45 分钟速通线性代数

线性代数要点概览

孔一江, 丁乐其, 袁成希

2023 年 12 月 12 日

西安交通大学

Xi'an Jiaotong University

作品信息

- **标题**：45 分钟速通线性代数: 线性代数要点概览
- **作者**：孔一江, 丁乐其, 袁成希
- **出品时间**：2023 年 12 月 12 日
- **总页数**：45

目录

♯ 标题中的数字 45 实际是指本书页数.

♭ 标有紫色的部分可能超出考试要求, 作为拓展.

Chapter 1

线性空间与线性映射

定义 1 (线性空间). 设有数域 $\mathbb{F}$, 集合 V, $k,l \in \mathbb{F}, \alpha, \beta, \mathbf{x} \in S$. 如果:

1 V 上定义加法运算且封闭;

2 V 上定义数量乘法且封闭;

3 加法和数乘满足:

A 加法交换律

B 加法结合律

C 数乘结合律 $(kl)\mathbf{x} = k(l\mathbf{x})$

D 分配律 *I* $k(\alpha + \beta) = k\alpha + k\beta$

E 分配律 *II* $(k+l)\mathbf{x} = k\mathbf{x} + l\mathbf{x}$

F 存在零元 $\exists\ \mathbf{0} \in S,\ \forall \mathbf{x} \in V,\ \mathbf{x} + \mathbf{0} = \mathbf{x}$

G 存在负元 $\exists\ -\mathbf{x} \in S,\ \forall \mathbf{x} \in V,\ \mathbf{x} + (-\mathbf{x}) = \mathbf{0}$

H $1 \cdot \mathbf{x} = \mathbf{x}$

则称 V 为 $\mathbb{F}$ 上的线性空间. (数域为 $\mathbb{R}$ 的也称为实线性空间.)

性质 1. 线性空间中零元以及任一元素的负元唯一.

定义 2 (子空间). 设 $\emptyset \neq W \subset$ 线性空间 V, 如果 W 也是线性空间, 则称它是 V 的子空间.

定理 1. 设 $\emptyset \neq W \subset$ 线性空间 V, W 对线性运算封闭 $\Leftrightarrow W$ 是 V 的子空间.

定义 3 (基和坐标)**.** *设有线性无关的 $\alpha_1 \dots \alpha_n \in$ 线性空间 V, 使得 $\forall \beta \in V, \exists\, x_1 \dots x_n, \beta = \sum_{i=1}^n x_i \alpha_i$, 则称: $\alpha_1 \dots \alpha_n$ 为 V 的一组基; $\dim V = n$; β 的坐标为 $(x_1, \dots, x_n)$*

只有向量 $\mathbf{0}$ 的空间维数为 0, 没有基.

定理 2 (子空间的交)**.** *两个子空间的交也是子空间.*

定理 3 (子空间的和)**.** *设 V_1, V_2 为 V 的子空间, 则 $V_1 + V_2$ 为子空间 $\{\alpha + \beta | \alpha \in V_1, \beta \in V_2\}$*

定理 4 (子空间维数关系)**.** $\dim V_1 + \dim V_2 = \dim(V_1 + V_2) + \dim(V_1 \cap V_2)$

证明. 以下是在部分下的情况简略证明. 设有 $V_1 \cap V_2$ 的基 $\eta_1, \dots, \eta_b$, 则可以扩充 V_1, V_2 的基 $\eta_1, \dots, \eta_b, \alpha_1, \dots, \alpha_{s_1}$ 和 $\eta_1, \dots, \eta_b, \beta_1, \dots, \beta_{s_2}$. 同时根据定义, $\forall \mathbf{x} \in (V_1 + V_2), \mathbf{x} = k\alpha + l\beta$, 其中 $\alpha \in V_1, \beta \in V_2$, 所以存在系数 $\mathbf{k}, \mathbf{l}$, 使得

$$\mathbf{x} = \left(\sum_{i=1}^{b} k_i \eta_i + \sum_{i=1}^{s_1} k_{i+b} \alpha_i\right) + \left(\sum_{i=1}^{b} l_i \eta_i + \sum_{i=1}^{s_2} l_{i+b} \beta_i\right) = \sum p_i \alpha_i + \sum q_i \beta_i + \sum r_i \eta_i.$$

显然线性无关的 $\{\alpha_i, \beta_i, \eta_i\}$ 也是 $V_1 + V_2$ 的基, $\dim(V_1 + V_2) = b + s_1 + s_2$. $\dim V_1 + \dim V_2 = 2b + s_1 + s_2 = \dim(V_1 + V_2) + \dim(V_1 \cap V_2)$. □

定义 4 (直和)**.** *设 V_1, V_2 都是 V 的子空间, 如果 $\forall \alpha \in V_1 + V_2$, $\alpha = \alpha_1 + \alpha_2$ $(\alpha_1 \in V_1, \alpha_2 \in V_2)$ 的表示唯一, 则称这个和为 V_1, V_2 的直和, 记作 $V_1 \bigoplus V_2$.*

定义 5 (过渡矩阵)**.** *对于已知的两个基 $\alpha_1 \dots \alpha_n$ 和 $\beta_1 \dots \beta_n$, 存在矩阵 A 使得 $\begin{bmatrix} \beta_1 & \cdots & \beta_n \end{bmatrix} = \begin{bmatrix} \alpha_1 & \cdots & \alpha_n \end{bmatrix} A$. A 是从基 $\alpha_{1\dots n}$ 到 $\beta_{1\dots n}$ 的过渡矩阵.*

定理 5 (基与坐标变换)**.** *对于已知的两个基 $(A) : \alpha_{1\dots n}$, $(B) : \beta_{1\dots n}$, (A) 到 (B) 的过渡矩阵为 M, 同一向量在 $(A), (B)$ 基下的坐标分别为 $\mathbf{x}, \mathbf{y}$, 则 $\mathbf{y} = A^{-1}\mathbf{x}$.*

证明. 设 $B = \begin{bmatrix} \beta_1 & \cdots & \beta_n \end{bmatrix}$, $A = \begin{bmatrix} \alpha_1 & \cdots & \alpha_n \end{bmatrix}$, $B = AM$. $A\mathbf{x} = B\mathbf{y} \Rightarrow \mathbf{y} = B^{-1}A\mathbf{x} = M^{-1}\mathbf{x}$ □

基变换是同一线性空间中的变换, 对于两个不同的线性空间之间的, 也有特殊的映射:

定义 6 (同构)**.** *V_1, V_2 是数域 $\mathbb{F}$ 上的两个线性空间, $f : V_1 \to V_2$. 如果 f 是双射且满足线性运算规则 (i.e. $\forall \alpha, \beta \in V_1,\ f(\alpha + \beta) = f(\alpha) + f(\beta)$ 以及 $\forall \alpha \in V_1, k \in \mathbb{F},\ f(k\alpha) = kf(\alpha)$), 则称 f 为同构映射.*

如果两个线性空间中能够建立同构映射, 就说这两个空间同构. 显然, 取坐标是任意 n 维空间到其数域 $\mathbb{F}$ 的线性双射, 因此 $\mathbb{F}^n$ 和 $\mathbb{F}$ 上任一 n 维空间同构.

性质 2. $f : V_1 \to V_2$ 是同构映射, 则

1 $f(\mathbf{0}\ \mathit{of}\ V_1) = \mathbf{0}\ \mathit{of}\ V_2$

2 线性性质 $\forall \alpha_i \in V_1, k_i \in \mathbb{F},\ f(\sum k_i\alpha_i) = \sum k_i f(\alpha_i)$

定理 6. 一个数域上有限维度的两个线性空间同构 $\Leftrightarrow$ 这两个空间维数相同.

证明. $\dim V_1 = n \Rightarrow V_1$ 与 $\mathbb{F}^n$ 同构. 同理 V_2 和 $\mathbb{F}^n$ 同构. 因此 V_1 与 V_2 同构. □

可见, 维数是有限维线性空间唯一不变的本质特征. 那么当向量组中向量个数小于维数的时候, 一定无法表示所有向量. 对此, 有以下定理

定理 7 (扩充定理)**.** 设 $\alpha_1, \dots, \alpha_r$ 是 n 维线性空间 V 中一个线性无关的向量组 $(r < n)$, 则必能找到 $n-r$ 个向量 $\alpha_{r+1}, \dots, \alpha_n$, 使得 $\alpha_1, \dots, \alpha_n$ 是 V 的基.

1.1 Euclid 空间

定义 7 (内积)**.** 在实线性空间 V 上, 对于任意两个向量 $\mathbf{a}, \mathbf{b}$ 定义运算 $(\mathbf{a}, \mathbf{b})$, 并且该运算满足

1 $(\mathbf{a}, \mathbf{b}) = (\mathbf{b}, \mathbf{a})$

2 $(\mathbf{a}+\mathbf{b}, \mathbf{c}) = (\mathbf{a}, \mathbf{c}) + (\mathbf{b}, \mathbf{c})$

3 $(k\mathbf{a}, \mathbf{b}) = k(\mathbf{a}, \mathbf{b})$

4 $(\mathbf{a}, \mathbf{a}) \geq 0$, 等号仅在 $\mathbf{a} = \mathbf{0}$ 时成立.

, 则称 $(\mathbf{a}, \mathbf{b})$ 为 $\mathbf{a}, \mathbf{b}$ 的内积, V 为 *Euclid* 空间.

常见的 Euclid 空间和其内积有 $\mathbb{R}^n$ 与 $(\mathbf{x}, \mathbf{y}) = \mathbf{x}^{\mathrm{T}}\mathbf{y}$, $C[0,1]$ 与 $(f, g) = \int_0^1 f(t)g(t)\mathrm{d}t$ 等.

定义 8 (范数)**.** $\|\mathbf{x}\| = \sqrt{(\mathbf{x}, \mathbf{x})}$.

定理 8 (Cauchy-Schwarz 不等式)**.** $\|(\mathbf{a}, \mathbf{b})\| \leq \|\mathbf{a}\|\|\mathbf{b}\|$.

证明. 线性相关时显然成立. 设有线性无关的 $\mathbf{a}, \mathbf{b}$ 与 $f(t) = (t\mathbf{a}+\mathbf{b}, t\mathbf{a}+\mathbf{b}) = t^2(\mathbf{a}, \mathbf{a}) + 2t(\mathbf{a}, \mathbf{b}) + (\mathbf{b}, \mathbf{b}) \geq 0$. 则二次方程判别式 $(2 \cdot (\mathbf{a}, \mathbf{b}))^2 - 4(\mathbf{a}, \mathbf{a}) \cdot (\mathbf{b}, \mathbf{b}) \leq 0$. □

定义 9 (向量夹角)**.** *Euclid* 空间中非零向量 $\mathbf{a}, \mathbf{b}$ 夹角为

$$\phi = \arccos\left(\frac{(\mathbf{a}, \mathbf{b})}{\|\mathbf{a}\|\,\|\mathbf{b}\|}\right).$$

定义 10 (距离)**.** $\mathrm{dis}(\mathbf{a},\mathbf{b}) = \|\mathbf{a}-\mathbf{b}\|$.

定义 11 (标准正交基)**.** *在 n 维 Euclid 空间中, n 个两两正交的非零向量形成的向量组是 V 的正交基. 如果每个向量都还是单位向量, 则称这个正交基为标准正交基.*

1.2 Gram-Schmidt 正交化和 QR 分解

我们常用这样的方法来获取一组标准正交基:

定义 12 (Gram-Schmidt)**.** *Euclid 空间中有一个基 $\alpha_1,\dots,\alpha_n$, 令*

$$\beta_i = \begin{cases} \alpha_i & i=1 \\ \alpha_i - \sum_{j=1}^{i-1} \frac{(\alpha_i,\beta_j)}{(\beta_j,\beta_j)}\beta_j & other \end{cases}$$

则 $\mathbf{e}_i = \beta_i / \|\beta_i\|$ 是生成的标准正交基.

将标准正交向量组组成方阵, 形成的是正交矩阵. 在我们的教材中定义如下

定义 13 (正交矩阵)**.** *如果实数方阵 A 满足 $A^{\mathrm{T}}A = I$, 则称它为正交矩阵.*

从 Gram-Schmidt 正交化 β 的生成方法可以看出, 原向量组中的任一向量 α_i 等于其在各个 $\beta_j(j\le i)$ 上投影相加的和. 通过对 β(或者 $\mathbf{e}$) 进行列操作得到:

$$\begin{bmatrix}\alpha_1 & \cdots & \alpha_n\end{bmatrix} = \begin{bmatrix}\mathbf{e}_1 & \cdots & \mathbf{e}_n\end{bmatrix} \begin{bmatrix} (\mathbf{e}_1,\alpha_1) & (\mathbf{e}_1,\alpha_2) & \cdots & (\mathbf{e}_1,\alpha_n) \\ 0 & (\mathbf{e}_2,\alpha_2) & \cdots & (\mathbf{e}_2,\alpha_n) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & (\mathbf{e}_n,\alpha_n) \end{bmatrix}$$

i.e.$A = QR$: 如果 A 可逆, 则 Q 为正交矩阵, R 为上三角矩阵.

1.3 线性变换

定义 14 (线性变换)**.** *设 $\mathbb{F}$ 有线性空间 V, W, $T: V \to W$. 如果满足 $\forall \mathbf{a},\mathbf{b} \in V, k \in V$, $T(\mathbf{a}+\mathbf{b}) = T(\mathbf{a}) + T(\mathbf{b})$ and $T(k\mathbf{a}) = kT(\mathbf{a})$, 则称 T 是 V 到 W 的线性变换, 记作 $T \in \mathfrak{L}(V,W)$. 如果 $W = V$, 则写成 $T \in \mathfrak{L}(V)$, 此时称 T 为 V 上的线性算子.*

类比同构映射 (但是注意线性变换不一定是单射):

性质 3. *线性映射 T 满足:*

1 $T(\mathbf{0})=\mathbf{0}$

2 $T(\sum k_i\mathbf{a}_i)=\sum k_iT(\mathbf{a}_i)$

定义 15 (kernel, range)**.** 现有 $T\in\mathfrak{L}(V,W)$, 则称:

1 集合 V 为 T 的定义域

2 集合 $\{\mathbf{a}|T(\mathbf{a})=\mathbf{0},\mathbf{a}\in V\}$ 为 T 的零空间或核 $\ker T$

3 集合 $\{T(\mathbf{a})|\mathbf{a}\in V\}$ 为 T 的像空间或值域 $R(T)$ *or* $T(V)$

通过运算的封闭性容易证明 $\ker T$,$R(T)$ 都是线性空间.

定义 16. $\operatorname{rank}T=\dim R(T)$, $\text{nullity }T=\dim\ker T$.

定理 9. $T\in\mathfrak{L}(V,W)$, $\dim V=n$, 那么 $\text{nullity }T+\operatorname{rank}T=n$

证明. 思路: 取 $\ker T$ 的基 $\mathbf{e}_1,\dots,\mathbf{e}_m$, 扩充成为 V 的基 $\mathbf{e}_1,\dots,\mathbf{e}_n$. 再证明 $T(\mathbf{e}_{m+1})\dots T(\mathbf{e}_n)$ 是 $R(T)$ 的基, 那么 $\operatorname{rank}T=n-m$, 证明成功. 这一步要分别证明 $R(T)$ 由这些向量组合生成, 还要证明它们线性无关. □

定义 17 (映射的乘积)**.** $S:U\to V,K:V\to W$ 是两个映射, 定义乘积 $KS:U\to W$ 为 $KS(\mathbf{a})=K(S(\mathbf{a}))$.

定义 18 (映射的线性运算)**.** $K,S,T\in\mathfrak{L}(V,W),c\in\mathbb{F}$, 定义 $(K+S)(\mathbf{a})=K(\mathbf{a})+S(\mathbf{a})$, 以及 $(cK)(\mathbf{a})=c(K(\mathbf{a}))$.

定义 19 (映射的逆)**.** 设有映射 $S:V\to W$, 若存在一个映射 $K:W\to V$, 使得 $KS=I_V,SK=I_W$, 其中 I_V,I_W 是 V,W 上的恒等映射, 则称 S 为可逆映射, $K=S^{-1}$.

定理 10. 设 T,P 为线性变换, 则乘积 TP 为线性变换. 如果 T 可逆, T^{-1} 也是线性变换.

定理 11. $T\in\mathfrak{L}(V,W)$, 以下条件互相等价:

1 T 是单射

2 V 中线性无关组的像仍然为线性无关组

3 $\operatorname{rank}T=\dim V$

4 $\ker T=\{\mathbf{0}\}$

定理 12. $T\in\mathfrak{L}(V,W)$, $\dim V=\dim W=n$, 则以下条件等价:

1 T 是单射

2 T 可逆

3 T 是满射

4 $\operatorname{rank} T = n$

5 $\ker T = \{\mathbf{0}\}$

以下内容关于有限维空间中线性映射的表达:

定义 20 (线性变换的矩阵)**.** 设有 n, m 维线性空间 V, W, 基分别为 $B_V = \{\alpha_1, \dots, \alpha_n\}$ 和 $B_W = \{\beta_1, \dots, \beta_m\}$, 对于映射 $T \in \mathfrak{L}(V, W)$ 存在常矩阵 A 使得

$$\begin{bmatrix} T(\alpha_1) & \dots & T(\alpha_n) \end{bmatrix} = \begin{bmatrix} \beta_1 & \dots & \beta_m \end{bmatrix} A,$$

则称 $A_{m\times n}$ 为 T 在 B_V, B_W 下的矩阵. 当 $V = W$ 时简称方阵 A 为 T 在 B_V 下的矩阵.

若设 $\mathbf{x}, \mathbf{y}$ 分别为原像和像在 B_V, B_W 中的坐标, 则 $\mathbf{y} = A\mathbf{x}$. 根据线性变换矩阵的定义易见, 线性变换的乘积以及线性运算就对应于相应矩阵的乘积和线性运算.

定理 13. $\mathfrak{L}(V, W)$ 和 $\mathbb{F}^{m\times n}$ 同构, 维数为 $m \times n$.

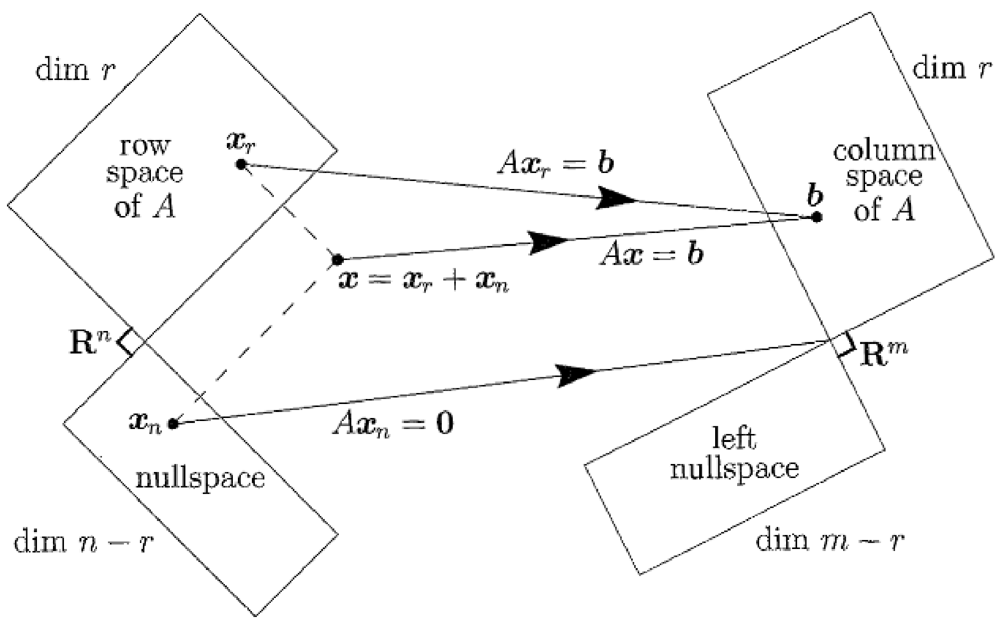

图 1.1: Four fundamental subspaces

性质 4 (4 fundamental subspaces)**.** 对于一个矩阵 $A_{m\times n}$, 有以下四个空间*:*

1 列空间 *column space* $C(A)$ 是由 A 的列向量张成的空间.

2 行空间 *row space* $C(A^{\mathrm{T}})$ 是由 A 的行向量张成的空间.

3 零空间 *nullspace* $N(A)$ 是由所有满足 $A\mathbf{x}=\mathbf{0}$ 的 $\mathbf{x}$ 组成的空间.

4 左零空间 *left nullspace* $N(A^{\mathrm{T}})$ 是由所有满足 $\mathbf{y}^{\mathrm{T}}A=\mathbf{0}^{\mathrm{T}}$ 的 $\mathbf{y}$ 组成的空间.

$C(A)\bigoplus N(A^{\mathrm{T}})=\mathbb{F}^m$, $C(A^{\mathrm{T}})\bigoplus N(A)=\mathbb{F}^n$.

对于映射 $T(\mathbf{x})=A\mathbf{x}$, 任何 $\mathbf{x}$ 可以唯一分解为 $\mathbf{x}_r+\mathbf{x}_n$, 其中 $\mathbf{x}_n\in N(A), \mathbf{x}_r\notin N(A)$. $\mathbf{x}_n$ 的部分会被映射成 $\mathbf{0}$, 另一部分有 $T(\mathbf{x})=T(\mathbf{x}_r)$.

定理 14. *$T\in\mathfrak{T}(V,W)$ 在某个基下的矩阵为 A, 那么*

1 $R(T)$ 与 $C(A)$ 同构. $\operatorname{rank}T=\operatorname{rank}A$.

2 $\ker(T)$ 与 $N(A)$ 同构. $\operatorname{nullity}T=n-\operatorname{rank}A$.

最后的重点就是同一个映射的矩阵在不同基下会如何变化.

定理 15 (基变换I)**.** *设 $T\in\mathfrak{L}(V)$, $B_a=\{\alpha_1,\ldots,\alpha_n\}$, $B_b=\{\beta_1,\ldots,\beta_n\}$. T 在 B_a,B_b 下的矩阵分别为 A,D. B_a 到 B_b 的过渡矩阵为 C, 则 $D=C^{-1}AC$. 线性算子在不同基下的矩阵是相似的.*

证明. 令 $P_a=\begin{bmatrix}\alpha_1 & \ldots & \alpha_n\end{bmatrix}$, $P_b=\begin{bmatrix}\beta_1 & \ldots & \beta_n\end{bmatrix}$. 在基 B_a 下, $\mathbf{x}$ 的坐标是 $P_a^{-1}\mathbf{x}$, 基 B_b 下为 $P_b^{-1}\mathbf{x}$. 将 $\mathbf{x}$ 的像用标准基 $\{\epsilon_1=(1,0,0,\ldots),\epsilon_2=(0,1,0,\ldots),\ldots,\epsilon_n\}$ 表示, 那么 $T(\mathbf{x})=P_aA(P_a^{-1}\mathbf{x})=P_bD(P_b^{-1}\mathbf{x})$. 因为 $P_b=P_aC\Rightarrow P_a^{-1}P_b=C\Rightarrow D=C^{-1}AC$. □

定理 16 (基变换 II)**.** *设 $T\in\mathfrak{L}(V,W)$. V 有基 B_a,B_b, W 有基 B_w. T 在 B_a,B_b 下的矩阵分别为 A,D. B_a 到 B_b 的过渡矩阵为 C. 则 $D=AC$.*

证明. 同上个定理. $T(\mathbf{x})=B_wA(P_a^{-1}\mathbf{x})=B_wD(P_b^{-1}\mathbf{x})\Rightarrow D=AC$. □

Chapter 2

向量组与线性方程组

线性代数是由线性方程组发展起来的, 线性方程组是线性代数的核心概念和根基, 它作为桥梁连接起矩阵与 n 维向量空间, 从而使得线性映射和线性空间这两大线性代数研究的主要内容联系起来, 因此研究线性方程组的解的理论与求解方法是线性代数的基本任务之一. 在前面章节行列式与矩阵等基本知识的基础之上, 本章的目标在于探讨: [3]

1) 线性方程组的求解方法;

2) 线性方程组解的情况与判别准则;

3) 线性方程组解的结构.

其中由于 n 元线性方程组的解是 n 元向量, 研究方程解的结构即解与解之间关系, 就是研究 n 元向量之间的关系, 因此本章中将引入 n 元向量理论, n 维向量空间也是下一章研究线性空间的基础. n 元向量理论中包含四大问题, 即:

1) 相关性问题 (研究一组向量之间的关系)

2) 表示性问题 (研究一个向量与一组向量之间的关系)

3) 代表性问题 (研究极大线性无关组)

4) 等价性问题 (研究一组向量与一组向量之间的关系)

这篇笔记将对以上问题进行逐一梳理. n 元向量理论中定理、推论、二级结论均繁多, n 维向量、线性方程组与矩阵三者之间的互化也需要认真理解, 是本章的难点所在.

2.1 线性方程组的解法与解的情况判定

2.1.1 线性方程组的表达形式

为了统一地研究线性方程组, 约定把常数项写在等号的右边, 含未知量的项写在等号左边. 由此可得一个 $m \times n$ 线性方程组的**一般形式**:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \qquad\qquad\qquad\qquad \cdots\cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n \end{cases}$$

利用矩阵相等和矩阵乘法的定义, 可将上述方程组简写为以下**矩阵方程形式**:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

或

$$Ax = b$$

由于线性方程组的增广矩阵可完全确定一个线性方程组, 因此我们可以将线性方程组中的未知量隐去, 以**增广矩阵形式**进行更为简捷地表达:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & b_n \end{bmatrix}$$

再回看第二种矩阵方程形式, 我们对系数矩阵 A 进行列分块, 可得

$$x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{bmatrix} + \cdots \quad + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{nn} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

或

$$x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n = b$$

这就是线性方程组的第四种表达形式——**向量形式**. 利用这种形式, 我们可将研究线性方程的问题与研究向量的问题相互转化.

2.1.2 线性方程组的解法

- $Gauss - Jordan$ 消元法

  求解线性方程组的基本思路是: 通过三种初等变换方法消去一些未知量 (消元) , 变成阶梯型方程组. 可以证明: 经过初等变换得到的方程组与原方程组同解, 从而经过一系列初等变换变成的阶梯形方程组与原方程同解.

  由于线性方程组与其增广矩阵一一对应, 因此线性方程组的初等变换也一一对应于增广矩阵的初等行变换. 第二章中我们学过, 任何一个矩阵都能经过一系列初等行变换化成阶梯形矩阵, 并且能用初等行变换进一步化成唯一确定的简化行阶梯形矩阵.

  于是应用 $Gauss - Jordan$ 消元法求解线性方程组的过程, 即为对增广矩阵进行系列初等行变换化为简化行阶梯形矩阵的过程, 这是一个上三角化的过程.

- 矩阵分解式求解法 (拓展)

  Gauss 消元的本质就是利用矩阵 L 左乘 A 得到上三角矩阵. 因此很自然地能够将 A 在消元法中分解为上、下三角矩阵. 在计算机求解线性方程组 $Ax = b$ 时, 求解 A^{-1} 再做乘法是复杂且没有必要的. 首先对系数矩阵 A 进行 LU 分解, 那么求解方程组 $Ax = b$ 只需依次求解 $Ly = b$ 和 $Ux = y$. 求解这两个方程组都较容易 (直接回代求解), 问题也就迎刃而解. 三角分解根据哪个矩阵的对角线元素均为 1 分为以下三种:

  1.$Doolittle$ 分解:$A = L\widetilde{U}$, 其中 L 为单位下三角阵 (对角线均为 1 的下三角阵) , $\widetilde{U}$ 为上三角阵.

  2.LDU 分解:$A = LDU$, 其中 L 为单位下三角阵, D 为对角阵, U 为单位上三角阵.

  3.$Crout$ 分解: $A = \widetilde{L}U$, 其中 $\widetilde{L}$ 为下三角阵, U 为单位上三角阵.

  特别地, 当 A 为对称矩阵时, $LDU = A = A^{\mathrm{T}} = U^{\mathrm{T}}DL^{\mathrm{T}}$, 根据阶梯矩阵的唯一性, $U = L^{\mathrm{T}}$, i.e.$A = LDL^{\mathrm{T}}$. 此外需要注意的是, LU 分解的应用是有条件的, 需满足 A 的各阶顺序主子式均不等于零, 这是 LU 分解存在且唯一的充要条件. 有唯一解的方程能够顺利进行消元, 它们均满足这个条件.

2.1.3 线性方程组解的情况及判别准则

定理 1. 系数和常数项为有理数的 n 元线性方程组的阶段情况只有三种可能: 无解, 有唯一解, 有无穷多个解. 且存在以下三组充要条件:

1) 无解 $\iff r(A) = r$ 而 $r(\overline{A}) = r + 1$

2) 有唯一解 $\iff r(A) = r(\overline{A}) = n$

3) 有无穷多解 $\iff r(A) = r(\overline{A}) < n$

(以上定理证明见书 p135-137. 如果用线性空间来理解, 那么无解 $\Leftrightarrow b \notin C(A)$.)

根据定理 1, 可得到针对于 n 元齐次线性方程组的两条推论:

推论 1.n 元齐次线性方程组有非零解 $\iff$ 它的系数矩阵经过初等行变换化成的阶梯形矩阵中, 非零行的数目 $r < n$.

推论 2. 对于 n 元齐次线性方程组 $Ax = 0$, 如果方程的数目 m 小于未知量的数目 n, 则它一定有非零解.

证 由于 $r(A_{m\times n}) \leq m$, 而 $m < n$, 故 $r(A_{m\times n}) < n$, 由推论 1 即知 n 元齐次线性方程组 $Ax = 0$ 必有非零解.

基于推论 2, 更一般地, 可得推论 3.

$\star$ **推论 3.** 当矩阵 $A_{m\times n}$ 的秩为 $r(A) = m < n$, 非齐次线性方程组 $Ax = b$ 一定有无穷多组解.

证 $r(A) = r(\overline{A}) = m < n$, 这是一个充要的结论.

2.2 向量组的线性相关性

2.2.1 数域上的 n 维向量及其线性运算

定义 1(数域). 若数的集合 $\mathbb{F}$ 满足:

(1) $0, 1 \in \mathbb{F}$;

(2) $a, b \in \mathbb{F} \implies a \pm b, ab \in \mathbb{F}$,

$a, b \in \mathbb{F}$ 且 $b \neq 0 \implies a/b \in \mathbb{F}$,

则称 $\mathbb{F}$ 是一个数域.

本章内容是在任一数域上进行讨论.

书上在此处引入了众多概念, 包括 n 维行列向量、零向量、负向量、n 维向量相等、n 维向量空间等等的定义, 笔者在此处略写, 若有概念不清可参考书 p143-144.

需要注意是: 维数不同的零向量虽然记号相同, 但是实则不同.

2.2.2 线性表示与等价向量组

定义 2(线性组合与线性表示). 设 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 是一组 n 维向量, $k_1, k_2, \cdots, k_s$ 是一组常数, 则称向量

$$k_1\alpha_1 + k_2\alpha_2 \quad + \cdots + k_s\alpha_s,$$

为 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 的一个线性组合. 若 β 可以表示为

$$\beta = k_1\alpha_1 \quad + k_2\alpha_2 + \cdots \quad + k_s\alpha_s,$$

则称向量 β 可由向量组 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 线性表示或线性表出.

由此定义可得到三条结论:

结论 1. 零向量是任意向量组的线性组合.

结论 2. 向量组中每一个向量都可由该向量组线性表示.

结论 3. 任一 n 维向量 $\alpha = [\alpha_1, \alpha_2, \ldots \alpha_n]^{\mathrm{T}}$ 都可由 n 维基本单位向量组唯一线表.(依据定义 2 即可完成证明)

此处应当指出, 这里所提到的 n **维基本单位向量组**有着特殊的性质, 即向量组的秩 $r = n$, 这使得它常常在后续秩相关证明题中, 作为"任一"n 维向量组中的"特殊"向量组取用, 以便于证明.

对比线表定义式与前面所提到的线性方程组的向量形式, 可知满足线表定义式的一组数 $k_1, k_2, \cdots, k_s$ 是线性方程组

$$x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n = \beta$$

的一个解. 由此可得一组重要的互推关系:

$\star$ **定理 2.** 向量 β 可由向量组 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 线性表示

$\Longleftrightarrow$ 线性方程组 $x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n = \beta$ 有解

$\iff r(\alpha_1,\alpha_2,\cdots,\alpha_s)=r(\alpha_1,\alpha_2,\cdots,\alpha_s,\beta)$

以上我们讨论了一个向量由一组向量线表的问题, 将其推广, 若向量组 (I) 中的每一个向量都可由向量组 (II) 线表, 则称向量组 (I) 可由向量组 (II) 线表.

特别地, 若与此同时向量组 (II) 也可由向量组 (I) 线表, 则称向量组 (I) 与向量组 (II) **等价**. 容易证明, 向量组的等价关系具有自反性、对称性和传递性. 那么我们如何判定两个向量组等价呢？总体而言, 有以下三种判定方法:

1) 利用**定义**——即证明两个向量组 (的极大无关组) 可互相线表;

2) 利用**定义和秩**——即证明 $r(A)=r(B)$ 且向量组 (I) 可由向量组 (II) 线表;

3) 纯利用**秩**——即证明 $r(A)=r(B)=r(A,B)$.

对方法二、三的正确性进行如下证明:

证 (方法二)　设 $r(I)=r(II)=r;(I)$ 的极大无关组为 $\alpha_1,\alpha_2,\cdots,\alpha_r;(II)$ 的极大无关组为 $\beta_1,\beta_2,\cdots,\beta_r$. (I) 可由 (II) 线表, 则 (I) 的极大无关组可由 (II) 的极大无关组线表. 另外易知向量组 (II) 中任一向量 β_i 必可由 (II) 的极大无关组线表, 所以向量组 $\alpha_1,\alpha_2,\cdots,\alpha_r,\beta_i$ 必可由 (II) 的极大无关组线表, 而该向量组中向量个数为 $(r+1)>r$, 由后文中我们将证明的定理可知, $\alpha_1,\alpha_2,\cdots,\alpha_r,\beta_i$ 线性相关, 又因为 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 线性无关, 所以 β_i 必可由 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 线表, 由 β_i 的任意性可知 $\beta_1,\beta_2,\cdots,\beta_r$ 可由 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 线表, 因此 (II) 可由 (I) 线表. □

证 (方法三)　$r(I)=r(I,II) \implies (II)$ 可由 (I) 线表;

$r(II)=r(I,II) \implies (I)$ 可由 (II) 线表;

由此可推出 (I) 与 (II) 等价. □

此外, 应当指出, 等价的概念在此前矩阵等价时也曾提出, 而矩阵又可看成是由列向量组或行向量组构成的, 但矩阵的等价与向量组的等价有着诸多不同点, 这里我们来谈谈它们的区别与联系.

首先由定义, 两个等价的矩阵必须是同型矩阵, 同型矩阵所含列 (行) 向量的个数是相同的 (有限个) ; 两个等价的向量组必须是同维数的向量组但它们所含向量的个数可以不同 (可以有无穷个) .eg. 向量组 $A:\alpha=(0,1)$ (只含一个向量) 与向量组 $B:[\beta=k\alpha|k\in R]$ (有无穷多个向量) 等价.

两个等价又有着怎样的联系呢？谈论联系, 只有在向量组含有有限个向量时才有意义, 此时不妨将向量组也作成矩阵. 对矩阵 A 做初等行 (列) 变换化为矩阵 B, 则 A 与 B 行

(列) 等价, 它们的行 (列) 向量组也等价; 但若 A 化为 C 既有行变换又有列变换, 那么 A 与 C 的行向量组与列向量组就都未必等价了.

2.2.3 线性相关与线性无关

定义 3 (线性相关与线性无关) . 设 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 是一组 n 维向量, 如果存在一组不全为零的常数 $k_1, k_2, \cdots, k_s$, $s.t.$

$$k_1\alpha_1 + k_2\alpha_2 \quad + \cdots + k_s\alpha_s = 0$$

则称向量组 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 是线性相关的. 反之, 若上式仅在 $k_1 = k_2 = \cdots = k_s = 0$ 时才成立, 则称向量组 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 是线性无关.

注意, 根据线性无关的向量组的定义去判断一个向量组线性无关的方法, 是最基本、最重要的方法.

由以上定义, 我们容易得出两种特殊的情况:

(1) 仅含一个向量的向量组线性相关 $\iff$ 该向量为零向量;

(2) 仅含两个向量的向量组线性相关 $\iff$ 这两个向量的对应分量成比例.

我们可以从以下几个角度来思考向量组线性相关与线性无关的本质区别: 线性组合、线性表出、齐次线性方程组、行列式、向量组的秩、向量组与其部分组的关系.

此处笔者略写, 供读者自行思考总结. 读者还可自己提出一些与向量组相关性有关的问题, 并尝试通过以上几个角度进行转化, 进而进行解答. 比如, 从齐次线性方程组的角度, 可以证明“初等行变换不改变矩阵任意列向量之间的相关性”这一结论.

基于以上思考, 可以得到两个特殊的、重要的推论:

推论 4.n 个 n 维向量 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 线性相关 (线性无关) $\iff$ $\det[\alpha_1, \alpha_2, \cdots, \alpha_n] = 0 (\neq 0)$.

$\star$ **推论 5.** 若 $s > n$, 则 s 个 n 维向量 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 必线性相关. 特别地, $n+1$ 个 n 维向量必线性相关.

证明思路 从秩的角度证明, 将向量组写做矩阵, 行秩等于列秩, 所以必列降秩, 列向量组线性相关; 也可从线性方程组的角度理解: 方程个数小于未知量个数的齐次线性方程组必有非零解.

对于向量组的线性相关性, 有如下一些重要定理:

⋆ **定理 3.** 向量组 $\alpha_1, \alpha_2, \cdots, \alpha_s (s \geq 2)$ 线性相关 $\Longleftrightarrow$ 该组中至少存在 1 个向量可由该组中其余 $s-1$ 个向量线表.

⋆ **定理 4.** 若向量组 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 线性无关, 而向量组 $\alpha_1, \alpha_2, \cdots, \alpha_s, \beta$ 线性相关, 则 β 可由向量组 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 线表, 且表示法唯一.

证明思路 用定义证线表, 反证法证唯一性.

定理 5. 如果向量组 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 有一个部分组线性相关, 则向量组 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 线性相关. 即部分相关则整体相关, 整体无关则任一部分无关.

另外, 由于零向量是线性相关的, 所以容易得出: 线性无关向量组必不含零向量.

2.3 向量组的秩

2.3.1 向量组的极大无关组与向量组的秩

定义 4 (极大无关组) . 如果向量组 U 的一个部分组 $\alpha_1, \alpha_2, \cdots, \alpha_r$ 满足:

(1) $\alpha_1, \alpha_2, \cdots, \alpha_r$ 线性无关;

(2) U 中任意向量 α 都可由向量组 $\alpha_1, \alpha_2, \cdots, \alpha_r$ 线表, 即从这个向量组的其余向量 (如果还有的话) 中任取一个添进去, 得到的新的部分组都线性相关.

则称 $\alpha_1, \alpha_2, \cdots, \alpha_r$ 为向量组 U 的一个极大 (线性) 无关组.

向量组的任意两个极大无关组所含向量的个数是否相等？我们通过以下几个定理和推论加以讨论:

定理 6. 向量组与它的极大无关组等价. (用互相线表证)

推论 6. 向量组的任意两个极大无关组等价. (由定理 6 和等价的对称性、传递性证)

推论 7.β 可以由向量组 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 的线表, 当且仅当 β 可以由向量组 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 的一个极大无关组线表. (用线表传递性证)

定理 7. 设向量组 $\beta_1, \beta_2, \cdots, \beta_r$ 可以由向量组 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 线表, 若 $r > s$, 则 $\beta_1, \beta_2, \cdots, \beta_r$ 线性相关. (用线性相关定义和推论 5 证)

推论 8. 设向量组 $\beta_1, \beta_2, \cdots, \beta_r$ 可以由向量组 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 线表, 若 $\beta_1, \beta_2, \cdots, \beta_r$ 线性无关, 则 $r \leq s$. (定理 7 的逆否命题)

推论 9. 等价的线性无关的向量组所含向量的个数相等. (由推论 8 证出)

推论 10. 向量组的任意两个极大无关组所含向量的个数相等. (由推论 6 和推论 9 证)

由推论 10, 引出下述重要概念:

定义 5 (向量组的秩) . 如果向量组 U 仅含零向量, 规定 U 的秩为零; 否则, 称 U 的极大无关组所含向量的个数为向量组 U 的秩, U 的秩记为 $r(U)$.

由向量组的秩和极大无关组的定义立刻可得出:

$\star$ **定理 8.** 向量组 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 线性无关 $\iff$ 它的 (行/列) 秩等于它所含 (行/列) 向量的个数.

仅凭一个自然数——向量组的秩就能判断向量组的线性相关性, 秩概念的重要性可见一斑!

定理 9. 若向量组 (I) 可由向量组 (II) 线表, 则 $r(I) \leq r(II)$.

证明思路 有关向量组的秩的证明题通常都要取向量组的一个极大无关组.

定理 10. 等价的向量组有相等的秩.

注意: 秩相等的两个向量组不一定等价. 反例,

$$\alpha_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \alpha_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}; \beta_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \beta_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}.$$

向量组等价的充要条件详见前文等价向量组的三种判定方法.

定理 11 (向量组的秩与矩阵秩的关系) .

$$\forall A, r(A) = A \text{ 的列秩} = A \text{ 的行秩}$$

此定理证明详见书 pp157-158. 此处编者想补充一点: 其实通过 CR 分解, 是理解定理 11 最直观的方法.

[**CR 分解**] C 由 A 的线性无关列组成 (即选择 R 中首非零元对应的 A 的列), 显然 C 与 A 有相同列空间. R 为 A 的简化行阶梯形矩阵 (并删去了零行), 各行线性无关. 而根据矩阵左乘的意义, A 的行是 R 中行的线性组合 $\Longrightarrow$ A 与 R 有相同的行空间.

于是, C 的列数为 A 的列秩, R 的行数等于 A 的行秩. 显然 C 的列数 $= R$ 的行数 $\Longrightarrow \forall A, A$ 的列秩 $= A$ 的行秩.

2.3.2 关于向量组的秩与矩阵的秩的重要结论及重要不等式

结论 1. 设矩阵 $A_{m\times n}, B_{n\times p}$ 满足 $AB=O$

$\Longrightarrow$ 1) 若 A 的列向量组线性无关, 则 $B=O$;

证明提示 将 B 列分块.

$\Longrightarrow$ 2) 若 B 的行向量组线性无关, 则 $A=O$;

证明提示 $AB=O \iff (AB)^{\mathrm{T}}=B^{\mathrm{T}}A^{\mathrm{T}}=0$, 再将 A^{T} 列分块.

$\Longrightarrow$ 3) 若 $B\neq 0$, 则 A 的列向量组线性相关; (结论 1 的逆否)

$\Longrightarrow$ 4) 若 $A\neq 0$, 则 B 的行向量组线性相关; (结论 2 的逆否)

$\star$ $\Longrightarrow$ 5)$r(A)+r(B)\leq n$. (详细证明见书 p169)

$\star$ **结论 2 ($r(A)$ 与 $r(A^*)$ 的关系) .**

$$r(A^*)=\begin{cases} n, r(A)=n \\ 1, r(A)=n-1 \\ 0, r(A)\leq n-2 \end{cases}$$

证明提示 考虑 A 与 A^* 关系的重要恒等式; 结论 1; A^* 的元素构成; 代数余子式 A_{ij} 本质上也是一个 $n-1$ 阶子式.

$\star$ **结论 3 (秩相关的重要等式与不等式) .**

1) $r(A)=r(kA)$, 其中 $k\neq 0$;

证 A 的每一行做倍乘初等变换, 秩不变.

2)$r(A+B)\leq r(A)+r(B)$;

3)$r(AB)\leq \min(r(A), r(B))$

(以上非常重要的结论 3 证明见书 p161-162)

$\star$ **结论 4 (证明秩相等的新方法) .** 若齐次线性方程组 $Ax=0$ 与 $Bx=0$ 同解 $\Longrightarrow$ $r(A)=r(B)$ (通过基础解系中所含向量个数来证)

$\star$ **结论 5.** 设矩阵 $A_{m\times n}, B_{n\times p}$, 且 $r(A)=n$, 则有 $r(AB)=r(B)$. 即矩阵左乘一个列满秩矩阵, 其秩不变.

证 法 1: (用结论 4, 通过齐次线性方程组同解来证秩相等)

若 x 是 $Bx=0$ 的解, 即 $Bx=0$, 两端左乘 A, 有 $ABx=0$. 若 x 是 $ABx=0$ 的解, 则 $y=Bx$ 是 $Ay=0$ 的解, 又因为 A 列满秩 $\implies$ $Ay=0$ 只有零解, 所以 $Bx=0$. 因此方程 $Bx=0$ 与 $ABx=0$ 同解, 故有 $r(A)=r(B)$ $\square$

法 2: (从初等变换不改变秩的角度证)

因 $r(A)=n$, 知 A 的行最简形矩阵即为 A 的秩标准形为 $\begin{bmatrix} I_n \\ O \end{bmatrix}_{m\times n}$, 并有 m 阶可逆矩阵 $P, s.t. PA=\begin{bmatrix} I_n \\ O \end{bmatrix}_{m\times n} \implies PAB=\begin{bmatrix} I_n \\ O \end{bmatrix} B=\begin{bmatrix} B \\ O \end{bmatrix}_{m\times n}$. 故 $r(AB)=r(PAB)=r\begin{bmatrix} B \\ O \end{bmatrix}=r(B)$ $\square$

2.4 线性方程组的解的结构

2.4.1 齐次线性方程组

若齐次线性方程组存在非零解, 则易证它的解满足如下三条性质:

性质 1. 若 $x=\xi_1$ 及 $x=\xi_2$ 都是齐次线性方程组 $Ax=0$ 的解, 则 $x=\xi_1+\xi_2$ 也是方程组 $Ax=0$ 的解.

性质 2. 若 $x=\xi_1$ 是齐次线性方程组 $Ax=0$ 的解, k 为任意常数, 则 $x=k\xi_1$ 也是方程组 $Ax=0$ 的解.

性质 3. 若方程组 $Ax=0$ 有非零解, 则必有无穷多个非零解.

类比于 n 维向量理论中“极大无关组”的概念, 线性方程组的解空间中也存在这样的极大无关组, 于是有如下定义:

定义 6 (基础解系) . 若齐次线性方程组 $Ax=0$ 的一组解向量 $\xi_1,\xi_2,\cdots,\xi_t$, 满足:

(1) $\xi_1,\xi_2,\cdots,\xi_t$ 线性无关;

(2) 方程组 $Ax=0$ 的任一解都可由 $\xi_1,\xi_2,\cdots,\xi_t$ 线表;

则称 $\xi_1,\xi_2,\cdots,\xi_t$ 是方程组 $Ax=0$ 的一个基础解系.

$\star$ **定理 12** 设 A 为 $m\times n$ 矩阵, 若 $r(A)=r<n$, 则 n 元齐次线性方程组 $Ax=0$ 必存在基础解系, 且基础解系含 $n-r$ 个解向量. (详细证明见书 p165-166)

同样类比于极大无关组, 可得:

推论 11. 设 A 为 $m\times n$ 矩阵, $r(A)=r<n$, 则 n 元齐次线性方程组 $Ax=0$ 的任何 $n-r$ 个线性无关的解向量都可作为 $Ax=0$ 的基础解系.

应当指出, 基于定理 12 和推论 11, 我们可以得到判定基础解系的另一个角度: 向量组中所含有向量的个数. 那么如何求基础解系呢？下面给出求解的一般方法:

首先求出自由未知量表示的通解, 此后有两种基本解法. 其一是在通解中令 $n-r$ 个自由未知量分别取 $n-r$ 组值, 使得相应得到的 $n-r$ 个解向量线性无关, 这组解向量就构成了方程的基础解系; 其二是将由自由未知量表示的通解改写成参数向量形式, 也可相应地得到基础解系. 两种作法的结果是一样的. 基于上述性质与定理, 可以推出齐次线性方程组的结构解的表达式:

定理 13 (齐次线性方程组的结构解) 若 $\xi_1,\xi_2,\cdots,\xi_t$ 是方程组 $Ax=0$ 的一个基础解系, 则方程组 $Ax=0$ 的结构解为 $x=c_1\xi_1+c_2\xi_2+\cdots+c_t\xi_t$ (其中 $c_1,c_2,\cdots,c_t$ 为任意常数) .

2.4.2 非齐次线性方程组

若非齐次线性方程组有无穷多解, 则易证它的解满足如下两条性质:

性质 4. 若 $x=\eta_1$ 及 $x=\eta_2$ 都是非齐次线性方程组 $Ax=b$ 的解, 则 $\eta_1-\eta_2$ 也是方程组 $Ax=b$ 的解.

性质 5. 若 $x=\eta$ 是非齐次线性方程组 $Ax=b$ 的解,ξ 是对应齐次线性方程组 $Ax=0$ 的解, 则 $\eta+\xi$ 是 $Ax=b$ 的解.

基于上述性质, 可以推出非齐次线性方程组的结构解的表达式:

定理 14 (非齐次线性方程组的结构解) . 若 $\xi_1,\xi_2,\cdots,\xi_t$ 是方程组 $Ax=0$ 的一个基础解系, η 是非齐次线性方程组 $Ax=b$ 的一个特解, 则方程组 $Ax=b$ 的结构解为 $x=\eta+c_1\xi_1+c_2\xi_2+\cdots+c_t\xi_t$ (其中 $c_1,c_2,\cdots,c_t$ 为任意常数) .

我们可以发现一个有意思的事实, 线性方程组虽然可能存在无穷多组解, 但这些解总可以通过有限个向量的线性组合来表达出.

Chapter 3

特征值与特征向量

3.1 特征值和特征向量

特征值可以帮助求解常微分方程组, 求矩阵的幂与指数函数, 求一些矩阵分解等, 是理解许多矩阵应用的关键所在. 本章内的所有矩阵若无特殊说明均是复数方阵.

定义 1 (特征值, 特征向量)**.** 对于方阵 A, 若存在复数 λ 及非零向量 $\mathbf{x}$, 使得 $A\mathbf{x} = \lambda\mathbf{x}$, *i.e.* $(A - \lambda I)\mathbf{x} = 0$, 则称 λ 为 A 的一个特征值, $\mathbf{x}$ 为 λ 对应的特征向量.

通过给 A 右乘特征向量 $\mathbf{x}$ 并不断利用 $A\mathbf{x} = \lambda\mathbf{x}$, 易证

性质 1. 若 λ 和 $\mathbf{x}$ 是 A 的特征值和特征向量, 则 λ^n 和 $\mathbf{x}$ 是 A^n 的特征值和特征向量. $(n \in \mathbb{N})$ 特别地, 如果 A 可逆, 那么 $n = -1$ 也成立.

另外, 实系数多项式方程 $\det(A - \lambda I) = 0$ 的虚数根一定是两两共轭的, 于是有

性质 2. 若 λ 和 $\mathbf{x}$ 是 A 的特征值和特征向量, 则它们也是关于 A 的多项式 $P(A)$ 的特征值和特征向量.

性质 3 (实矩阵的特征值共轭)**.** 若 λ 是实矩阵 A 的特征值, 则 $\overline{\lambda}$ 也是其特征值.

矩阵对角化需要每一特征值都能找到对应的特征向量. 为了判断重特征值的性质, 我们引入以下概念:

定义 2 (特征多项式, 代数重数)**.** 关于 λ 的多项式 $\det(A - \lambda I)$ 是 A 的特征多项式. 若特征多项式在 $\lambda = \lambda_r$ 时有 k 重根, 则称 λ_r 的代数重数为 k.

定义 3 (几何重数)**.** 几何重数等于 $\dim N(A - \lambda I)$, *i.e.* $A - \lambda I$ 解空间的维数.

可见, 代数重数决定相同特征值的个数, 几何重数代表该特征值对应的线性无关的特征向量个数. 以下有两个为相似对角化作准备的定理:

定理 1. *属于互不相同的一组特征值 $\lambda_1 \cdots \lambda_m$ 的特征向量 $\mathbf{x}_1 \cdots \mathbf{x}_m$ 线性无关.*

证明. 设有一组数 $k_1 \cdots k_m$ 使得 $\sum_{i=1}^m k_i \mathbf{x}_i = 0$. $A\mathbf{x}_i = \lambda_i \mathbf{x}_i \Rightarrow \sum_{i=1}^m \lambda_i k_i \mathbf{x}_i = 0$.

左右同时左乘 A 得到

$$\sum_{i=1}^m \lambda_i k_i (A\mathbf{x}_i) = \sum_{i=1}^m \lambda_i^2 k_i \mathbf{x}_i = 0$$

反复左乘 A $(m-1)$ 次, 得到:

$$\begin{bmatrix} k_1\mathbf{x}_1 & k_2\mathbf{x}_2 & \cdots & k_m\mathbf{x}_m \end{bmatrix} \begin{bmatrix} 1 & \lambda_1 & \lambda_1^2 & \cdots & \lambda_1^{m-1} \\ 1 & \lambda_2 & \lambda_2^2 & \cdots & \lambda_2^{m-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_m & \lambda_m^2 & \cdots & \lambda_m^{m-1} \end{bmatrix} = O$$

因为 λ_i 互不相同, Vandermonde 矩阵可逆,

$$\begin{bmatrix} k_1\mathbf{x}_1 & k_2\mathbf{x}_2 & \cdots & k_m\mathbf{x}_m \end{bmatrix} = O \Rightarrow k_i = 0$$

故这些特征向量线性无关. □

定理 2. *几何重数 $\leq$ 代数重数.*

证明. 设 V_0 为 $n \times n$ 矩阵 A 的特征值 λ_0 的特征空间. $\dim V_0 = k$. V_0 有标准正交基 $\alpha_1, \cdots, \alpha_k$, 现将其扩充成 $\mathbb{C}^n$ 的标准正交基 $\alpha_1, \cdots \alpha_n$.

$$\begin{aligned} AP = A\begin{bmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_n \end{bmatrix} &= \begin{bmatrix} \lambda_0\alpha_1 & \cdots & \lambda_0\alpha_k & A\alpha_{k+1} & \cdots & A\alpha_n \end{bmatrix} \\ &= P\begin{bmatrix} \lambda_0 I_{k\times k} & B_{k\times(n-k)} \\ O_{(n-k)\times k} & D_{(n-k)\times(n-k)} \end{bmatrix} \end{aligned}$$

$$B = P^{-1}AP,\ \det(B-\lambda I) = (\lambda_0-\lambda)^k \det D$$

显然, λ_0 至少是 B 的 k 重特征值, 由于 A 与 B 相似, 特征值相同, 故 A 的特征值 λ_0 的代数重数 $\geq k =$ 几何重数. □

3.2 相似对角化

定义 4 (矩阵相似)**.** 如果存在可逆矩阵 P, 使得 $B = P^{-1}AP$, 则称 A 和 B 相似, 变换 $P^{-1}AP$ 为相似变换.

性质 4. 若矩阵 A, B 相似, 则 $\det A = \det B$, $\operatorname{rank} A = \operatorname{rank} B$.

"相似"变换中保持不变的量是矩阵的特征值:

定理 3. 相似矩阵有相同的特征多项式, 从而也有相同的特征值.

证明. 设 $B = P^{-1}AP$,

$$\begin{aligned}\det(B-\lambda I) &= \det\left(P^{-1}AP - P^{-1}\lambda P\right)\\ &= \det\left(P^{-1}\right)\det(A-\lambda I)det(P)\\ &= \det(A-\lambda I)\end{aligned}$$

□

结合之前的证明过程, 根据 $AP = PD$ 的形式, 不难发现: 只要能找到由特征向量组成的可逆的 P, 矩阵就能完成对角化. 因此矩阵对角化的条件是:

定理 4 (对角化充要条件)**.** $n \times n$ 矩阵 A 有 n 个线性无关的特征向量.

证明.

$$\begin{aligned}\exists P = [\mathbf{x}_1 \cdots \mathbf{x}_n]\ (\det P \neq 0),\quad & P^{-1}AP = \operatorname{diag}(\lambda_1, \cdots, \lambda_n)\\ \Leftrightarrow \quad A[\mathbf{x}_1 \cdots \mathbf{x}_n] =& [\mathbf{x}_1 \cdots \mathbf{x}_n]\operatorname{diag}(\lambda_1, \cdots, \lambda_n)\\ \Leftrightarrow \quad [A\mathbf{x}_1 \cdots A\mathbf{x}_n] =& [\lambda_1\mathbf{x}_1 \cdots \lambda_n\mathbf{x}_n] \Leftrightarrow \quad A\mathbf{x}_i = \lambda_i\mathbf{x}_i\end{aligned}$$

i.e. 非零向量 $\mathbf{x}_i$ 是特征向量, λ_i 为特征值. $\det P \neq 0$ 说明这些特征向量线性无关, 于是, 可对角化的矩阵必有 n 个线性无关的特征向量. 从下到上倒推, 可证明 A 若有 n 个线性无关的特征向量, 则可以组成可逆矩阵 P, 因而 A 可以对角化. □

显然, 如果一个矩阵的特征值都是单特征值, 则每个特征值对应的特征向量线性无关, 因此这样的矩阵必然可以对角化.

定理 5 (对角化充要条件)**.** 矩阵 A 每个特征值的几何重数 = 代数重数.

综上, 一般矩阵对角化的步骤为

1 求出 A 的所有特征值 $\lambda_1 \cdots \lambda_n$.

2 对于单特征值 λ_i, 求其对应的特征向量 ξ_i.

3 对于 k 重特征值 $\lambda_j \cdots \lambda_{j+k-1}$, 求其特征空间的一个基 $\xi_j \cdots \xi_{j+k-1}$. 因为特征空间维数等于代数重数, 故基中同样有 k 个向量.

4 $P = \begin{bmatrix} \xi_1 & \xi_2 & \cdots & \xi_n \end{bmatrix}$ 可逆, 且 $P^{-1}AP = \mathrm{diag}(\lambda_1, \cdots, \lambda_n)$

3.3 特殊矩阵

3.3.1 Hermitian matrices

前面已经提及, 本章内容可以用于复矩阵. 现在考虑一个 $\mathbb{C}^2$ 向量 $\mathbf{x} = (1, i)^{\mathrm{T}}$, 当我们使用 $\mathbf{x}^{\mathrm{T}}\mathbf{x}$ 来计算其模长时会发现结果为 0. 我们知道一个复数乘以自己的共轭等于其模长平方. 因此在复数域上, 类比转置, 我们定义矩阵的另一个运算

定义 5 (Hermitian matrix)**.** $(A^{\mathrm{H}})_{ij} = \overline{A_{ji}}$. *如果 $A = A^{\mathrm{H}}$ 则称其为 Hermitian 矩阵.*

可以发现, Hermitian 运算对于实矩阵就是转置. 因此实对称矩阵也是 Hermitian 矩阵.

性质 5. *如果 $A = A^{\mathrm{H}}$, 则 A 的所有特征值都是实数.*

证明. 设有 $A = A^{\mathrm{H}}$ 并且 $A\mathbf{x} = \lambda\mathbf{x}$. 左右同取共轭并转置得到 $\mathbf{x}^{\mathrm{H}}A^{\mathrm{H}} = \overline{\lambda}\mathbf{x}^{\mathrm{H}}$

$$\overline{\lambda}\mathbf{x}^{\mathrm{H}}\mathbf{x} = (A\mathbf{x})^{\mathrm{H}}\mathbf{x} = \mathbf{x}^{\mathrm{H}}(A\mathbf{x}) = \lambda\mathbf{x}^{\mathrm{H}}\mathbf{x}$$

$$\Rightarrow (\lambda - \overline{\lambda})\mathbf{x}^{\mathrm{H}}\mathbf{x} = (\lambda - \overline{\lambda})\|\mathbf{x}\|^2 = 0$$

因为 $\mathbf{x}$ 不是零向量, 故 $\lambda = \overline{\lambda}$, 为实数. □

性质 6. *Hermitian 矩阵不同特征值对应的特征向量正交.*

证明. 已知 $A = A^{\mathrm{H}}$, $A\mathbf{x}_1 = \lambda_1\mathbf{x}_1, A\mathbf{x}_2 = \lambda_2\mathbf{x}_2$. $\lambda_{1,2} \in \mathbb{R}$.

$$\lambda_2\mathbf{x}_2^{\mathrm{H}}\mathbf{x}_1 = (A\mathbf{x}_2)^{\mathrm{H}}\mathbf{x}_1 = \mathbf{x}_2^{\mathrm{H}}(A\mathbf{x}_1) = \lambda_1\mathbf{x}_2^{\mathrm{H}}\mathbf{x}_1$$

因为 $\lambda_1 \neq \lambda_2$, 所以 $\mathbf{x}_2^{\mathrm{H}}\mathbf{x}_1 = 0$, $\mathbf{x}_1, \mathbf{x}_2$ 正交. □

定理 6. *若 A 是实对称矩阵, 则必然存在一个正交矩阵 P, 使得 $P^{-1}AP$ 为对角矩阵.*

以上定理证明见书 pp239-241. 具体计算时, 单特征值取单位特征向量, 重特征值时利用 Gram-Schmidt 正交化求特征空间的一个标准正交基, 用这些向量组成正交矩阵 P.

3.3.2 Unitary matrices

正交矩阵是满足 $P^{\mathrm{T}}P = I$ 的矩阵. 当矩阵扩展到复数, 转置扩展为 Hermitian 时, 很自然地产生了以下概念

定义 6 (Unitary matrix)**.** *若一个复矩阵满足 $U^{\mathrm{H}}U = I$, 则称其为酉矩阵.*

实数的正交矩阵显然也是酉矩阵. 和正交矩阵的推导类似, 酉矩阵的特征值满足:

性质 7. *酉矩阵的特征值 λ 满足 $\|\lambda\| = 1$.*

证明. 假设有 $U\mathbf{x} = \lambda\mathbf{x}$, 则 $\mathbf{x}^{\mathrm{H}}U^{\mathrm{H}} = \overline{\lambda}\mathbf{x}^{\mathrm{H}}$. $\mathbf{x}^{\mathrm{H}}U^{\mathrm{H}}U\mathbf{x} = \overline{\lambda}\lambda\mathbf{x}^{\mathrm{H}}\mathbf{x} \Rightarrow \|\mathbf{x}\|^2 = \|\lambda\|^2\|\mathbf{x}\|^2$ 因为 $\|\mathbf{x}\| \neq 0$, 所以 $\|\lambda\| = 1$. □

于是, 实正交矩阵的实特征值一定是 1 或 -1. 并且当行列式为 -1 时, $\det = \prod_{i=1}^{n} \lambda_i = -1$, 其中共轭的虚特征值两两乘积均为 1, 剩下的实特征值必定含有 -1.

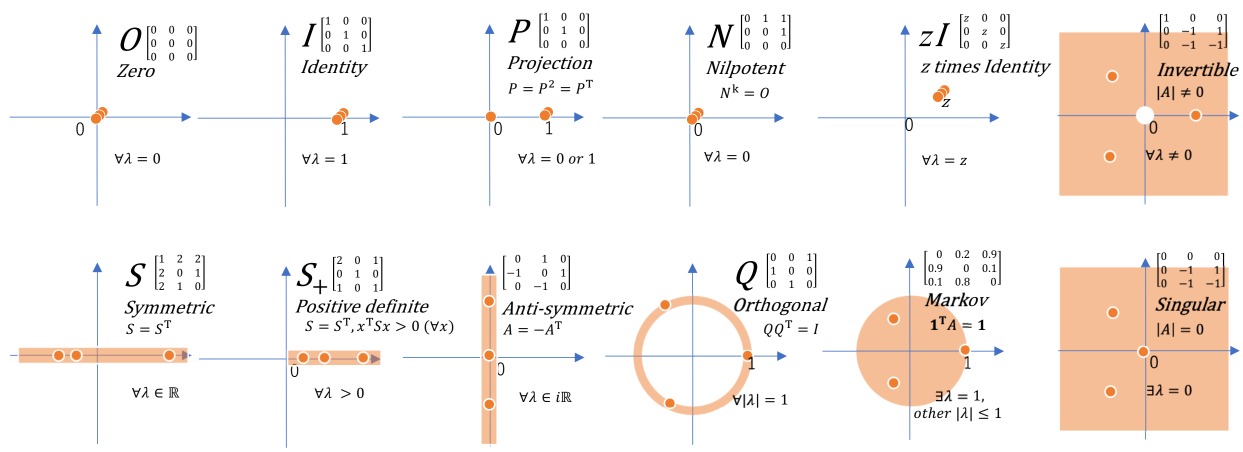

图 3.1: 特殊矩阵的特征值分布 [1]

3.4 特征值与微分, 差分方程

3.4.1 Markov: 离散模型

对于人口迁移, 存款利润等问题, 一般会使用离散模型, 用上一年 (或几年, 几个月) 的数据预测下一刻的情况. 例如经典的 Markov 问题: 城市中每年有 6% 的市区人口迁至郊区, 2% 的郊区人口迁入市区. 给定初始城区, 郊区居民占比分别为 30%, 70%, 设总人口数量不变, 试推算人口分布随时间的变化.

设市区, 郊区人口占比分别为 p_u, p_c, $\mathbf{x} = (p_u, p_c)^{\mathrm{T}}$.

$$\mathbf{x}_{n+1} = A\mathbf{x}_n, \quad \text{这个例子中} A = \begin{bmatrix} 0.94 & 0.02 \\ 0.06 & 0.98 \end{bmatrix}$$

根据递推关系, $\mathbf{x}$ 的通项为 $\mathbf{x}_n = A^n\mathbf{x}_0 = S\Lambda S^{-1}\mathbf{x}_0$. 记 $\mathbf{c} = S^{-1}\mathbf{x}_0$, $\mathbf{x}_n = \sum_{i=1}^n c_i\lambda_i^n\mathbf{x}_i$. $\mathbf{x}$ 的稳定 (i.e.$\mathbf{x}_\infty$ 是否存在) 取决于特征值的取值:

1. $\forall|\lambda_i| < 1$: 稳定.
2. $\exists|\lambda_j| = 1, \text{other} < 1$: 临界状态.
3. $\exists|\lambda_j| > 1$: 不稳定, 趋向于无穷大或反复震荡.

因为 Markov 矩阵的列和为 1, 所以 $(A^{\mathrm{T}} - I)(1, 1, \cdots, 1)^{\mathrm{T}} = \mathbf{0}$. 另外根据 Markov 链的实际意义, **Markov 矩阵的特征值 $|\lambda| \le 1$, 而且 $\exists\lambda_1 = 1$.** 随着 $n \to \infty$, $\mathbf{x}_\infty$ 趋向于稳定态 ξ_1, 即特征值 1 对应的特征向量. (注意: 当存在特征值 -1 时, 不存在稳定态. 一般应用中不会出现这种情况.)

3.4.2 ODE System: 连续模型

解常微分方程中常会遇到关于多个变量 $x(t), y(t), z(t)\ldots$ 以及它们导数的方程组, 这里设向量 $\mathbf{u} = (x, y, z\ldots)^{\mathrm{T}}$. 当然许多时候对一元高阶的线性方程, 我们也会用 $\mathbf{u} = (y, y\prime, y\prime\prime\ldots)^{\mathrm{T}}$ 的方法将其化成一阶方程来处理.

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u} = A\mathbf{u}$$

其中 A 为常矩阵. 我们知道, 这样一个方程的解形如

$$\mathbf{u}(t) = \sum_{i=1}^n c_i e^{\lambda_i t}\mathbf{x}_i = \begin{bmatrix} \mathbf{x}_1 & \cdots & \mathbf{x}_n \end{bmatrix} \begin{bmatrix} e^{\lambda_1 t} & & \\ & \ddots & \\ & & e^{\lambda_n t} \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix},$$

如果这里 $n = 1$, 那么结果就是 $u(t) = e^{at}u(0)$, 以此类推, 我们希望让结果呈现出如 $\mathbf{u}(t) = e^{At}\mathbf{u}(0)$ 般优雅美丽的形态. 于是我们希望定义矩阵的 exp 函数. 对于对角矩阵 $D = \mathrm{diag}(d_1, \cdots, d_n)$, $e^D = \mathrm{diag}(e^{d_1}, \cdots, e^{d_n})$. 那么上式中间的对角矩阵就可以写为 $e^{\Lambda t}$. 对于一般的矩阵, 我们定义 [2]

定义 7 (Exp of matrix)**.** $e^M = I + M + \frac{1}{2!}M^2 + \frac{1}{3!}M^3 + \cdots = I + \sum_{r=1}^\infty \frac{1}{r!}A^r$

性质 8. 对于可相似对角化的矩阵 $A = S\Lambda S^{-1}$, $e^A = Se^\Lambda S^{-1}$.

证明. 由 $A^r = S\Lambda^r S^{-1}$ 易得. □

Remark 1: 对于不可对角化的矩阵, 一般也能利用 Jordan 分解 $At = T(Jt)T^{-1}$, 将问题简化为求 e^{Jt}. 从定义式可以看出, 如果两个方阵 $KS = SK$, 则 $e^K e^S = e^{K+S}$. (可以它们的其定义式相乘, 并比对 $K^p S^q$ 项的系数) 如果 K 或者 S 是幂零矩阵和对角阵, 就能算出答案. 这里可以取对角的 D 为 Jordan 标准型对角线上的元素 (实际上就是特征值); 取 $N = J - D$, 它只有对角线上方的 1, 是幂零的. 这样 $ND = DN$, $e^{Jt} = e^{Dt}e^{Nt}$. 另外, 一个 Jordan block 去掉对角线元素之后的矩阵 (以 $n = 3$) 为例, 为

$$Nt = \begin{bmatrix} 0 & t & 0 \\ 0 & 0 & t \\ 0 & 0 & 0 \end{bmatrix}, \quad (Nt)^2 = \begin{bmatrix} 0 & 0 & t^2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad e^{Nt} = \begin{bmatrix} 1 & t & \frac{1}{2}t^2 \\ 0 & 1 & t \\ 0 & 0 & 1 \end{bmatrix},$$

形式非常简洁, 其中出现的 x 幂次的项也说明一般公式中 $C_0 + C_1 x + C_2 x^2 + \cdots$ 的部分和以上推导结果本质上是相同的.

于是, 微分方程的通解就是

$$\mathbf{u}(t) = e^{At}\mathbf{u}(0) = Se^{Jt}S^{-1} \begin{bmatrix} c_1' \\ \vdots \\ c_n' \end{bmatrix} = Se^{Jt} \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}.$$

Remark 2: 当有共轭特征值时, 特征向量也共轭, 此时 $\mathbf{c} = S^{-1}\mathbf{c}'$, 其中 $\mathbf{c}'$ 为实向量, $\mathbf{c}$ 不代表 n 个任意实数常数. 为了方便地写成实数形式, 假设 $\lambda_k, \overline{\lambda_k}$ 对应 $\mathbf{x}_k, \overline{\mathbf{x}_k}$, 则在通解中这两项为 $c_{k1}\,\mathrm{Re}(e^{\lambda t}\mathbf{x}_k)$, $c_{k2}\,\mathrm{Im}(e^{\lambda t}\mathbf{x}_k)$

Chapter 4

二次曲面与正定矩阵

4.1 曲面与空间曲线

这一部分属于对于第三章直线与平面的扩展, 着重强调对于具有规律, 可以使用代数方程进行表达的二次曲面. 对比第三章我们处理直线与平面使用的一次方程进行联立表达, 二次曲面依赖于典型的二次方程, 一般为三元二次方程.

本章的重点在于合理使用先前所学习的第六章特征向量、正交分解、矩阵对角化等知识解决实际在几何中颇有现实意义的问题, 并且得到一些十分有用的结论.

4.1.1 曲面与空间曲线的方程

- 曲面的方程

  一个曲面, 一般地, 我们可以使用直角坐标系的约束关系对于曲面进行表达。当一个方程约束了曲面的三个坐标时, 这个方程可以被称为曲面的直角坐标方程。

  一个三维的空间曲面 $F(x, y, z)$ 如果拥有两个自由变量 (自变量), 而通过所建立的约束方程 F 从而得到剩余的另一个变量的变化规律。因此我们可以认为, 三维曲面可以建立一个 $\mathbb{R}^2 \mapsto \mathbb{R}^3$ 的一个一一映射, 即空间曲面是在坐标 $(u, v) \in \mathbb{R}^2$ 映射下 (x, y, z) 的像。这就是曲面的参数方程形式。

  考虑一个球面, 其曲面上的点可以写为球坐标系表达的参数形式。

$$\begin{cases} x = x_0 + R\sin\theta\cos\phi \\ y = y_0 + R\sin\theta\sin\phi \\ z = z_0 + R\cos\theta \end{cases}$$

- 曲线的方程

  类似的, 对于三维空间中的一条曲线, 当这条曲线不再是所研究的简单直线时, 这条曲线的形式将会更加多样化。

  1. 使用两个曲面相交构成曲线: 即考虑曲面的联立解。
  2. 使用参数方程形式表达: 找到一个独立自由变化的因变量, 在对应的每一个值都能唯一确定一组 (x, y, z)。

- 三种基本曲面

  1. **柱面**: 由一条定直线与一条空间定曲线 [**准线**] 两个基本要素由一条平行于定直线的直线 [**母线**] 沿着准线所移动构成的空间曲面.

     特殊情况: 当母线平行于坐标轴, 只需取这个坐标轴的任意一个剖面, 展示另外两个坐标的约束关系 [准线方程], 即可得到此柱面方程。

     一般方程:

     $$\begin{cases} F(x,y,z) = 0 \text{ 准线方程} \\ l(x,y,z) = 0 \text{ 母线方程} \end{cases}$$

     对于一个已经给定的准线方程与母线方程, 如何才能获取它的一般直角坐标方程呢?

     我们可以任意在柱面上取一点 (x, y, z), 根据柱面的性质, 过这个点其余柱面上的点可以写为关于这个点与母线方向向量的参数方程。注意到柱面的定义是母线绕着准线所构成的空间曲线移动而产生。因此必然存在这样的一个位于母线上的一点, 使得其满足准线方程。求解出此点, 使用 (x, y, z) 表示, 代入联立方程即可得到一般方程。[见课本 P264]

     $$\text{例题 } l(x,y,z): x = y = z$$
     $$\Gamma: \begin{cases} x^2 + y^2 + z^2 = a^2 \text{ 准线方程} \\ x + y + z = 0 \text{ 母线方程} \end{cases}$$

     使用上述方法代入参数方程计算可以得到 $2x^2+2y^2+2z^2-2xy-2xz-2yz = 3a^2$.

     *　　　*　　　*

     你能够使用之后学习的二次型的知识验证这是一个垂直于平面 $x+y+z=0$ 的曲面吗?

解答: 写出曲面的二次型方程, 表达为 $x^{\mathrm{T}}Sx = x^{\mathrm{T}}\begin{bmatrix}2 & -1 & -1\\ -1 & 2 & -1\\ -1 & -1 & 2\end{bmatrix}x$;

使用正交分解基变换做坐标代换 $y = \begin{bmatrix}v_1 & v_2 & v_3\end{bmatrix}^{-1}x$, 则有 $x^{\mathrm{T}}Sx = y^{\mathrm{T}}\begin{bmatrix}0 & 0 & 0\\ 0 & 3 & 0\\ 0 & 0 & 3\end{bmatrix}y$;

原方程即化为: $(\frac{\sqrt{2}}{2}(x_1-x_3))^2+(\frac{1}{\sqrt{6}}x_1-\frac{2}{\sqrt{6}}x_2+\frac{1}{\sqrt{6}}x_3)^2+0\cdot(x_1+x_2+x_3)^2 = a^2$ 注意到 $y_3 = x_1+x_2+x_3$ 的系数为 0, 因此我们对于在这个方向上原始坐标 (x_1, x_2, x_3) 的任意变化都将被转化至此坐标系下 (y_1, y_2, y_3) 中y_3 的分量，对于曲面没有任何影响。因此这是一个在此坐标系下的 $(y_1)^2 + (y_2)^2 = 0$ 的方程, 显然是一个柱面!

2. **锥面**: 空间中一条定曲线 Γ(锥面的**准线**) 与一个定点 O(**锥面的顶点**), 一条定直线 l 绕着这个顶点 O 并沿着定曲线移动. 本书涉及的例题中只出现正锥面, 即准线平行于某一坐标平面。

   求解一般方程:

   对于一个锥面, 需要两个要素确定这个空间曲面。最为关键的是将锥面上的任意一点与我们已知的准线进行联系。由于准线是建立在一个给定的平面, 通常是 $z \equiv C$, 因此我们要将锥面上任意一点使用**参数方程**过渡至已知准线建立约束。

   锥面的参数方程:
$$\begin{cases} x = x_0 + (f_1(u) - x_0)v \\ y = y_0 + (f_2(u) - y_0)v \\ z = z_0 + (f_3(u) - z_0)v \end{cases}$$

   锥面的直角坐标方程:
$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = \frac{z^2}{c^2}$$
   我们可以通过判断当 $x, y = 0$ 时, 曲面的剖面是两条直线。对于垂直 z 轴的方向的剖面得到的是一个个不断扩大的椭圆, 因此称为椭圆锥面。当 $a = b$, 椭圆锥面退化为圆锥面。

3. **旋转面**: 在一个坐标平面的曲线围绕着构成这个坐标平面的一根轴 (一般为坐标轴) 进行旋转。

剖面方程:

$$L: \begin{cases} f(y,z) = 0 \\ x = 0 \end{cases}$$

对于上述例子, 若按照 z 轴旋转, 只需将轴面方程的参数保持旋转轴不变 [旋转不改变中心轴性质], 而根据距离相等替换另一个变量为 $\sqrt{x^2+y^2}$。因此旋转面还有一个很重要的性质就是非旋转轴两个坐标应该是对称的, 即做一个垂直于旋转轴的剖面是一个圆。我们在写完方程后可以观察方程形式验证具有**对称性的变量**是否属于非旋转方向。

- **五种典型的二次曲面**

  1. 椭球面: $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}+\dfrac{z^2}{c^2}=1(a,b,c>0)$
     - 作 $x=0,y=0,z=0$ 三个过坐标轴的平面, 显然是三个方向分别轴向长度为 a,b,c 的椭圆。
     - 作平行于坐标轴的平面, 在三个方向均为椭圆的类似形。
  2. 单叶双曲面: $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}-\dfrac{z^2}{c^2}=1$
     - 垂直于 z 轴方向的平行于 xOy 平面进行剖切, 得到的是一些等比例放缩的椭圆, 且 xOy 平面剖切的曲面最小 [不过原点]。由于在此双曲面中 z 方向是连续的, 因此我们称之为单叶双曲面, 与双叶进行区分。
     - 对于含有中心轴 z 的平面进行剖切, 即固定 x *or* y 的一个, 得到均为实轴位于 x *or* y 的双曲线。因此此双曲面约束为 $|x|>a$ 与 $|y|>b$.
  3. 双叶双曲面: $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}-\dfrac{z^2}{c^2}=-1$
     - 类似单叶双曲面, 我们首先作平行于 xOy 的剖面. 注意到左侧 $(\frac{x^2}{a^2}+\frac{y^2}{b^2})\geq 0$, 因此 $|z|>c$, 即在 z 轴方向有间断。因此我们称之为双叶双曲面。
     - 作平行于 xOz *or* yOz 平面的剖面, 需要注意的是这时实轴是 z.
  4. 椭圆抛物面: $\dfrac{x^2}{p}+\dfrac{y^2}{q}=2z$
     - 作 zOx 与 zOy 两个过坐标轴的平面, 得知均为抛物面, 同时 z 轴正向均为抛物线的开口方向。
     - 作平行于 xOy 的平面, 可以观察知是一个随 z 的坐标增大而增大的椭圆。

5. 双曲抛物面: $\dfrac{x^2}{p}-\dfrac{y^2}{q}=2z$

   - 作 zOx 与 zOy 两个过坐标轴的平面, 得知均为抛物面, 但是由于左侧符号差异, 构成的抛物线开口方向相反。
   - 双曲抛物面又称为马鞍面, 在 Oxy 平面的交线为两条相交于原点的直线。而平行于 Oxy 平面的交线为实轴不断增大的双曲线。

事实上, 当二次型曲面关于两个坐标平面对称时, 即方程中的两个参数是对称的, 我们的二次型曲面可以使用旋转方式构造。例如曲面 $x^2+y^2=z$ 可以看作关于 z 轴在 Oyz 平面上的曲线 $z=y^2$ 旋转而成。而二次型曲面就是在旋转的基础上对于某些方向进行拉伸得到的更广义的曲面。

当方程中 (x,y,z) 均次数为两次时, 其曲面具有中心对称性, 双叶单叶双曲面, 椭球, 锥面均为有心二次曲面, 而椭圆, 双曲抛物面称为无心二次曲面。

4.1.2 曲线在坐标面上的投影

对于空间中的一条曲线, 假设它是由两个平面相交所构成的, 即:

$$\Gamma:\begin{cases}F_1(x,y,z)=0\\F_2(x,y,z)=0\end{cases}$$

注意到这条曲线对于某一个平面的投影可以理解为垂直于这个平面构建的一个柱面, 而这条曲线一定处于这个柱面之上。因此我们约束这条曲线的方程组中必然**满足**柱面方程。因此假设我们可以从曲线 Γ 中解出一个柱面方程 $\varphi(x,y)=0$, 则这个方程一定代表过这个曲线的柱面, 也就是这个曲线在这个方向的投影。

因此我们在解题时只需要**消去对应投射方向的因变量**即可得到对应投射曲线的方程。

4.1.3 空间简图

[空间简图常用在重积分中, 目前涉及较少, 因此笔者此块略写, 下方将给出一道例题]

* * *

例 (P227 7.1(B)Q2): 已知点 A 与点 B 的直角坐标分别为 $(1,0,0)$ 和 $(0,1,1)$, 线段 AB 绕 z 轴旋转一周所形成的曲面为 S, 并利用定积分求由 S 与平面 $z=0,z=1$ 所围成的空间立体的体积。

解: 对于关于坐标轴旋转的问题, 解题策略均为构建一个定量 (曲面上的点到 y 轴距离保持不变)。利用旋转体积分方法 $V=\int_0^1 \pi x(y)^2 dy$ 以 y 轴进行积分则得到体积。

根据端点写出直线 AB 方程为 $\dfrac{x-1}{-1}=\dfrac{y}{1}=\dfrac{z}{1}$. 对于任意给定高度 z_0, 我们使用此直线方程解得一组 (x,y), 从而得到曲面上高度为 z_0 点到 z 轴的距离。

写出旋转面依赖半径不变的曲面方程:$\pm\sqrt{x^2+y^2}=\sqrt{z^2+(1-z)^2}$, 化简得到曲面方程为 $x^2+y^2=2z^2-2z+1$.

代入积分计算: $V=\int_0^1 \pi(2z^2-2z+1)dz=\frac{2}{3}\pi$

例 (P299 综合练习题 Q7) 求准线为:

$$\begin{cases} xy=4 \\ z=0 \end{cases}$$

母线平行于向量 $(1,-1,1)^{\mathrm{T}}$ 的柱面方程。

对于柱面上的一点,(x,y,z), 使用母线参数方程可以得到柱面一条线上所有点的方程:$a=(x+t,y-t,z+t)$. 代入满足准线方程的一点 $z'=0, s.t.\ z=-t$, 解得这一点坐标为 $(x-z,y+z,0)$. 代入准线方程得: $(x-z)(y+z)=4$. 柱面方程即为 $xy+xz+yz-z^2-4=0$. 显然这不是平行于任何坐标轴的柱面。

4.2 实二次型

4.2.1 二次型及其标准型

- **二次型**: 称关于 n 个变量 $x_1, x_2, \ldots, x_n$ 的**二次齐次**多项式函数 $f(x) = a_{11}x_1^2 + 2a_{12}x_1x_2 + 2a_{13}x_1x_3 + \cdots + 2a_{nn}x_n^2$ 为一个 n 元二次型。

  注意到上方所写的二次多项式中含有 2 的项均有 $a_{ij}, i < j$。不妨构造对称矩阵 $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$ 则有 $f(x) = x^{\mathrm{T}}Ax, x \in \mathbb{R}^3$。

  我们考虑的二次型均为实数域范围, 因此我们将实对称矩阵 A 称为二次型矩阵, 将实对称矩阵 A 的秩称为这个二次型的秩。

  利用矩阵表达二次型带来了极大的便利, 因此我们计算多项式方程可以使用矩阵的形式清晰展现每一项的系数。而我们所构造的对称矩阵也为我们进行标准化带来了极大的便利。需要说明的是, 我们所使用的二次型实对称矩阵可以唯一地确定某一个二次型。当题目中矩阵不是标准二次型矩阵时, 使用变换 $S = \dfrac{A + A^{\mathrm{T}}}{2}$ 既可获得二次型矩阵。

- **二次型的标准型**

  我们希望使得我们的二次型拥有一个十分简洁的形式, 即

$$f(x) = x^{\mathrm{T}} \begin{bmatrix} c_{11} & 0 & \cdots & 0 \\ 0 & c_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & c_{nn} \end{bmatrix} x$$

  如果我们选择一个合适的 $x = Qy$, 并且要求这个 Q 可逆。这样使得代入后可以将内部的二次型矩阵化简为对角阵, 既可以完成我们的目标。又因为我们构造的属于实对称矩阵, 因此我们一定可以使用矩阵正交分解, 选择 $Q\ s.t.\ Q^{\mathrm{T}}AQ = D$, 既可以实现二次型的简化。

  (二次型的标准型) 我们称形如 $f(x) = d_1y_1^2 + d_2y_2^2 + \cdots + d_ny_n^2$ 的表达式为这个二次型的标准型。

4.2.2 正定二次型

- **合同变换与惯性定理**

在二次型中我们对于坐标 x 进行线性变换 $x = Cy$, 从而二次型的矩阵变换为 $C^{\mathrm{T}}AC$. 为了表达变换的不变量与一致性关系, 我们定义一种新的矩阵间的关系: **合同关系**。为了更好地回顾我们已学的三种关系, 在这里将三种矩阵间的关系列举如下。

1. (**合同关系**) 设 A, B 为 n 阶矩阵, 如果存在 n 阶可逆矩阵 C, 使得

$$C^{\mathrm{T}}AC = B$$

   则称矩阵 A, B 是合同的。

2. (**等价关系**) 设 A, B 都是 $m \times n$ 阶矩阵, 等价关系的充要条件为:

$$A\text{与}B\text{等价} \iff \text{存在可逆矩阵}P_{m\times m}, Q_{n\times n}, \text{使得}PAQ = B$$

3. (**相似关系**) 设 A, B 都是 n 阶方阵, 如果存在一个 n 阶可逆矩阵 P, 使得:

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

   则称矩阵 A, B 是相似的。

注: 矩阵的三种关系一般而言并没有明确的包含, 等价等关系, 因此实际情况中我们需要认真从矩阵性质中进行分析。

- 当我们考虑的矩阵是方阵时, 更特别的是实对称矩阵时, 相似关系可以理解为合同关系的升级版, 因为合同关系需要保证正惯性指数与负惯性指数相同, 而相似关系在实对称矩阵中进一步保证了特征值完全相同, 因此正负号自然满足惯性指数规律。需要注意的是, 合同矩阵显然在任何情况下不能直接推出相似。
- 合同关系要求 C 是可逆矩阵,C^{T} 与 C 自然满足等价关系的定义, 因此满足合同关系的矩阵一定等价。类似的, 相似矩阵也一定等价。因此具有合同关系与相似关系的矩阵必然存在相同的秩, 也一定拥有相同个数特征值为 0 的特征向量, 并且在标准形中具有相同的非零项个数。

我们重新回到介绍相互具有合同关系的矩阵在我们所研究的二次型范围内的作用。类似于其他的矩阵关系, 合同关系满足**自反性, 对称性, 传递性**。这里我们简单叙述一下合同关系的传递性:

$$B = C^{\mathrm{T}}AC,\ Z = Q^{\mathrm{T}}BQ \implies Z = (CQ)^{\mathrm{T}}A(CQ)$$

(惯性定理) 对于一个二次型 $f(x) = x^{\mathrm{T}}Ax$ 的秩为 r, 它的系数矩阵化为标准形时, 有 $n - r(A)$ 个非零元 (等价自然满足)。剩余的非零元素的正负数量不随着任何可逆矩

阵的线性变换而改变, 也就是正惯性指数 (标准型中大于 0 的元素) 与负惯性指数个数是保持不变的。

[几何解释] 根据对于实对称矩阵的理解, 我们实际上可以为二次曲面的计算给出一个十分优美的解释。对于任意一个实对称矩阵, 由于其必然可以进行正交分解, 因此我们可以把矩阵 S 分解为 $S = QDQ^{\mathrm{T}}$。因此我们将一个矩阵 S 作用给任何一个向量时, 都有 $Sx = QD(Q^{\mathrm{T}}x)$。棕色的部分属于对于 x 的正交变换, 根据 (Spectrum Theory) 谱分解, 我们的作用分为将基位于特征空间 Q 上的向量旋转至一般直角坐标系, 进行拉伸, 再旋转至特征向量的基。因此我们的二次型 $x^{\mathrm{T}}Sx$ 实际上就是在旋转回到原始特征向量的坐标空间上时再与原始向量做一次内积。显然在特征空间上 $\sum \lambda_i y_i^2$。因此当我们将坐标系直接选择为 $x = Qy$, 直接将我们的向量理解为一个处于特征空间中的向量, 把它在特征空间的坐标作为自变量, 这时就可以理解为内积计算得到结果。

$$f(x) = x^{\mathrm{T}}Sx(\text{think x in Original bases}) = y^{\mathrm{T}}Dy(\text{think input y in C(Q)})$$

$$f(x) = (y^{\mathrm{T}}Q^{\mathrm{T}})\ (QDQ^{\mathrm{T}})\ (Qy) = y^{\mathrm{T}}Dy$$

得到我们上方所解释的计算结果: $\sum \lambda_i y_i^2$.

- **二次型的规范型**

  我们已经介绍了如何将一个二次型使用线性变换 $x = Cy$ 化为二次型的标准型 $f(x) = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots \lambda_n y_n^2$。在介绍了惯性定理后, 我们可以将先前介绍的标准型统一地变化为一种形式, 我们对于每一个 $\lambda_i y_i$, 再次进行规范化, 即 $y_i' = \frac{y_i}{\sqrt{\lambda_i}}$, 则可以获得二次型的规范型。

  惯性定理实际上反映了二次曲面几何上的不变性, 即无论做什么线性变换, 反映这个曲面的关键性质, 二次项正负数量与对应特征值 (零元) 的数目必须是保持不变, 一个二次曲面具有唯一的规范形。而我们做的规范化可以从几何上理解为将特征空间的 (基) 向量取的更短, 从而将拉伸直接融入计算特征空间分量的大小 [但是进行规范化后合同变换的矩阵在特征值不等于 1 时不再是正交矩阵, 只是普通可逆矩阵]。

- **正定二次型**

  (正定二次型) 对于一个确定二次型的矩阵 $A_n \times n$, 如果 $\forall \frown \in \mathbb{R}^n$, $x \neq 0$, 都有 $f(x) = x^{\mathrm{T}}Ax > 0$, 则称 f 是正定二次型, 实对称矩阵 A 为正定矩阵。

  (保正定性) 对于任意一个可逆矩阵 P, 做变换 $x = Py$, 对于新的实对称矩阵 $P^{\mathrm{T}}AP$ 而言, 正定型不变。

证明是显然的, 对于可逆变换而言,x 坐标空间与 y 是同构的, 可以建立一个从 $x \leftrightarrow y$ 的一个一一映射。因此当已知某一个正定矩阵与另一个矩阵等价 (经过**合同变换**) 的二次型, 验证另一个矩阵的正定性只需要将坐标 y 使用两个坐标空间的过渡矩阵转换回原正定矩阵, 而原矩阵显然可以保证在它的任意坐标是满足大于 0 条件的。

<u>正定矩阵的判定准则</u>

1. (充要条件)实对称矩阵的所有特征值大于 0.
2. (充要条件)实对称矩阵的所有主元 (化简为阶梯型) 大于 0.

   感兴趣的同学可以探索一下特征值与主元正负的联系。特征值与主元对于同一个矩阵而言符号完全相同。
3. (充要条件)实对称矩阵 A 存在可逆矩阵 M, 使得:

   $$A = M^{\mathrm{T}}(I)M \iff A\text{与单位矩阵合同}.$$

   这样的 M 可以取为 $Pdiag(\sqrt(\lambda_1), \sqrt(\lambda_2), ..., \sqrt(\lambda_n))P^{\mathrm{T}}$,$P$ 为矩阵 A 相似对角化的正交矩阵。

   A 同时可以进行 LDL^{T} 分解, 因此可以取 M 为 $L^{\mathrm{T}}diag(\sqrt(\lambda_1), \sqrt(\lambda_2), ..., \sqrt(\lambda_n))$.
4. (必要条件) 矩阵行列式大于 0。

   如果存在偶数个负特征值也可以使得行列式大于 0, 因此不是充分条件。同时, 此必要条件的逆反命题为行列式小于 0 的矩阵一定不是正定矩阵。
5. (充要条件)实对称矩阵 A 的各阶顺序主子式大于 0。
   - 充分性: 主元与顺序主子式有着密切的关系。$d_k = \dfrac{\det(A_k)}{\det(A_{k-1})}$ 若每一个主子式大于 0, 两者的商大于 0, 因此矩阵的主元均大于 0. 根据判定准则 2 的充要性得出。
   - 必要性: 使用反证法进行验证。

     假设它的某一阶顺序主子式小于 0[为了简便, 设它的 $n-1$ 阶顺序主子式小于 0], 根据判定的必要条件, 这个子阵 (A') 一定不是正定的。

     通过定义, 我们可以任意的构造一个 x, 使得 $x^{\mathrm{T}}Ax > 0$。我们取 $x \in \mathbb{R}^n = [x_1, x_2, \dots x_n - 1, 0]$, 使得 $f(x') = [x_1, x_2, \dots x_n - 1]^{\mathrm{T}}A[x_1, x_2, \dots x_n - 1] < 0$

     根据矩阵乘法, 我们可以得知 $f(x) = x^{\mathrm{T}}Ax = x'^{\mathrm{T}}Ax' < 0$, 因此与 A 是正定矩阵矛盾, 固顺序主子式一定大于 0. 反言之,<u>正定矩阵所有主子式都是正定矩阵</u>。

6. (充要条件) 实对称矩阵 A 的各阶主子式都是正的。只需证明必要性。类似的使用选择合适的输入 x, 获得分块主子式矩阵二次型计算表达, 得到分块子式的正定性。根据必要条件 4, 保证行列式必然大于 0.

   (其他类型) 类似的, 我们可以定义不严谨大于 0 的二次型矩阵为半正定矩阵 (Semipositive definite matrix). 根据正定与负定的对称关系, 可以类似的定义定义负定, 半负定矩阵。

   注: 课本对于半正定矩阵的定义为对于任意的向量 x, $f(x) = x^{\mathrm{T}}Ax \geq 0$, $\exists x \in \mathbb{R}^n, s.t. x^{\mathrm{T}}Ax = 0$. 因此我们不能将正定矩阵包含于半正定的定义内。

课本习题精选

1. 已知 A 是正定矩阵, 证明 A^2, A^{-1}, A^* 都是正定矩阵。

   判断正定矩阵的充要条件知, 我们只需要通过对角化方法得到所有特征值均大于 0, 则说明矩阵是正定的。当然前提条件是这个对角化是合理的, 我们在使用合同变换进行换参不能使用不可逆矩阵, 否则参数不独立在 x 不等于 0 的情况下得到 y 等于 0 的半正定矩阵。对于上述三个矩阵, 特征值分别为 $\lambda^2, \dfrac{1}{\lambda}, \dfrac{\det(A)}{\lambda}$. 正定矩阵的行列式必然大于 0, 因此有三个矩阵特征值均大于 0.

2. 实对称矩阵 A 是正定矩阵的一个必要条件是对角线元素 a_i 都是正的。

   通过我们先前所证明的判定准则 6 的充要性可以直接得出, 对角线上的元素可以看做取一行一列的主子式。

3. 实数 a_1, a_2, a_3 满足什么条件时, 二次型 $f(x_1, x_2, x_3) = (x_1 + a_1x_2)^2 + (x_2 + a_2x_3)^2 + (x_3 + a_3x_1)^2$ 为正定矩阵。

   我们在判断两个通过一个矩阵变换后得到的二次型矩阵与原矩阵是否具有相同的性质时, 书上明确要求变换矩阵 M 必须可逆。因此我们将当前每个平方项看做一个整体, 记作 y, 原矩阵与当前 y 作为参数的正定矩阵具有一样的性质当且仅当 M 可逆。通过表达式, 我们可以写出 $\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & a_3 \\ a_1 & 1 & 0 \\ 0 & a_2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ 计算得到 $1 + a_1a_2a_3 \neq 0$. 我们也可以通过理解, 正定矩阵要求就是对于任意的 x, 都有二次型输出大于 0. 从表达式看出大于等于 0 是显然的, 那么只需要说明 x 的任意取法不会使得每一个 y 均为 0。

4. 对于任意的 x, 都有 $x^{\mathrm{T}}Ax = x^{\mathrm{T}}Bx$, 则 $A = B$.

只需证, $x^{\mathrm{T}}(A-B)x=0$当且仅当$A=B$. 二次型的值恒为 0, 说明其规范方程系数均为 0. 即矩阵 $A-B$ 与零矩阵合同。由于合同必然等价, 因此原矩阵 $A-B=0$. 这个结论说明了二次型的唯一性。不会存在两个实对称二次型具有相同的函数值但是形式不同。

5. 实对称矩阵 A 是半正定的充分必要条件是正惯性指数与秩相等, 同时 $r(A)<n$ 或矩阵 A 的特征值大于等于 0, 且存在至少一个特征值为 0. [书有误]

   对角化后发现 $f(x)=\lambda_1y_1+\lambda_2y_2+\cdots+\lambda_ry_r$. 如果有负的特征值, 只需取这个 y_i 不等于 0, 函数值小于 0, 负定. 因此特征值均大于等于 0. 特别的, 矩阵等价确保了零特征值个数就是矩阵的秩, 剩下的均为正惯性指数。同时我们要求必须有一个 x, $s.t. x^{\mathrm{T}}Ax=0$. 因此我们必须要求矩阵 A 的秩小于其行列数。

6. $*$ A 是 n 阶正定矩阵, 证明下列结论成立:

   - $\det(A)\le a_{nn}\cdot\det(A_{n-1})$
   - $\det(A)\le a_{11}a_{22}\cdots a_{nn}$

   第二小问是第一小问的拓展, 这里只阐述第一小问。为了证明这一结论, 我们首先构造矩阵 $B=\begin{bmatrix}A_{n-1} & \\ & a_{nn}\end{bmatrix}$ 为了能够与第二个矩阵行列式产生关联, 使用高斯消元法, 将矩阵 $A=\begin{bmatrix}A_{n-1} & v\\ v^{\mathrm{T}} & a_{nn}\end{bmatrix}$ 最后一列的 v 消为 0 向量。由于 $\det A_{n-1}>0$, 则必然存在矩阵 P, $s.t. PA_{n-1}=v^{\mathrm{T}}$[行变换]. 因此矩阵 A 变换为 $\begin{bmatrix}A_{n-1} & v\\ 0 & a_{nn}-v^{\mathrm{T}}A_{n-1}v\end{bmatrix}$. 考虑到主对角线的主子式矩阵也是正定矩阵, 因此对于任意的一个向量 $v\in\mathbb{R}^{n-1}$, 都有 $-v^{\mathrm{T}}A_{n-1}v<0$. 因此 $a_{nn}-v^{\mathrm{T}}A_{n-1}v<a_{nn}$.

$$\det(A)=\det\left(\begin{bmatrix}A_{n-1} & v\\ 0 & a_{nn}-v^{\mathrm{T}}A_{n-1}v\end{bmatrix}\right)<\det\left(\begin{bmatrix}A_{n-1} & \\ & a_{nn}\end{bmatrix}\right)=a_{nn}\cdot(A_{n-1})$$

7. (Priciple Direction) 记实数集合 $f(x)=\left\{x^{\mathrm{T}}Ax\,|x\in\mathbb{R}^n,\|x\|=1\right\}$. 证明矩阵 A 的二次型变换值域为 $[\lambda_1,\lambda_2]$.λ_1 是最小特征值,λ_2 为最大特征值.

   对于一个实对称矩阵, 必然可以进行正交分解。假设 A 可以分解为 QDQ^{T},Q满秩, 因此 $f(x)=x^{\mathrm{T}}(Q^{\mathrm{T}})^{\mathrm{T}}DQ^{\mathrm{T}}x$. 因此我们做变换 $x=Qy$, 由于 Q 是正交矩阵, 因此得到的变换基坐标 y 应该与 x 具有相同的模长 [保距性]. 我们可以将表达式重新写为 $f(x)=\lambda_1y_1^2+\lambda_2y_2^2+\cdots+\lambda_ny_n^2$. 根据线性组合性质显然位于 $[\lambda_1,\lambda_2]$

Chapter 5

几何向量与应用

向量的线性运算运算包括加法与数乘运算。

定理 1. 向量的三角不等式

- 非负性: $\|\mathbf{a}\| \geq 0$, 且 $\|\mathbf{a}\| = 0 \iff \mathbf{a} = 0$;
- 齐性: $\|\lambda \mathbf{a}\| = |\lambda| \, \|\mathbf{a}\|$;
- 三角不等式: $\|\mathbf{a} + \mathbf{b}\| \leq \|\mathbf{a}\| + \|\mathbf{b}\|$ 等号成立 $\iff$ $\mathbf{a}$ 与 $\mathbf{b}$ 同向。

定理 2. 向量共线的充要条件

$$\text{存在不全为0的} k_1, k_2, \ s.t. \ k_1\mathbf{a} + k_2\mathbf{b} = 0$$

定理 3. 向量共面的充要条件

$$\text{存在不全为0的} k_1, k_2, k_3 \ s.t. \ k_1\mathbf{a} + k_2\mathbf{b} + k_3\mathbf{c} = 0$$

推论 1. 零向量与任意向量都共线, 平行

定理 4. 向量的直角坐标与空间余弦

$$\begin{cases} \cos\alpha = \dfrac{x}{\|\mathbf{a}\|} = \dfrac{x}{\sqrt{x^2+y^2+z^2}} \\ \cos\beta = \dfrac{y}{\|\mathbf{a}\|} = \dfrac{y}{\sqrt{x^2+y^2+z^2}} \\ \cos\gamma = \dfrac{z}{\|\mathbf{a}\|} = \dfrac{z}{\sqrt{x^2+y^2+z^2}} \end{cases} \tag{5.1}$$

定理 5. 坐标表示下的共线共面条件

$$\mathbf{a}\text{与}\mathbf{b}\text{共线} \iff \mathbf{a} = \lambda\mathbf{b} \iff x_1 : x_2 : x_3 = y_1 : y_2 : y_3$$

$$\begin{cases} x_1k_1 + y_1k_2 + z_1k_3 \\ x_2k_1 + y_2k_2 + z_2k_3 \\ x_3k_1 + y_3k_2 + z_3k_3 \end{cases} \iff \det\begin{bmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{bmatrix} \tag{5.2}$$

定理 6. 正交射影

$$\mathbf{Proj}_a b = \|\mathbf{b}\| \cos\theta\, \widehat{\mathbf{a}} = (b \cdot \mathbf{a})\ \hat{a}$$

定理 7. 向量的数量积 *(内积, 点积)*

- $\mathbf{a}\cdot\mathbf{b} \;=\; \mathbf{b}\cdot\mathbf{a}$
- $(\mathbf{a}+\mathbf{b})\cdot c = (\mathbf{a}\cdot\mathbf{c}) + (\mathbf{b}\cdot\mathbf{c})$
- $(k\mathbf{a})\cdot\mathbf{b} = k(\mathbf{a}\cdot\mathbf{b})$
- $\mathbf{a}\cdot\mathbf{a} = \|a\|^2 \geq 0 \iff \mathbf{a} = \mathbf{0}$

定理 8. 数量积的应用

$$\|\mathbf{a}\| = \sqrt{a\cdot a} = \sqrt{a_x^2 + a_y^2 + a_z^2}$$
$$\cos(\mathbf{a},\mathbf{b}) = \frac{\mathbf{a}\cdot\mathbf{b}}{\|\mathbf{a}\|\,\|\mathbf{b}\|}$$

定理 9. 两个向量的外积 *(叉乘)*

$$\|\mathbf{a}\times\mathbf{b}\| = \|\mathbf{a}\|\,\|\mathbf{b}\|\sin(\mathbf{a},\mathbf{b})\text{方向为右手螺旋方向}$$

定理 10. 向量叉乘的性质

- $\mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a}$
- $(k\mathbf{a}) \times \mathbf{b} = a \times (k\mathbf{b}) = k(\mathbf{a} \times \mathbf{b})$
- $\mathbf{a} \times (\mathbf{b} + \mathbf{c}) = \mathbf{a} \times \mathbf{b} + \mathbf{a} \times \mathbf{c}$

定理 11. 向量叉乘的坐标表示

$$\mathbf{a} \times \mathbf{b} = \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_x & a_y & a_z \\ b_x & b_y & b_z \end{bmatrix}$$

定理 12. 向量的混合积

$$[\mathbf{a}\ \mathbf{b}\ \mathbf{c}] = (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}$$

平行六面体体积公式:

$$V = \|\mathbf{a} \times \mathbf{b}\| \cdot (c)_{a \times b}$$

四面体体积公式

$$V = \frac{1}{6} \|\mathbf{a} \times \mathbf{b}\| \cdot (c)_{a \times b}$$

定理 13. 混合积的坐标计算

$$[\mathbf{a}\ \mathbf{b}\ \mathbf{c}] = \det \begin{bmatrix} a_x & a_y & a_z \\ b_x & b_y & b_z \\ c_x & c_y & c_z \end{bmatrix}$$

$$(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} = (\mathbf{b} \times \mathbf{c}) \cdot \mathbf{a} = (\mathbf{c} \times \mathbf{a}) \cdot \mathbf{b}$$

$$[\mathbf{b}\ \mathbf{a}\ \mathbf{c}] = -[\mathbf{a}\ \mathbf{b}\ \mathbf{c}]$$

定理 14. 向量共面的充要条件 2

$$\mathbf{a}, \mathbf{b}, \mathbf{c}\text{共面} \iff \text{混合积}[\mathbf{a}\ \mathbf{b}\ \mathbf{x}] = 0$$

定理 15. 空间平面的方程

[平面的点法式方程] $A(x-x_0)+B(y-y_0)+C(z-z_0)=0$. (A,B,C)是平面的法向量

[平面的一般方程]

- 通过原点 $Ax+By+Cz=0$
- 平行于坐标轴 $Ax+By+D=0$
- 通过坐标轴 $Ax+By=0$

[平面的截距式方程] $\frac{x}{a}+\frac{y}{b}+\frac{z}{c}=1$ [平面的参数式方程]

$$\begin{cases} x=x_0+sL_1+tL_2 \\ y=y_0+sM_1+tM_2 \\ z=z_0+sN_1+tN_2 \end{cases}$$

定理 16. 空间平面的夹角

$$\cos\theta=\frac{|\mathbf{n_1}\cdot\mathbf{n_2}|}{\|\mathbf{n_1}\|\,\|\mathbf{n_2}\|}$$

定理 17. 平面的三种位置关系

- π_1 与 π_2 相交 $\iff$ $\mathbf{n_1}$ 与 $\mathbf{n_2}$ 不平行 $iff a_1:b_1:c_1\neq a_2:b_2:c_2$
- π_1 与 π_2 平行但不重合 $\frac{a_1}{a_2}=\frac{b_1}{b_2}=\frac{c_1}{c_2}\neq\frac{d_1}{d_2}$
- π_1 与 π_2 平行且重合 $\frac{a_1}{a_2}=\frac{b_1}{b_2}=\frac{c_1}{c_2}=\frac{d_1}{d_2}$

定理 18. 空间直线的方程

[直线的参数式方程]

$$\begin{cases} x=x_0+lt \\ y=y_0+mt \\ z=z_0+nt \end{cases}$$

[直线的对称式方程]

$$\frac{x-x_0}{l}=\frac{y-y_0}{m}=\frac{z-z_0}{n}$$

[直线的一般式方程]

$$a_1x+b_1y+c_1z+d_1=0$$
$$a_2x+b_2y+c_2z+d_2=0$$

定理 19. 直线的夹角

$$\cos\theta = \frac{|\mathbf{l_1} \cdot \mathbf{l_2}|}{\|\mathbf{l_1}\| \, \|\mathbf{l_2}\|}$$

定理 20. 两条直线的位置关系

- L_1 与 L_2 异面 $\iff$ 两条直线间一个向量 $\overrightarrow{P_1P_2}, \mathbf{a_1}, \mathbf{a_2}$ 不共面
- L_1 与 L_2 相交 $\iff$ $\overrightarrow{P_1P_2}, \mathbf{a_1}, \mathbf{a_2}$ 共面, 且 $\mathbf{a_1}$ 与 $\mathbf{a_2}$ 不平行
- L_1 与 L_2 平行 $\iff$ $\overrightarrow{P_1P_2}, \mathbf{a_1}, \mathbf{a_2}$ 共面, 且 $\mathbf{a_1}$ 与 $\mathbf{a_2}$ 平行但与 $\overrightarrow{P_1P_2}$ 不平行
- L_1 与 L_2 重合 $\iff$ $\overrightarrow{P_1P_2}, \mathbf{a_1}, \mathbf{a_2}$ 平行

定理 21. 直线与平面的夹角

$$\sin\varphi = |\cos(\mathbf{a}, \mathbf{n})| \, (\varphi \in [0, \frac{\pi}{2}])$$

定理 22. 距离

$$[\text{点到平面的距离}] \; d = \frac{Ax_1 + By_1 + Cz_1 + D}{\sqrt{a^2 + b^2 + c^2}}$$

$$[\text{点到直线的距离}] \; d = \frac{\overrightarrow{P_0P_1} \times \mathbf{a}}{\|a\|}$$

$$[\text{异面直线的距离}] \; d = \frac{\left|(\mathbf{a_1} \times \mathbf{a_2}) \cdot \overrightarrow{P_1P_2}\right|}{\|\mathbf{a_1} \times \mathbf{a_2}\|}$$

参考文献

[1] Kenji Hiranabe. *Github: The-Art-of-Linear-Algebra.*

[2] Gilbert Strang. *Linear Algebra and Its Applications.* Thomson Brooks/Cole, 2006.

[3] 李继成，魏战线. 线性代数与解析几何. 高等教育出版社, 2019.