的基本概念**定理 4.3: 基变换**

设 V 是定义在数域 F 上的线性空间, $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ 和 $\{\vec{f}_1, \vec{f}_2, \dots, \vec{f}_n\}$ 是 V 的两个基, 那么存在唯一的非奇异矩阵 P , 使得:

$$(\vec{f}_1, \vec{f}_2, \dots, \vec{f}_n) = (\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n)P$$

这个矩阵 P 被称作从基 $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ 到基 $\{\vec{f}_1, \vec{f}_2, \dots, \vec{f}_n\}$ 的基变换矩阵。

注意 (Note)

千万注意: 这里的矩阵 P 是把旧基表示成新基的线性组合所对应的矩阵, 而不是把新基表示成旧基的线性组合所对应的矩阵。也就是说, 如果我们有:

$$\vec{f}_j = a_{1j}\vec{e}_1 + a_{2j}\vec{e}_2 + \dots + a_{nj}\vec{e}_n$$

那么矩阵 P 的第 j 列就是 $(a_{1j}, a_{2j}, \dots, a_{nj})^T$ 。这也是为什么我们要右乘矩阵 P 。 P 的每一列其实对应的是坐标。

定理 4.4: 坐标变换

设 V 是定义在数域 F 上的线性空间, $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ 和 $\{\vec{f}_1, \vec{f}_2, \dots, \vec{f}_n\}$ 是 V 的两个基, P 是从基 $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ 到基 $\{\vec{f}_1, \vec{f}_2, \dots, \vec{f}_n\}$ 的基变换矩阵。那么对于 V 中的任意向量 \vec{v} , 其在基 $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ 下的坐标为 $(k_1, k_2, \dots, k_n)^T$, 在基 $\{\vec{f}_1, \vec{f}_2, \dots, \vec{f}_n\}$ 下的坐标为 $(l_1, l_2, \dots, l_n)^T$, 则有:

$$(l_1, l_2, \dots, l_n)^T = P^{-1}(k_1, k_2, \dots, k_n)^T$$

也就是说, 坐标变换矩阵是基变换矩阵的逆矩阵。

坐标变换之所以是在原坐标的基础上左乘基变换矩阵的逆矩阵, 是因为我们要把旧坐标表示成新坐标的线性组合。具体来说, 如果我们有:

$$\vec{v} = k_1 \vec{e}_1 + k_2 \vec{e}_2 + \dots + k_n \vec{e}_n = l_1 \vec{f}_1 + l_2 \vec{f}_2 + \dots + l_n \vec{f}_n$$

那么我们可以把新基表示成旧基的线性组合, 代入上式, 就可以得到旧坐标和新坐标之间的关系。经过推导, 我们可以得到上面的公式。

4.3 欧氏空间

欧氏空间 (Euclidean Space) 是线性空间的一个特殊类型, 它不仅具有线性空间的结构, 还具有内积 (inner product) 的结构。内积为我们提供了测量向量长度和向量之间夹角的工具。配备了内积的线性空间被称为欧氏空间, 也称为内积空间。

定义 4.5: 内积

内积: 设 V 是定义在数域 F 上的线性空间, 如果在 V 中定义了一个映射 $\langle \cdot, \cdot \rangle : V \times V \rightarrow F$, 满足以下四条公理, 那么我们称这个映射为 V 上的一个内积, 线性空间 V 配备了这个内积后称为欧氏空间。

内积公理:

- 正定性: 对于任意的 $\vec{v} \in V$, 有 $\langle \vec{v}, \vec{v} \rangle \geq 0$, 且当且仅当 $\vec{v} = \vec{0}$ 时取等号。
- 对称性: 对于任意的 $\vec{u}, \vec{v} \in V$, 有 $\langle \vec{u}, \vec{v} \rangle = \langle \vec{v}, \vec{u} \rangle$ 。
- 线性性: 对于任意的 $\alpha \in F$ 和 $\vec{u}, \vec{v}, \vec{w} \in V$, 有 $\langle \alpha \vec{u} + \vec{v}, \vec{w} \rangle = \alpha \langle \vec{u}, \vec{w} \rangle + \langle \vec{v}, \vec{w} \rangle$ 。
- 共轭对称性: 对于任意的 $\alpha, \beta \in F$ 和 $\vec{u}, \vec{v} \in V$, 有 $\langle \vec{u}, \alpha \vec{v} + \beta \vec{w} \rangle = \bar{\alpha} \langle \vec{u}, \vec{v} \rangle + \bar{\beta} \langle \vec{u}, \vec{w} \rangle$, 其中 $\bar{\alpha}$ 表示 α 的共轭。

定义了内积的实线性空间被称为实欧氏空间, 定义了内积的复线性空间被称为复欧氏空间。由内积我们可以衍生出度量 (metrics) 等其他概念。

定义 4.6: 长度、夹角与距离

- **向量的长度:** 设 V 是定义在数域 F 上的欧氏空间, 对于任意的向量 $\vec{v} \in V$, 我们定义向量 \vec{v} 的长度为:

$$\|\vec{v}\| = \sqrt{\langle \vec{v}, \vec{v} \rangle}$$

向量的长度满足以下性质:

- 非负性: 对于任意的向量 $\vec{v} \in V$, 有 $\|\vec{v}\| \geq 0$, 且当且仅当 $\vec{v} = \vec{0}$ 时取等号。
- 齐次性: 对于任意的标量 $\alpha \in F$ 和向量 $\vec{v} \in V$, 有 $\|\alpha\vec{v}\| = |\alpha|\|\vec{v}\|$ 。
- 三角不等式: 对于任意的向量 $\vec{u}, \vec{v} \in V$, 有 $\|\vec{u} + \vec{v}\| \leq \|\vec{u}\| + \|\vec{v}\|$ 。

这里的三角不等式又称为柯西-施瓦茨不等式 (Cauchy-Schwarz Inequality), 是内积空间中的一个重要不等式。不等式取等号的充分必要条件是向量 \vec{u} 和 \vec{v} 线性相关。

- **向量之间的夹角:** 设 V 是定义在数域 F 上的欧氏空间, 对于任意的非零向量 $\vec{u}, \vec{v} \in V$, 我们定义向量 \vec{u} 和向量 \vec{v} 之间的夹角 θ 为:

$$\cos \theta = \frac{\langle \vec{u}, \vec{v} \rangle}{\|\vec{u}\| \|\vec{v}\|}$$

夹角 θ 满足以下性质:

- 夹角的范围: 夹角 θ 的取值范围为 $0 \leq \theta \leq \pi$ 。
- 正交性: 如果 $\langle \vec{u}, \vec{v} \rangle = 0$, 则称向量 \vec{u} 和向量 \vec{v} 正交, 此时夹角 $\theta = \frac{\pi}{2}$ 。
- **向量之间的距离:** 设 V 是定义在数域 F 上的欧氏空间, 对于任意的向量 $\vec{u}, \vec{v} \in V$, 我们定义向量 \vec{u} 和向量 \vec{v} 之间的距离 $d(\vec{u}, \vec{v})$ 为:

$$d(\vec{u}, \vec{v}) = \|\vec{u} - \vec{v}\|$$

距离 $d(\vec{u}, \vec{v})$ 满足以下性质:

- 非负性: 对于任意的向量 $\vec{u}, \vec{v} \in V$, 有 $d(\vec{u}, \vec{v}) \geq 0$, 且当且仅当 $\vec{u} = \vec{v}$ 时取等号。
- 对称性: 对于任意的向量 $\vec{u}, \vec{v} \in V$, 有 $d(\vec{u}, \vec{v}) = d(\vec{v}, \vec{u})$ 。
- 三角不等式: 对于任意的向量 $\vec{u}, \vec{v}, \vec{w} \in V$, 有 $d(\vec{u}, \vec{w}) \leq d(\vec{u}, \vec{v}) + d(\vec{v}, \vec{w})$ 。

4.3.1 正交基与正交化

定义 4.7: 正交基

1. **正交基:** 设 V 是定义在数域 F 上的欧氏空间, 如果向量组 $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ 满足:

- 向量组 $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ 线性无关
- 对于任意的 $i \neq j$, 有 $\langle \vec{e}_i, \vec{e}_j \rangle = 0$

则称向量组 $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ 为欧氏空间 V 的一组正交基。

度量矩阵: 设 V 是定义在数域 F 上的欧氏空间, $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ 是 V 的一组基, 那么我们定义度量矩阵 G 为:

$$G = \begin{pmatrix} \langle \vec{e}_1, \vec{e}_1 \rangle & \langle \vec{e}_1, \vec{e}_2 \rangle & \cdots & \langle \vec{e}_1, \vec{e}_n \rangle \\ \langle \vec{e}_2, \vec{e}_1 \rangle & \langle \vec{e}_2, \vec{e}_2 \rangle & \cdots & \langle \vec{e}_2, \vec{e}_n \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \vec{e}_n, \vec{e}_1 \rangle & \langle \vec{e}_n, \vec{e}_2 \rangle & \cdots & \langle \vec{e}_n, \vec{e}_n \rangle \end{pmatrix}$$

在后面的学习中我们会知道, 度量矩阵 G 是一个对称正定矩阵。

对于在这个基下的两个向量 α 和 β , 其坐标分别为 $(a_1, a_2, \dots, a_n)^T$ 和 $(b_1, b_2, \dots, b_n)^T$, 则有:

$$\langle \alpha, \beta \rangle = (a_1, a_2, \dots, a_n) G (b_1, b_2, \dots, b_n)^T$$

我们定义度量矩阵的主要目的是为了简化计算。

2. **标准正交基:** 设 V 是定义在数域 F 上的欧氏空间, 如果向量组 $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ 满足:

- 向量组 $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ 线性无关
- 对于任意的 $i \neq j$, 有 $\langle \vec{e}_i, \vec{e}_j \rangle = 0$
- 对于任意的 i , 有 $\langle \vec{e}_i, \vec{e}_i \rangle = 1$

则称向量组 $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ 为欧氏空间 V 的一组标准正交基。

定理 4.5: 格拉姆-施密特正交化

设 V 是定义在数域 F 上的欧氏空间, $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$ 是 V 的一组线性无关向量组, 那么我们可以通过格拉姆-施密特正交化过程得到一组正交基 $\{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n\}$, 其过程如下:

- 设 $\vec{u}_1 = \vec{v}_1$
- 对于 $k = 2, 3, \dots, n$, 设

$$\vec{u}_k = \vec{v}_k - \sum_{j=1}^{k-1} \frac{\langle \vec{v}_k, \vec{u}_j \rangle}{\langle \vec{u}_j, \vec{u}_j \rangle} \vec{u}_j$$

则向量组 $\{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n\}$ 就是欧氏空间 V 的一组正交基。

定义 4.8: 正交投影与正交矩阵

- * **正交投影:** 设 V 是定义在数域 F 上的欧氏空间, W 是 V 的一个子空间, 对于任意的向量 $\vec{v} \in V$, 存在唯一的向量 $\vec{w} \in W$, 使得向量 $\vec{v} - \vec{w}$ 与子空间 W 中的任意向量都正交。这个向量 \vec{w} 被称作向量 \vec{v} 在子空间 W 上的正交投影。(此部分考试不涉及)
- **正交矩阵:** 设 V 是定义在数域 F 上的欧氏空间, 如果矩阵 Q 满足 $Q^T Q = I$, 则称矩阵 Q 为一个正交矩阵。正交矩阵具有以下性质:
 - 保长度: 对于任意的向量 $\vec{v} \in V$, 有 $\|Q\vec{v}\| = \|\vec{v}\|$ 。
 - 保内积: 对于任意的向量 $\vec{u}, \vec{v} \in V$, 有 $\langle Q\vec{u}, Q\vec{v} \rangle = \langle \vec{u}, \vec{v} \rangle$ 。
 - 行列式的绝对值为 1: 有 $\det(Q) = \pm 1$ 。

事实上我们在内积空间中讨论正交矩阵大多是通过下面这个定义完成的:

定义 4.9: 正交矩阵

设 A 是一个 n 阶实矩阵, 如果 A 的列向量组构成 \mathbb{R}^n 的一组标准正交基, 则称 A 为一个正交矩阵。即:

$$A = (\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n), \quad \langle \vec{a}_i, \vec{a}_j \rangle = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

定理 4.6

设 A 是一个 n 阶实矩阵, 则 A 为正交矩阵的充分必要条件是 $A^T = A^{-1}$ 。同时, A^{-1} 和 A^T , A^* 也是一个正交矩阵。

定理 4.7

设 A 和 B 是两个 n 阶正交矩阵, 则 AB 也是一个 n 阶正交矩阵。

4.4 线性变换

定义 4.10: 线性变换

1. 线性变换的定义: 设 V 和 W 是定义在数域 F 上的两个线性空间, 如果存在一个映射 $T : V \rightarrow W$, 满足对于任意的 $\vec{u}, \vec{v} \in V$ 和任意的 $\alpha \in F$, 有:

- $T(\vec{u} + \vec{v}) = T(\vec{u}) + T(\vec{v})$
- $T(\alpha\vec{v}) = \alpha T(\vec{v})$

则称 T 为线性空间 V 到线性空间 W 的一个线性变换 (linear transformation)。

特别的, 对于一类线性变换 $T : V \rightarrow V$, V 是全体 n 阶对称矩阵, P 是给定的 n 阶可逆矩阵, 则称 $\sigma(A) = P^T AP$ 为合同变换。我们会在后面讨论这种线性变换。

另一类经常见到的线性变换是 n 阶方阵的相似变换, 即 $T(A) = P^{-1}AP$, 其中 P 是给定的 n 阶可逆矩阵。我们会在后面讨论这些内容。

2. 线性变换的值域, 核, 秩和零度: 设 $T : V \rightarrow W$ 是线性空间 V 到线性空间 W 的一个线性变换, 则:

- 值域 (image): 线性变换 T 的值域是线性空间 W 的一个子空间, 记作 $\text{Im}(T) = \{T(\vec{v}) \mid \vec{v} \in V\}$ 。
- 核 (kernel): 线性变换 T 的核是线性空间 V 的一个子空间, 记作 $\text{Ker}(T) = \{\vec{v} \in V \mid T(\vec{v}) = \vec{0}\}$ 。
- 秩 (rank): 线性变换 T 的秩是其值域的维数, 记作 $\text{rank}(T) = \dim(\text{Im}(T))$ 。
- 零度 (nullity): 线性变换 T 的零度是其核的维数, 记作 $\text{nullity}(T) = \dim(\text{Ker}(T))$ 。

定义: 对于给定空间上的线性变换 T , 我们定义:

定理 4.8: 秩-零度定理

设 $T : V \rightarrow W$ 是线性空间 V 到线性空间 W 的一个线性变换，则有：

$$\text{rank}(T) + \text{nullity}(T) = \dim(V)$$

其中 $\dim(V)$ 表示线性空间 V 的维数。

定义 4.11: 线性变换的运算

设 V 为定义在数域 F 上的线性空间， $k \in F$, T_1 和 T_2 是两个线性变换，定义：

- $(T_1 + T_2)(v) = T_1(v) + T_2(v)$, 对任意 $v \in V$, 称为线性变换的和。
- $(kT_1)(v) = k(T_1(v))$, 对任意 $v \in V$, 称为线性变换的纯量乘积。
- $(T_2T_1)(v) = T_2(T_1(v))$, 对任意 $v \in V$, 称为线性变换的积。

性质 4.2: 线性变换的性质

设 T 是定义在 F 上的线性空间 V 的线性变换，则有：

- $T(\vec{0}) = \vec{0}$
- $T(-\vec{v}) = -T(\vec{v})$, 对任意的 $\vec{v} \in V$
- $T(\sum_{i=1}^n \alpha_i \vec{v}_i) = \sum_{i=1}^n \alpha_i T(\vec{v}_i)$, 对任意的 $\alpha_i \in F$ 和 $\vec{v}_i \in V$

也就是说，线性变换保持向量的线性组合结构。线性相关的向量经过线性变换后仍然线性相关，但是线性无关的向量经过线性变换后可能变为线性相关。

定理 4.9: 线性变换的矩阵表示

线性变换在给定基下的矩阵表示: 设 V 和 W 是定义在数域 F 上的两个线性空间, $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ 是 V 的一组基, $\{\vec{f}_1, \vec{f}_2, \dots, \vec{f}_m\}$ 是 W 的一组基, $T: V \rightarrow W$ 是线性空间 V 到线性空间 W 的一个线性变换。那么对于任意的 $i = 1, 2, \dots, n$, 存在唯一的一组数 $a_{1i}, a_{2i}, \dots, a_{mi}$, 使得:

$$T(\vec{e}_i) = a_{1i}\vec{f}_1 + a_{2i}\vec{f}_2 + \dots + a_{mi}\vec{f}_m$$

我们将这些数 a_{ji} 按列排列成一个 $m \times n$ 的矩阵:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

称矩阵 A 为线性变换 T 在基 $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ 和基 $\{\vec{f}_1, \vec{f}_2, \dots, \vec{f}_m\}$ 下的矩阵表示。进而, 两组基之间有如下表示:

$$(T(\vec{e}_1), T(\vec{e}_2), \dots, T(\vec{e}_n)) = (\vec{f}_1, \vec{f}_2, \dots, \vec{f}_m)A$$

对于 V 中的任意向量 \vec{v} , 其在基 $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ 下的坐标为 $(k_1, k_2, \dots, k_n)^T$, 有:

$$T(\vec{v}) = T\left(\sum_{i=1}^n k_i \vec{e}_i\right) = \sum_{i=1}^n k_i T(\vec{e}_i) = \sum_{i=1}^n k_i \left(\sum_{j=1}^m a_{ji} \vec{f}_j\right) = \sum_{j=1}^m \left(\sum_{i=1}^n a_{ji} k_i\right) \vec{f}_j.$$

如果我们将 $T(\vec{v})$ 在基 $\{\vec{f}_1, \vec{f}_2, \dots, \vec{f}_m\}$ 下的坐标表示为 $(l_1, l_2, \dots, l_m)^T$, 则有:

$$(l_1, l_2, \dots, l_m)^T = A(k_1, k_2, \dots, k_n)^T$$

千万注意 A 在坐标变换是左乘的, 因为我们是把新坐标表示成旧坐标的线性组合。然而在线性变换的基变换中, 我们是把旧基表示成新基的线性组合, 所以是右乘的。

定理 4.10: 线性变换在不同基下矩阵的关系

设 $\vec{\epsilon}_1, \vec{\epsilon}_2, \dots, \vec{\epsilon}_n$ 是线性空间 V 的一组基, $\vec{\epsilon}'_1, \vec{\epsilon}'_2, \dots, \vec{\epsilon}'_n$ 是 V 的另一组基。设 P 是从基 $\{\vec{\epsilon}_1, \vec{\epsilon}_2, \dots, \vec{\epsilon}_n\}$ 到基 $\{\vec{\epsilon}'_1, \vec{\epsilon}'_2, \dots, \vec{\epsilon}'_n\}$ 的基变换矩阵。设线性变换 $T: V \rightarrow V$ 在基 $\{\vec{\epsilon}_1, \vec{\epsilon}_2, \dots, \vec{\epsilon}_n\}$ 下的矩阵表示为 A , 在基 $\{\vec{\epsilon}'_1, \vec{\epsilon}'_2, \dots, \vec{\epsilon}'_n\}$ 下的矩阵表示为 B 。则有:

$$B = P^{-1}AP$$

换言之, 线性变换在不同基下的矩阵表示是相似的。注意是左乘 P 的逆, 右乘 P 。这是因为我们要把旧基表示成新基的线性组合。而 P 是把旧基表示成新基的线性组合所对应的矩阵。

4.5 相似对角化

在上一章节中我们提到, 线性变换在不同基下的矩阵表示是相似的。如果我们能够找到一个合适的基, 使得线性变换在该基下的矩阵表示为对角矩阵, 则称该线性变换是相似对角化的。这种基使得该种线性变换表现出的几何特征是沿着各个坐标轴独立伸缩的。这是更明白和清晰的表现形式。

我们先对上一章的末尾提到的线性变换在不同基下的矩阵关系进行一些变形:

$$B = P^{-1}AP$$

我们实质上想寻找的是一个合适的基变换矩阵 P , 使得 B 是一个对角矩阵。也就是说, 我们想找到一个可逆矩阵 P , 使得 $P^{-1}AP$ 是一个对角矩阵。这就是即将我们在本章节中讨论的矩阵的相似对角化问题。

对上述方程, 我们左乘 P 得到:

$$AP = PB$$

将 P 表征为列向量的格式: 即为:

$$P = (\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n)$$

那么事实上, 上式可以写成:

$$\forall i \in \{1, 2, \dots, n\}, \quad A\vec{v}_i = \lambda_i \vec{v}_i$$

而我们的目标就可以相应的变成求解对应的 λ 和 \vec{v}_i , 使得上式成立。这就是我们即将在下面内容中讨论的矩阵的特征值和特征向量的问题。

定义 4.12: 特征值与特征向量

特征值与特征向量: 设 V 是定义在数域 F 上的线性空间, $T : V \rightarrow V$ 是线性空间 V 到自身的一个线性变换。如果存在一个非零向量 $\vec{v} \in V$ 和一个标量 $\lambda \in F$, 使得:

$$T(\vec{v}) = \lambda \vec{v}$$

则称 λ 为线性变换 T 的一个特征值, \vec{v} 为对应于特征值 λ 的一个特征向量。

通常的, 为了更好的进行运算, 我们默认采用单位矩阵基下的线性变换, 我们用矩阵 A 来表示线性变换 T 。则上式可以写成:

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

这就是我们在前面章节中讨论的矩阵的特征值和特征向量的定义。

对于这个定义, 我们进行移项操作:

$$(A - \lambda I)\vec{v} = \vec{0}$$

由于 \vec{v} 是非零向量, 所以矩阵 $A - \lambda I$ 是奇异矩阵, 即:

$$\det(A - \lambda I) = 0$$

这是 λ 是线性变换 T 的特征值的充分必要条件。我们称多项式 $\det(A - \lambda I)$ 为线性变换 T 的 **特征多项式**。

同时, 对于每一个特征值 λ , 我们可以通过解线性方程组 $(A - \lambda I)\vec{v} = \vec{0}$ 来求得对应的特征向量 \vec{v} 。而这样求得的特征向量并不唯一。对于任意的非零标量 $k \in F$, 如果 \vec{v} 是对应于特征值 λ 的一个特征向量, 则 $k\vec{v}$ 也是对应于特征值 λ 的一个特征向量。

因此, 对应于每一个特征值 λ , 我们通过求解方程组可以得到一个通解, 进而可以得到一个由所有对应特征向量构成的子空间, 称为 **特征子空间**, 更精确的, 是矩阵 A 对应于特征值 λ 的 **特征子空间**。

特别的, 特征子空间的维度称为该特征值的 **几何重数**, 它表示对应于该特征值的线性无关特征向量的个数。

因此, 我们可以梳理出求解矩阵的特征值和特征向量的步骤:

算法步骤 4.1: 求解特征值与特征向量

- 计算特征多项式方程 $\det(A - \lambda I) = 0$ 。
- 解特征多项式方程, 求得所有的特征值 λ 。
- 对于每一个特征值 λ , 解线性方程组 $(A - \lambda I)\vec{v} = \vec{0}$, 求得对应的特征向量 \vec{v} 。
- 对于每一个特征值 λ , 通过求解方程组得到基础解系。
- 对每一个特征值 λ , 由对应的基础解系构造出特征子空间。

然而, 我们在上面的章节中了解到, 所有的矩阵本质上就是一种线性变换, 那么是否存在某种线性变换, 不能够被对角化呢?

这就涉及到了矩阵可否相似对角化的判定方法。我们先给出一个定义:

对于特征多项式, 它总可以写成复数域上多个不可约因子的乘积形式:

$$|\lambda I - A| = (\lambda - \lambda_1)^{m_1}(\lambda - \lambda_2)^{m_2} \cdots (\lambda - \lambda_k)^{m_k}$$

其中 $\lambda_1, \lambda_2, \dots, \lambda_k$ 是矩阵 A 的不同特征值, 我们定义 m_i 为该特征值的 **代数重数**。

然后我们就可以给出下述定理:

定理 4.11

对任意矩阵 A 的任意特征值, 其几何重数不大于代数重数。

证明. 设 λ_0 是 n 阶矩阵 A 的一个特征值。记其几何重数为 s , 代数重数为 t 。则根据定义:

- $s = \dim(\text{Ker}(\lambda_0 I - A))$ 那么我们有理由把矩阵 A 重写为分块矩阵形式:

$$A = M \begin{pmatrix} \lambda_0 I_s & O \\ O & B \end{pmatrix} M^{-1}$$

- 而代数重数可以通过计算特征多项式来得到:

$$\det(\lambda I - A) = \det \left(\lambda I - M \begin{pmatrix} \lambda_0 I_s & O \\ O & B \end{pmatrix} M^{-1} \right) = \det(M^{-1}(\lambda I - A)M) = \det \begin{pmatrix} ((\lambda - \lambda_0)I_s & O \\ O & \lambda I - B \end{pmatrix}$$

因此, 我们有:

$$\det(\lambda I - A) = (\lambda - \lambda_0)^s \det(\lambda I - B)$$

从而, λ_0 作为特征多项式的根的重数至少为 s 。因此, 我们得出结论: $s \leq t$, 即几何重数不大于代数重数。

□

定理 4.12: 矩阵相似对角化的充分必要条件

一个 n 阶矩阵 A 可以相似对角化的充分必要条件是对于 A 的每一个特征值，其几何重数等于代数重数。

证明. 设 A 是一个 n 阶矩阵，我们先证明充分性。假设对于 A 的每一个特征值，其几何重数等于代数重数。设 A 的不同特征值为 $\lambda_1, \lambda_2, \dots, \lambda_k$ ，对应的几何重数和代数重数均为 m_1, m_2, \dots, m_k 。则根据定义，对于每一个特征值 λ_i ，存在 m_i 个线性无关的特征向量 $\vec{v}_{i1}, \vec{v}_{i2}, \dots, \vec{v}_{im_i}$ 。将所有这些特征向量组合成一个矩阵：

$$P = (\vec{v}_{11}, \vec{v}_{12}, \dots, \vec{v}_{1m_1}, \vec{v}_{21}, \dots, \vec{v}_{km_k})$$

由于这些特征向量线性无关，矩阵 P 是可逆的。现在我们计算 $P^{-1}AP$ ：

$$P^{-1}AP = \begin{pmatrix} \lambda_1 I_{m_1} & O & \cdots & O \\ O & \lambda_2 I_{m_2} & \cdots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \cdots & \lambda_k I_{m_k} \end{pmatrix}$$

这就是一个对角矩阵。因此，我们证明了充分性。

接下来我们证明必要性。假设矩阵 A 可以相似对角化，即存在一个可逆矩阵 P ，使得：

$$P^{-1}AP = D$$

其中 D 是一个对角矩阵。设 D 的对角元素为 d_1, d_2, \dots, d_n 。则这些对角元素就是矩阵 A 的特征值。设 λ 是矩阵 A 的一个特征值，则存在若干个对角元素等于 λ 。设这些对角元素的个数为 m 。则对应于这些对角元素，存在 m 个线性无关的特征向量。因此，矩阵 A 的特征值 λ 的几何重数至少为 m 。同时，由于对角矩阵 D 的特征多项式可以直接计算出来，可知 λ 的代数重数也为 m 。因此，我们证明了必要性。□

因此，在实际计算中，如果我们发现一个矩阵的某个特征值的几何重数小于代数重数，那么就省事了，该矩阵一定不能相似对角化。

下面给出一些性质：

性质 4.3: 特征向量的线性无关性

一个矩阵的任意两个特征向量一定线性无关。

证明. 本定理显然， P 必须可逆，必须满秩，列向量必然线性无关。□

性质 4.4

如果一个矩阵有 n 个不同的特征值，则该矩阵一定可以相似对角化。

性质 4.5: 迹的性质

一个矩阵的特征值的和等于该矩阵的迹 $\text{tr}(A)$, 即主对角线元素之和。

证明. 设 A 是一个 n 阶矩阵, 其特征多项式为:

$$\det(\lambda I - A) = (\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_n)$$

展开右边, 我们得到:

$$\det(\lambda I - A) = \lambda^n - (\lambda_1 + \lambda_2 + \cdots + \lambda_n)\lambda^{n-1} + \cdots$$

同时, 根据行列式的性质, 我们也可以直接展开左边, 得到:

$$\det(\lambda I - A) = \lambda^n - \text{tr}(A)\lambda^{n-1} + \cdots$$

我们之所以可以得到这个结论, 主要是因为 $n-1$ 次项的系数只可能来自矩阵 A 的主对角线元素, 即迹 $\text{tr}(A)$ 。比较两个展开式中 λ^{n-1} 的系数, 我们得到:

$$\lambda_1 + \lambda_2 + \cdots + \lambda_n = \text{tr}(A)$$

因此, 我们证明了该性质。 □

性质 4.6: 行列式的性质

一个矩阵的特征值的积等于该矩阵的行列式 $\det(A)$ 。

证明. 我们先对矩阵 A 相似对角化得到:

$$P^{-1}AP = D$$

其中 D 是一个对角矩阵, 其对角元素为矩阵 A 的特征值 $\lambda_1, \lambda_2, \dots, \lambda_n$ 。则有:

$$\det(A) = \det(PDP^{-1}) = \det(D) = \lambda_1 \lambda_2 \cdots \lambda_n$$

因此, 我们证明了该性质。 □

定理 4.13: 零化多项式与极小多项式

设 A 是一个 n 阶矩阵, 如果存在一个非零多项式 $p(\lambda)$, 使得 $p(A) = 0$, 则称 $p(\lambda)$ 为矩阵 A 的一个零化多项式。在所有的零化多项式中, 存在唯一一个首一多项式 $m(\lambda)$, 使得对于任意的零化多项式 $p(\lambda)$, 都有 $m(\lambda)$ 整除 $p(\lambda)$ 。则称 $m(\lambda)$ 为矩阵 A 的极小多项式。

证明. 设 A 是一个 n 阶矩阵, 我们考虑所有的零化多项式构成的集合:

$$S = \{p(\lambda) \mid p(A) = 0, p(\lambda) \neq 0\}$$

由于多项式的次数是非负整数, 且每个多项式的次数有限, 因此集合 S 中存在一个次数最小的多项式, 记为 $m(\lambda)$ 。我们可以将 $m(\lambda)$ 写成首一形式, 即最高次项系数为 1。

现在, 我们证明 $m(\lambda)$ 整除任意的零化多项式 $p(\lambda)$ 。设 $p(\lambda)$ 是任意一个零化多项式, 则存在多项式 $q(\lambda)$ 和余数 $r(\lambda)$, 使得:

$$p(\lambda) = m(\lambda)q(\lambda) + r(\lambda)$$

其中 $r(\lambda)$ 的次数小于 $m(\lambda)$ 的次数。现在, 我们计算 $p(A)$:

$$p(A) = m(A)q(A) + r(A)$$

由于 $p(A) = 0$ 且 $m(A) = 0$, 因此有 $r(A) = 0$ 。但是 $r(\lambda)$ 的次数小于 $m(\lambda)$ 的次数, 根据 $m(\lambda)$ 是零化多项式中次数最小的多项式, 我们得到 $r(\lambda) = 0$ 。因此, 我们得出结论:

$$p(\lambda) = m(\lambda)q(\lambda)$$

即 $m(\lambda)$ 整除 $p(\lambda)$ 。

最后, 我们证明 $m(\lambda)$ 的唯一性。假设存在另一个首一多项式 $m'(\lambda)$, 使得对于任意的零化多项式 $p(\lambda)$, 都有 $m'(\lambda)$ 整除 $p(\lambda)$ 。则根据前面的证明过程, 我们可以得到:

$$m'(\lambda) = m(\lambda)q_1(\lambda)$$

$$m(\lambda) = m'(\lambda)q_2(\lambda)$$

其中 $q_1(\lambda)$ 和 $q_2(\lambda)$ 是某些多项式。代入第一个等式到第二个等式中, 我们得到:

$$m(\lambda) = m(\lambda)q_1(\lambda)q_2(\lambda)$$

由于 $m(\lambda)$ 是首一多项式, 因此 $q_1(\lambda)q_2(\lambda) = 1$ 。这意味着 $q_1(\lambda)$ 和 $q_2(\lambda)$ 都是常数多项式, 且它们的乘积为 1。因此, 我们得出结论:

$$m'(\lambda) = m(\lambda)$$

综上所述, 我们证明了极小多项式的存在性和唯一性。 □

定理 4.14

一切零化多项式都被极小多项式整除。一切零化多项式都被特征多项式整除, 极小多项式可被特征多项式整除。

我们在这里不再赘述证明过程, 考试也不会涉及。

定理 4.15

若 A 可相似对角化, 那么其转置, 共轭转置, 逆矩阵, 伴随矩阵也可相似对角化, 且它们的特征值与 A 经过对应代数操纵后的结果相同。特征向量不变。

证明. 设 A 可相似对角化, 则存在可逆矩阵 P , 使得:

$$P^{-1}AP = D$$

其中 D 是一个对角矩阵, 其对角元素为矩阵 A 的特征值 $\lambda_1, \lambda_2, \dots, \lambda_n$ 。

1. 对于转置矩阵 A^T , 我们有:

$$(P^{-1}AP)^T = D^T$$

因为对角矩阵的转置仍然是对角矩阵, 所以 A^T 也可相似对角化, 且其特征值与 A 相同。

2. 对于共轭转置矩阵 A^* , 我们有:

$$(P^{-1}AP)^* = D^*$$

因为对角矩阵的共轭转置仍然是对角矩阵, 所以 A^* 也可相似对角化, 且其特征值为 $\overline{\lambda_1}, \overline{\lambda_2}, \dots, \overline{\lambda_n}$ 。

3. 对于逆矩阵 A^{-1} , 我们有:

$$(P^{-1}AP)^{-1} = D^{-1}$$

因为对角矩阵的逆仍然是对角矩阵, 所以 A^{-1} 也可相似对角化, 且其特征值为 $\frac{1}{\lambda_1}, \frac{1}{\lambda_2}, \dots, \frac{1}{\lambda_n}$ 。

□

4.5.1 实对称矩阵的相似对角化

定理 4.16

实对称矩阵一定可以相似对角化。

证明. 我们采用归纳法完成本定理的证明。

基础步骤: 当 $n = 1$ 时, 实对称矩阵只有一个元素, 显然可以正交相似于对角矩阵, 且其特征值为该元素本身。

归纳步骤: 假设任意 $k < n$ 的实对称矩阵都可以正交相似于对角矩阵。现在考虑一个 n 阶实对称矩阵 A 。将 A 写成分块矩阵形式:

$$A = \begin{pmatrix} a_{11} & b^\top \\ b & B \end{pmatrix}$$

其中 a_{11} 是 A 的第一个元素, b 是一个 $(n - 1)$ 维列向量, B 是一个 $(n - 1) \times (n - 1)$ 的实对称矩阵。根据归纳假设, 矩阵 B 可以正交对角化, 即存在正交矩阵 P , 使得:

$$P^T B P = D$$

其中 D 是一个对角矩阵, 其对角元素为矩阵 B 的特征值 $\lambda_2, \lambda_3, \dots, \lambda_n$ 。构造正交矩阵:

$$Q = \begin{pmatrix} 1 & 0 \\ 0 & P \end{pmatrix}$$

则有:

$$Q^T A Q = \begin{pmatrix} a_{11} & b^T P \\ P^T b & D \end{pmatrix} =: M$$

记 $c = P^T b$, 则 $M = \begin{pmatrix} a_{11} & c^T \\ c & D \end{pmatrix}$, 且 M 是实对称矩阵。由于 M 是实对称矩阵, 它至少有一个实特征值 λ_1 和对应的单位特征向量 u_1 。将 u_1 扩展为 \mathbb{R}^n 的一组标准正交基 u_1, u_2, \dots, u_n , 并令 $R_1 = \begin{pmatrix} u_1 & u_2 & \dots & u_n \end{pmatrix}$, 则 R_1 是正交矩阵, 且

$$R_1^T M R_1 = \begin{pmatrix} \lambda_1 & 0 \\ 0 & A_{n-1} \end{pmatrix}$$

其中 A_{n-1} 是一个 $(n - 1) \times (n - 1)$ 实对称矩阵。对 A_{n-1} 应用归纳假设, 存在正交矩阵 R_2 使得

$$R_2^T A_{n-1} R_2 = \text{diag}(\lambda_2, \dots, \lambda_n)$$

$$R = R_1 \begin{pmatrix} 1 & 0 \\ 0 & R_2 \end{pmatrix}$$

则 R 是正交矩阵, 且

$$R^T M R = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \text{diag}(\lambda_2, \dots, \lambda_n) \end{pmatrix} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$$

最后, 记 $T = QR$, 由于 Q 和 R 均为正交矩阵, T 也是正交矩阵, 且

$$T^T A T = (QR)^T A (QR) = R^T (Q^T A Q) R = R^T M R = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$$

因此 A 正交相似于对角矩阵, 归纳步骤完成。由数学归纳法, 任意 n 阶实对称矩阵均可正交相似于对角矩阵。 \square

定理 4.17

实对称矩阵相似对角化得到的特征值均为实数。

本定理的证明略, 无需掌握。

定理 4.18

设 A 是一个 n 阶实对称矩阵，则存在 n 阶正交矩阵 Q ，使得：

$$Q^T A Q = D$$

其中 D 是一个对角矩阵，其对角元素为矩阵 A 的特征值 $\lambda_1, \lambda_2, \dots, \lambda_n$ 。同时，矩阵 A 的不同特征值对应的特征向量是正交的。在这里， $Q^T = Q^{-1}$

求解实对称矩阵正交相似对角阵的方法如下：

算法步骤 4.2: 求解实对称矩阵正交对角化

- 计算实对称矩阵 A 的特征多项式，求解特征多项式方程 $\det(\lambda I - A) = 0$ ，得到所有的特征值 $\lambda_1, \lambda_2, \dots, \lambda_k$ 及其代数重数 m_1, m_2, \dots, m_k 。
- 对于每一个特征值 λ_i ，解线性方程组 $(\lambda_i I - A)\vec{v} = \vec{0}$ ，求得对应的特征向量 $\vec{v}_{i1}, \vec{v}_{i2}, \dots, \vec{v}_{im_i}$ 。
- 对于每一个特征值 λ_i ，通过求解方程组得到的特征向量构成的子空间，使用格拉姆-施密特正交化方法，得到一组正交基 $\vec{u}_{i1}, \vec{u}_{i2}, \dots, \vec{u}_{im_i}$ 。
- 将所有的正交基组合成一个矩阵：

$$Q = (\vec{u}_{11}, \vec{u}_{12}, \dots, \vec{u}_{1m_1}, \vec{u}_{21}, \dots, \vec{u}_{km_k})$$

由于这些特征向量线性无关，矩阵 Q 是正交矩阵。

5 实二次型

5.1 定义和性质

二次型问题缘起于二次曲线和二次曲面。实二次型是指形如：

$$f(x_1, x_2, \dots, x_n) = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j$$

其中，我们规定： $a_{ij} = a_{ji}$ 。基于这一表达式，我们可以将实二次型写成矩阵的形式，也就是二次型的定义：

定义 5.1: 实二次型

设 A 是一个 n 阶实对称矩阵, 则称函数 $f: \mathbb{R}^n \rightarrow \mathbb{R}$ 定义为:

$$f(\vec{x}) = \vec{x}^\top A \vec{x}$$

其中 $\vec{x} = (x_1, x_2, \dots, x_n)^\top \in \mathbb{R}^n$, 为实二次型。

只含有平方项的实二次型称为标准实二次型, 即:

$$f(\vec{x}) = \lambda_1 x_1^2 + \lambda_2 x_2^2 + \dots + \lambda_n x_n^2$$

其中 $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{R}$ 。

而若标准二次型中所有的 λ_i 均为 1, -1, 或 0, 则称为规范实二次型, 即:

$$f(\vec{x}) = \underbrace{x_1^2 + x_2^2 + \dots + x_p^2}_{p \text{ 项}} - \underbrace{x_{p+1}^2 - x_{p+2}^2 - \dots - x_{p+q}^2}_{q \text{ 项}}$$

注意: $p + q \leq n$, 剩下的 $n - p - q$ 项均为 0。

定义 5.2: 非奇异线性变换

设 P 是可逆矩阵, 那么 $x = Py$ 是一个非奇异的线性变换。

定义 5.3: 合同矩阵

设 A 和 B 为 n 阶实对称矩阵, 分别对应实二次型 $f(\vec{x}) = \vec{x}^\top A \vec{x}$ 和 $g(\vec{x}) = \vec{x}^\top B \vec{x}$ 。如果存在一个可逆矩阵 P , 使得 $B = P^\top AP$, 则称矩阵 A 和 B 为 **合同矩阵**, 实二次型 f 和 g 为 **合同二次型**。

定理 5.1

一个二次型的非奇异线性变换得到的依然是二次型, 矩阵合同。

定理 5.2

任何一个实二次型都可以通过适当的非奇异线性变换化为标准实二次型。

求解一个实二次型的方式等价于求解其实对称矩阵正交相似对角阵。这里的过过程是完全相同且等价的, 不再赘述。

$$f = x^\top Ax \xlongequal{x=Qy} y^\top (Q^\top A Q)y = y^\top D y$$

5.2 惯性定理和正定二次型

定理 5.3: 惯性定理

任何实二次型必然存在非奇异线性替换化为规范标准型，且唯一。

定义 5.4: 惯性指数

设实二次型 $f(\vec{x}) = \vec{x}^T A \vec{x}$ 化为规范标准型后，其中正项个数为 p ，负项个数为 q ，则称 (p, q) 为实二次型 f 的惯性指数。其中， p 称为正惯性指数， q 称为负惯性指数， $n - p - q$ 称为零惯性指数。

定义 5.5: 正定性分类

设 A 是一个 n 阶实对称矩阵，则称实二次型 $f(\vec{x}) = \vec{x}^T A \vec{x}$ 为：

- 正定二次型：如果对于任意非零向量 $\vec{x} \in \mathbb{R}^n$ ，都有 $f(\vec{x}) > 0$ 。
- 半正定二次型：如果对于任意向量 $\vec{x} \in \mathbb{R}^n$ ，都有 $f(\vec{x}) \geq 0$ 。
- 负定二次型：如果对于任意非零向量 $\vec{x} \in \mathbb{R}^n$ ，都有 $f(\vec{x}) < 0$ 。
- 半负定二次型：如果对于任意向量 $\vec{x} \in \mathbb{R}^n$ ，都有 $f(\vec{x}) \leq 0$ 。
- 不定二次型：如果存在向量 $\vec{x}_1, \vec{x}_2 \in \mathbb{R}^n$ ，使得 $f(\vec{x}_1) > 0$ 且 $f(\vec{x}_2) < 0$ 。

定理 5.4

实二次型经过非奇异线性变换后，其正惯性指数和负惯性指数不变，正定性不改变。

定义 5.6: 顺序主子式

设 $A = (a_{ij})$ 是一个 n 阶矩阵，则称由矩阵 A 的前 k 行和前 k 列所构成的 k 阶子式：

$$D_k = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{vmatrix}$$

为矩阵 A 的第 k 个顺序主子式。

定理 5.5: 正定二次型的判定

设 A 是一个 n 阶实对称矩阵, 则下列命题等价:

- 实二次型 $f(\vec{x}) = \vec{x}^T A \vec{x}$ 为正定二次型。
- 矩阵 A 的所有特征值均为正数。
- 矩阵 A 的所有顺序主子式均为正数。
- 存在一个可逆矩阵 P , 使得 $A = P^T P$ 。
- A 合同于单位矩阵 I 。

定理 5.6: 负定二次型的判定

设 A 是一个 n 阶实对称矩阵, 则下列命题等价:

- 实二次型 $f(\vec{x}) = \vec{x}^T A \vec{x}$ 为负定二次型。
- 矩阵 A 的所有特征值均为负数。
- 矩阵 A 的顺序主子式 D_k 满足: 当 k 为奇数时, $D_k < 0$; 当 k 为偶数时, $D_k > 0$ 。
- 存在一个可逆矩阵 P , 使得 $A = -P^T P$ 。
- A 合同于 $-I$ 。

6 全书总结

6.1 线性公理

性质 6.1: 线性空间公理总结

线性空间是一个集合 V , 其元素称为向量, 并定义了两个运算: 向量加法和数乘运算。这些运算满足以下公理:

1. 向量加法封闭性: 对于任意的 $\vec{u}, \vec{v} \in V$, 有 $\vec{u} + \vec{v} \in V$ 。
2. 向量加法交换律: 对于任意的 $\vec{u}, \vec{v} \in V$, 有 $\vec{u} + \vec{v} = \vec{v} + \vec{u}$ 。
3. 向量加法结合律: 对于任意的 $\vec{u}, \vec{v}, \vec{w} \in V$, 有 $(\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w})$ 。
4. 零向量存在性: 存在一个零向量 $\vec{0} \in V$, 使得对于任意的 $\vec{v} \in V$, 有 $\vec{v} + \vec{0} = \vec{v}$ 。
5. 负向量存在性: 对于任意的 $\vec{v} \in V$, 存在一个向量 $-\vec{v} \in V$, 使得 $\vec{v} + (-\vec{v}) = \vec{0}$ 。
6. 数乘封闭性: 对于任意的标量 $c \in \mathbb{F}$ 和任意的 $\vec{v} \in V$, 有 $c\vec{v} \in V$ 。
7. 数乘分配律: 对于任意的标量 $c, d \in \mathbb{F}$ 和任意的 $\vec{v} \in V$, 有 $(c+d)\vec{v} = c\vec{v} + d\vec{v}$ 。
8. 数乘结合律: 对于任意的标量 $c, d \in \mathbb{F}$ 和任意的 $\vec{v} \in V$, 有 $c(d\vec{v}) = (cd)\vec{v}$ 。

6.2 秩-零化度定理的多种表示形式

定理 6.1: 秩-零化度定理等价表述

下面这些表述是等价的:

- 对于任意线性方程组, 自由位置量的个数加上主元的个数等于未知量个数。
- 任何一个矩阵, $\text{rank}(A) + \text{nullity}(A) = n$, 其中 n 为矩阵 A 的列数。
- 对于任意线性变换 $T : V \rightarrow W$, 有 $\dim(\text{Im}(T)) + \dim(\text{Ker}(T)) = \dim(V)$ 。

6.3 秩的多种含义

注意 (Note)

秩的多种几何与代数含义:

- 线性方程组有效方程的个数
- 行列式非零子式的最高阶数 (定义)
- 矩阵的行最简阶梯型中非零行的个数
- 矩阵行/列向量组中极大无关组的向量个数
- 线性变换的像空间的维数, 即线性变换保留的信息量

线性代数极强的对称性和连贯的逻辑使得这个学科的知识极具张力.