线性代数

胡越

北京邮电大学理学院

《线性代数》：公共必修课，为学好专业课打下良好基础，大学生数学竞赛，建模竞赛，考研等等都需要.

课程特点：知识点多，抽象性高，初学者不易掌握. 如何学好?

- 从简单到复杂，从例子到抽象，掌握核心概念和方法.
- 上课认真听讲，作业独立并认真完成.
- 自主学习，看推荐的教材，做适量的习题.

教材及参考书:

- 刘吉佑, 莫骄, 线性代数与几何 (第 2 版), 北京邮电大学出版社.
- 陈发来, 陈效群, 李思敏, 王新茂, 线性代数与解析几何 (第二版), 高等教育出版社.
- 刘吉佑, 莫骄, 李亚杰, 线性代数与几何学习辅导 (第 2 版), 北京邮电大学出版社.

课程考核: 平时作业 $30\% +$ 期末考试 70%.

作业: 教学云平台提交, 逾期成绩 $\times 70\%$.

完成好每周的功课, 加上有计划的复习, 期末才会游刃有余.

课程目标:

- 获得好成绩.
- 锻炼思维能力, 提升数学修养.
- 获得启发与快乐.

1 复数与向量

- 复数及其几何表示
- 空间向量
- 内积, 外积和混合积
- 直线与平面
- 高维向量

1.1

复数及其几何表示

- **数系的发展.**

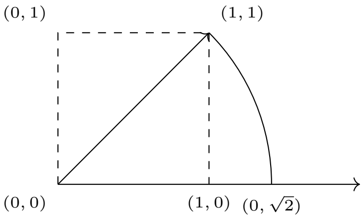
计数: 自然数 $\mathbb{N}$ (Natural): $\{0, 1, 2, \dots\}$.

亏欠, 减法: 整数 $\mathbb{Z}$ (Zahlen 德文): $\{\dots, -2, -1, 0, 1, 2, \dots\}$.

分割, 除法: 有理数 $\mathbb{Q}$ (Quotient): $\{\frac{m}{n} \mid m, n \in \mathbb{Z}, n \neq 0\}$.

注记 1.1

有理数的英文词根为 “ratio”, 是 “比例” 的意思, 和 “有理” 实际没有关系.



有理数不“连续”: 将有理数看成直线上的点, 则有理点之间的空隙为“无理数”.

实数 $\mathbb{R}$ (Real): 平面上 x 轴上所有的点, 称为**实直线**; 或者 $a \in \mathbb{R}$ 可以十进制表示为

$$\begin{aligned} a &= a_0.a_1a_2\cdots a_k\cdots \qquad (a_0\in\mathbb{Z},\ 0\le a_1,a_2,\cdots\le 9)\\ &= a_0+\frac{a_1}{10}+\frac{a_2}{100}+\cdots+\frac{a_k}{10^k}+\cdots. \end{aligned}$$

有理数：有限小数及无限循环小数. (为什么？)

无理数: 无限不循环小数.

二次方程求解: $x^2+1=0$ 没有实数解.

复数 $\mathbb{C}$ (Complex): $\{z=x+\mathrm{i}y \mid x,y\in\mathbb{R}\}$, 其中 $\mathrm{i}=\sqrt{-1}$ 是引进的满足 $\mathrm{i}^2=-1$ 的 “数”, 称为虚数单位.

$x=\mathrm{Re}(z)$ 称为 z 的实部, $y=\mathrm{Im}(z)$ 称为 z 的虚部.

▶ **复数的运算.** 　定义 $z=x+\mathrm{i}y$ 的**共轭**为 $\bar{z}=x-\mathrm{i}y$, 定义 $z=x+\mathrm{i}y$ 的**模**为 $|z|=\sqrt{x^2+y^2}$, 则有 $z=0 \Leftrightarrow |z|=0$.

在复数集 $\mathbb{C}$ 上, 可以定义通常的加, 减, 乘, 除运算:

$$z_1=x_1+\mathrm{i}y_1, \quad z_2=x_2+\mathrm{i}y_2,$$

$$z_1 \pm z_2=(x_1 \pm x_2)+\mathrm{i}(y_1 \pm y_2),$$

$$z_1 \cdot z_2=(x_1x_2-y_1y_2)+\mathrm{i}(x_1y_2+x_2y_1),$$

$$\frac{z_1}{z_2}=\frac{x_1x_2+y_1y_2}{x_2^2+y_2^2}+\mathrm{i}\frac{x_2y_1-x_1y_2}{x_2^2+y_2^2}, \quad (z_2 \neq 0).$$

且有通常的零元 0 $(0+a=a+0=a,\ 0\cdot a=a\cdot 0=0)$, 以及单位元 1 $(1\cdot a=a\cdot 1=a)$.

▶ 数域.

定义 1.2

设 $\mathbb{F}$ 是由一些复数构成的集合, 且 $0, 1 \in \mathbb{F}$. 如果对任意 $\mathbb{F}$ 中任意两个数, 它们的和, 差, 积, 商 (除数不为 0) 都仍然属于 $\mathbb{F}$, 称 $\mathbb{F}$ 是一个数域.

显然, 所有有理数, 实数, 复数组成的集合都构成数域, 分别记为 $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$. 事实上, 任何数域 $\mathbb{F}$ 都包含 $\mathbb{Q}$.

例 1.3

$\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2}, a, b \in \mathbb{Q}\}$ 也构成一个数域.

$$(a+b\sqrt{2})\pm(c+d\sqrt{2})=(a\pm c)+(b\pm d)\sqrt{2},$$

$$(a+b\sqrt{2})(c+d\sqrt{2})=(ac+2bd)+(ad+bc)\sqrt{2},$$

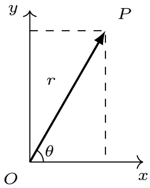
$$\frac{a+b\sqrt{2}}{c+d\sqrt{2}}=\frac{ac-2bd}{c^2-2d^2}+\frac{-ad+bc}{c^2-2d^2}\cdot\sqrt{2},\ \ c+d\sqrt{2}\neq 0.$$

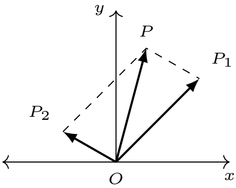
所有整数构成的集合 $\mathbb{Z}$ 也是代数研究的重要对象, 但它不是一个数域, 因为它在除法下不封闭.

这类仅要求有加法, 减法及乘法, 并在这些运算下封闭的代数对象, 我们统称为**环**. $\mathbb{Z}$ 称作**整数环**.

- **复平面.** 在平面上取直角坐标系 Oxy, 每个复数 $z = x + \mathrm{i}y$ 一一对应于平面上的一个坐标为 (x, y) 的点 P. 与复数这样对应的平面称为复平面, 可以记作 $\mathbb{R}^2$, x 轴与实数对应, y 轴与纯虚数对应.

记 $\vec{OP}$ 为以原点 O 为起点, P 为终点的向量, 则复数 $z = x + \mathrm{i}y$ 也与平面上向量 $\vec{OP}$ 一一对应. 并且复数的加法对应于向量的加法, 复数的模对应于向量的长度.





设复数 z 对应向量 $\vec{OP}$ 的长度为 r, 从 x 轴正向半轴逆时针旋转到达 $\vec{OP}$ 的角度记作 θ. 则 z 可以写作如下三角形式

$$z = r(\cos\theta + \mathrm{i}\sin\theta).$$

在三角形式下, 更容易看出复数乘法的几何含义:

$$z_1 = r_1(\cos\theta_1 + \mathrm{i}\sin\theta_1), \qquad z_2 = r_2(\cos\theta_2 + \mathrm{i}\sin\theta_2),$$

$$z_1 \cdot z_2 = r_1 r_2(\cos(\theta_1+\theta_2) + \mathrm{i}\sin(\theta_1+\theta_2)),$$

$$\frac{z_1}{z_2} = \frac{r_1}{r_2}(\cos(\theta_1-\theta_2) + \mathrm{i}\sin(\theta_1-\theta_2)), \qquad (z_2 \neq 0).$$

z 乘以另一个复数 z' 就是把 z' 逆时针旋转 θ, 长度再乘以 r.

推论 1.4 (棣莫弗公式)

对任意 $n \in \mathbb{Z}$, $z^n = r^n(\cos n\theta + \mathrm{i}\sin n\theta)$.

对任意 $n > 0$, 记

$$\varepsilon_k := \cos\frac{2k\pi}{n} + \mathrm{i}\sin\frac{2k\pi}{n}, \quad k = 0, 1, \ldots, n-1,$$

则由棣莫弗公式有 $\varepsilon_k^n = 1$, 从而它们是方程

$$x^n - 1 = 0$$

的 n 个不同的根, 称为 n 次**单位根**.

定理 1.5 (代数基本定理)

复系数的 n 次代数方程

$$x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 = 0$$

在复数域 $\mathbb{C}$ 中至少有一个根.

注记 1.6

这里我们略去这个定理的证明. 有趣的是, 这个定理叫代数基本定理, 但是人们一开始所知道的证明方法都要用到数学分析的知识. 一个纯代数的证明参见《代数学引论（第一卷）：基础代数》, 作者: （俄罗斯）柯斯特利金的第六章, 3.3 节部分.

推论 1.7

复系数多项式

$$f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$$

可以分解为 n 个一次因式的乘积

$$f(x) = (x - \alpha_1)(x - \alpha_2)\cdots(x - \alpha_n).$$

从而任何 n 次代数方程至多有 n 个根.

于是 $\varepsilon_k = \cos\dfrac{2k\pi}{n} + \mathrm{i}\sin\dfrac{2k\pi}{n},\ (k = 0, 1, \ldots, n-1)$ 就是方程 $x^n - 1 = 0$ 的所有根.

例 1.8

求复数 $z = 1 + \cos\theta + \mathrm{i}\sin\theta\ (-\pi \leq \theta < \pi)$ 的三角形式.

例 1.9

在平面直角坐标系 Oxy 中, 求将点 $P = (x, y)$ 绕原点逆时针旋转 θ 角后得到的点 P' 的坐标.

例 1.10

在平面直角坐标系 Oxy 中, 记 $P_1 = (x_1, y_1)$, $P_2 = (x_2, y_2)$, 证明以 $\vec{OP_1}$, $\vec{OP_2}$ 为邻边的平行四边形的面积为 $|x_1y_2 - x_2y_1|$.

1.2

空间向量

▶ **向量的定义.** 空间中既有大小又有方向的量称为向量.

例如: 力, 加速度等. 记为 $\vec{OA}$ 或 $\mathbf{a}$.

向量相等: $\mathbf{a} = \mathbf{b}$. 零向量和负向量: $\mathbf{0}$, $-\mathbf{a}$.

向量的长度或模: $|\mathbf{a}|$. 单位向量: 长度为 1 的向量.

如果 $\mathbf{a}$ 和 $\mathbf{b}$ 方向相同或相反, 称它们平行, 记为 $\mathbf{a} \parallel \mathbf{b}$. 如果 $\mathbf{a}$ 和 $\mathbf{b}$ 方向互相垂直, 称它们垂直或正交, 记为 $\mathbf{a} \perp \mathbf{b}$.

向量 $\vec{OA}$ 与 $\vec{OB}$ 所夹的角 $\angle AOB$ 称为它们的夹角.

我们已见过向量的加法, 它满足如下性质:

$$\mathbf{a}+\mathbf{b}=\mathbf{b}+\mathbf{a}, \quad (\text{交换律}) \tag{1.1}$$

$$(\mathbf{a}+\mathbf{b})+\mathbf{c}=\mathbf{a}+(\mathbf{b}+\mathbf{c}), \quad (\text{结合律}) \tag{1.2}$$

$$\mathbf{a}+\mathbf{0}=\mathbf{a}, \quad \mathbf{a}+(-\mathbf{a})=\mathbf{0}. \tag{1.3}$$

向量的另一个基本运算是数与向量的乘法.

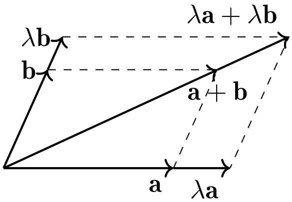
设 $\lambda \in \mathbb{R}$ 是一个实数, λ 乘以 $\mathbf{a}$ 是一个向量, 记为 $\lambda\mathbf{a}$, 其长度为 $|\lambda||\mathbf{a}|$, 当 $\lambda > 0$ 时与 $\mathbf{a}$ 同向, 当 $\lambda < 0$ 时与 $\mathbf{a}$ 反向. 称 $\lambda\mathbf{a}$ 为 λ 与 $\mathbf{a}$ 的数乘.

命题 1.11

设 $\lambda, \mu \in \mathbb{R}$ 是实数, $\mathbf{a}, \mathbf{b}$ 是向量, 则

$$\lambda(\mathbf{a}+\mathbf{b}) = \lambda\mathbf{a} + \lambda\mathbf{b}, \qquad (\lambda+\mu)\mathbf{a} = \lambda\mathbf{a} + \mu\mathbf{a}, \tag{1.4}$$

$$(\lambda\mu)\mathbf{a} = \lambda(\mu\mathbf{a}), \quad 0\mathbf{a} = \mathbf{0}, \quad 1\mathbf{a} = \mathbf{a}, \quad (-1)\mathbf{a} = -\mathbf{a}. \tag{1.5}$$



▶ 向量的共线和共面.

定义 1.12

一组向量, 如果通过平移使它们同处一条直线上, 称它们**共线**.

一组向量, 如果通过平移使它们同处一个平面上, 称它们**共面**.

可以看出, 两个向量一定共面.

命题 1.13

向量 $\mathbf{a}$, $\mathbf{b}$ 共线 $\iff$ 存在不全为 0 的实数 λ, μ 使得 $\lambda\mathbf{a}+\mu\mathbf{b}=0$.

向量 $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ 共面 $\iff$ 存在不全为 0 的实数 λ, μ, γ 使得 $\lambda\mathbf{a}+\mu\mathbf{b}+\gamma\mathbf{c}=0$.

定义 1.14

设 $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ 为一组向量, $\lambda_1, \lambda_2, \ldots, \lambda_n$ 为一组实数, 称向量

$$\lambda_1 \mathbf{a}_1 + \lambda_2 \mathbf{a}_2 + \cdots + \lambda_n \mathbf{a}_n$$

为向量 $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ 的线性组合.

定义 1.15

一组向量 $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ 称为线性相关, 如果存在不全为 0 的实数 $\lambda_1, \lambda_2, \ldots, \lambda_n$, 使得

$$\lambda_1 \mathbf{a}_1 + \lambda_2 \mathbf{a}_2 + \cdots + \lambda_n \mathbf{a}_n = 0.$$

反之就称这组向量线性无关.

换句话说, 如果 $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ 线性无关并且

$$\lambda_1\mathbf{a}_1 + \lambda_2\mathbf{a}_2 + \cdots + \lambda_n\mathbf{a}_n = 0,$$

则 $\lambda_1 = \lambda_2 = \cdots = \lambda_n = 0.$

结合命题 1.13, 我们有

- 一个向量 $\mathbf{a}$ 线性相关 $\iff \mathbf{a} = \mathbf{0}$.
- 两个向量线性相关 $\iff$ 它们共线.
- 三个向量线性相关 $\iff$ 它们共面.
- 四个向量呢?

- **仿射坐标系.** 在空间取定一组直角坐标系 $Oxyz$, 设与 x 轴, y 轴, z 轴方向相同的单位向量分别为 $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$. 则点 P 的坐标为 (x_1, x_2, x_3) 等价于 $\vec{OP} = x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k}$.

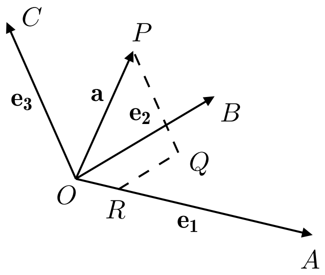
在直角坐标系下三个坐标轴两两垂直. 事实上, 坐标系可以推广到更一般的情形.

定理 1.16

设 $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ 为空间中三个不共面的向量, 则对每个向量 $\mathbf{a}$, 存在唯一的三元有序实数组 (x_1, x_2, x_3), 使得

$$\mathbf{a} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3.$$

证明:



过 P 作 OC 平行线与平面 AOB 交于点 Q, 再过 Q 作 OB 平行线与 OA 交于点 R. 于是 $\mathbf{a} = \vec{OP} = \vec{OR} + \vec{RQ} + \vec{QP}$. 由命题 1.13 知存在实数 x_1, x_2, x_3 使得

$$\vec{OR} = x_1\mathbf{e}_1, \quad \vec{RQ} = x_2\mathbf{e}_2, \quad \vec{QP} = x_3\mathbf{e}_3,$$

从而 $\mathbf{a} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3$. 下面证明唯一性. 如果

$$\mathbf{a} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3 = y_1\mathbf{e}_1 + y_2\mathbf{e}_2 + y_3\mathbf{e}_3,$$

则有

$$(x_1 - y_1)\mathbf{e}_1 + (x_2 - y_2)\mathbf{e}_2 + (x_3 - y_3)\mathbf{e}_3 = 0,$$

因为 $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ 不共面, 由命题 1.13 知 $x_1 = y_1$, $x_2 = y_2$, $x_3 = y_3$, 因此表示唯一.

推论 1.17

空间中任意四个向量 $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4$ 一定线性相关.

证明:　　如果 $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ 共面, 则 $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ 线性相关, 于是 $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4$ 也线性相关.

如果 $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ 不共面, 则由定理 1.16 知存在唯一一组实数 x_1, x_2, x_3 使得 $\mathbf{a}_4 = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + x_3\mathbf{a}_3$, 即

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + x_3\mathbf{a}_3 - \mathbf{a}_4 = 0,$$

从而 $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4$ 线性相关.

定义 1.18

空间中任意三个有序的不共面的向量 $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ 称为空间的一组**基**. 对于向量 $\mathbf{a}$, 若

$$\mathbf{a} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3,$$

称 (x_1, x_2, x_3) 为 $\mathbf{a}$ 在基 $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ 下的**仿射坐标**或简称**坐标**. 空间中任意一点 O 和一组基 $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ 合在一起称为空间的一个**仿射坐标系**, 记为 $[O; \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3]$.

引入仿射坐标系后, 下列三者存在一一对应关系:

$$\text{空间中的点 } P \longleftrightarrow \text{向量 } \vec{OP} \longleftrightarrow \text{坐标 } (x_1, x_2, x_3).$$

于是向量的运算可以化为其坐标的运算:

$$\mathbf{a} = (x_1, x_2, x_3), \quad \mathbf{b} = (y_1, y_2, y_3),$$

$$\mathbf{a} + \mathbf{b} = (x_1 + y_1, x_2 + y_2, x_3 + y_3),$$

$$\lambda\mathbf{a} = (\lambda x_1, \lambda x_2, \lambda x_3).$$

例 1.19

设 $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ 为空间的一组基. $\mathbf{a} = \mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3$, $\mathbf{b} = \mathbf{e}_1 - \mathbf{e}_2 + \mathbf{e}_3$, $\mathbf{c} = \mathbf{e}_1 + \mathbf{e}_2 - \mathbf{e}_3$

(1) 求向量 $\mathbf{d} = \mathbf{a} + 2\mathbf{b} - 3\mathbf{c}$ 在基 $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ 下的坐标.

(2) 证明 $\mathbf{a}, \mathbf{b}, \mathbf{c}$ 也是空间的一组基.

1.3

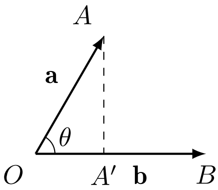
内积, 外积和混合积

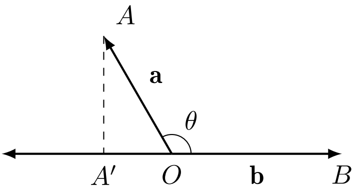
▶ **向量的内积.** 两个向量 $\mathbf{a} = \vec{OA}$ 与 $\mathbf{b} = \vec{OB}$ 的内积为一个实数, 记为 $\mathbf{a} \cdot \mathbf{b}$, 等于两个向量的模与两个向量夹角的余弦的乘积:

$$\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}||\mathbf{b}| \cos\theta.$$

设 $\vec{OA}$ 在 $\vec{OB}$ 所在直线上的投影为向量 $\vec{OA'}$, 可以看出

$$\vec{OA} \cdot \vec{OB} = \vec{OA'} \cdot \vec{OB}.$$

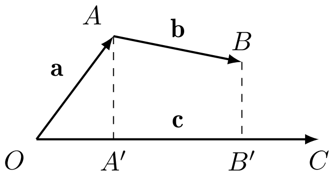




命题 1.20

对向量 $\mathbf{a}, \mathbf{b}, \mathbf{c}$ 及实数 λ, 内积满足如下性质:

1. $\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a}$,
2. $(\mathbf{a} + \mathbf{b}) \cdot \mathbf{c} = \mathbf{a} \cdot \mathbf{c} + \mathbf{b} \cdot \mathbf{c}$,
3. $(\lambda \mathbf{a}) \cdot \mathbf{b} = \lambda(\mathbf{a} \cdot \mathbf{b}) = \mathbf{a} \cdot (\lambda \mathbf{b})$,
4. $\mathbf{a} \cdot \mathbf{a} = |\mathbf{a}|^2 \geq 0$, 等号成立当且仅当 $\mathbf{a} = \mathbf{0}$.



性质 2 的证明

设 $[O;\mathbf{i},\mathbf{j},\mathbf{k}]$ 为一个直角坐标系, 由于 $\mathbf{i},\mathbf{j},\mathbf{k}$ 为两两正交的单位向量, 所以

$$\mathbf{i}\cdot\mathbf{i}=1,\quad \mathbf{j}\cdot\mathbf{j}=1,\quad \mathbf{k}\cdot\mathbf{k}=1,$$
$$\mathbf{i}\cdot\mathbf{j}=0,\quad \mathbf{i}\cdot\mathbf{k}=0,\quad \mathbf{j}\cdot\mathbf{k}=0.$$

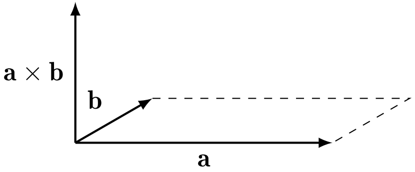
设 $\mathbf{a}=a_1\mathbf{i}+a_2\mathbf{j}+a_3\mathbf{k}$, $\mathbf{b}=b_1\mathbf{i}+b_2\mathbf{j}+b_3\mathbf{k}$, $\mathbf{a}$ 与 $\mathbf{b}$ 的夹角为 θ, 则在坐标表示下有:

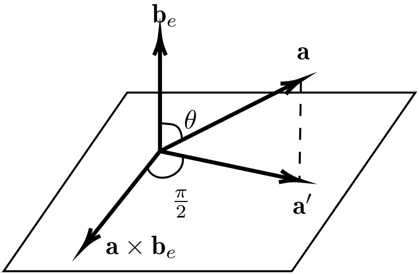
$$\mathbf{a}\cdot\mathbf{b}=(a_1,a_2,a_3)\cdot(b_1,b_2,b_3)=a_1b_1+a_2b_2+a_3b_3,$$
$$\cos\theta=\frac{\mathbf{a}\cdot\mathbf{b}}{|\mathbf{a}||\mathbf{b}|}=\frac{a_1b_1+a_2b_2+a_3b_3}{\sqrt{a_1^2+a_2^2+a_3^2}\sqrt{b_1^2+b_2^2+b_3^2}}.$$

- **向量的外积.** 　　向量 $\mathbf{a}$ 和 $\mathbf{b}$ 的外积是一个向量, 记为 $\mathbf{a}\times\mathbf{b}$, 其模为以 $\mathbf{a},\mathbf{b}$ 为邻边的平行四边形面积:

$$|\mathbf{a}\times\mathbf{b}| = |a||b|\sin\theta.$$

方向规定为: $\mathbf{a}\times\mathbf{b}$ 与 $\mathbf{a}$, $\mathbf{b}$ 均正交, 且 $\{\mathbf{a},\mathbf{b},\mathbf{a}\times\mathbf{b}\}$ 构成右手系.
设和 $\mathbf{b}$ 同方向的单位向量为 $\mathbf{b}_e$, 则 $\mathbf{a}\times\mathbf{b} = |\mathbf{b}|\mathbf{a}\times\mathbf{b}_e$.



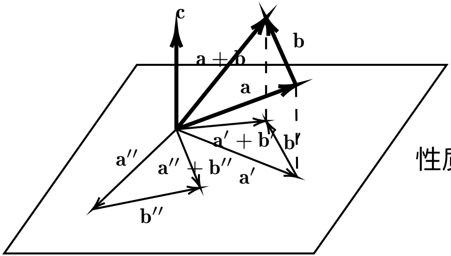


设 $\mathbf{b}_e$ 是单位向量, 求 $\mathbf{a}\times\mathbf{b}_e$ 的方法如下: 取 $\mathbf{b}_e$ 的垂直平面, 设 $\mathbf{a}$ 在该平面的投影为向量 $\mathbf{a}'$, 将 $\mathbf{a}'$ 顺时针旋转 $\frac{\pi}{2}$ 就得到 $\mathbf{a}\times\mathbf{b}_e$.

命题 1.21

向量外积满足如下性质:

1. $\mathbf{a}\times\mathbf{b}=-\mathbf{b}\times\mathbf{a}$, $\mathbf{a}\times\mathbf{b}=\mathbf{0} \iff \mathbf{a}\parallel\mathbf{b}$.
2. $(\mathbf{a}+\mathbf{b})\times\mathbf{c}=\mathbf{a}\times\mathbf{c}+\mathbf{b}\times\mathbf{c}$,
3. $(\lambda\mathbf{a})\times\mathbf{b}=\lambda(\mathbf{a}\times\mathbf{b})=\mathbf{a}\times(\lambda\mathbf{b})$.



性质 2 的证明

设 $[O;\mathbf{i},\mathbf{j},\mathbf{k}]$ 为一个直角坐标系, 则

$$\mathbf{i}\times\mathbf{i}=\mathbf{0},\quad \mathbf{j}\times\mathbf{j}=\mathbf{0},\quad \mathbf{k}\times\mathbf{k}=\mathbf{0},$$

$$\mathbf{i}\times\mathbf{j}=\mathbf{k},\quad \mathbf{j}\times\mathbf{k}=\mathbf{i},\quad \mathbf{k}\times\mathbf{i}=\mathbf{j}.$$

设 $\mathbf{a}=a_1\mathbf{i}+a_2\mathbf{j}+a_3\mathbf{k}$, $\mathbf{b}=b_1\mathbf{i}+b_2\mathbf{j}+b_3\mathbf{k}$, 则在坐标表示下有:

$$\mathbf{a}\times\mathbf{b}=(a_1,a_2,a_3)\times(b_1,b_2,b_3) \tag{1.6}$$

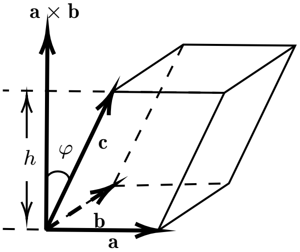
$$=(a_2b_3-a_3b_2,a_3b_1-a_1b_3,a_1b_2-a_2b_1). \tag{1.7}$$

例 1.22

求垂直于向量 $\mathbf{a}=(-1,2,1)$ 和 $\mathbf{b}=(1,0,3)$ 的单位向量.

- **向量的混合积.**　$(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}$ 称为 $\mathbf{a},\mathbf{b},\mathbf{c}$ 的混合积.

设 V 是以 $\mathbf{a},\mathbf{b},\mathbf{c}$ 为棱的平行六面体的体积. 当 φ 为锐角, $\mathbf{a},\mathbf{b},\mathbf{c}$ 构成右手系, $V=(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}$; 当 φ 为钝角, $\mathbf{a},\mathbf{b},\mathbf{c}$ 构成左手系, $V=-(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}$. 因此混合积表示以 $\mathbf{a},\mathbf{b},\mathbf{c}$ 为棱的平行六面体的 “有向体积”.



设 $\mathbf{a}=(a_1,a_2,a_3)$, $\mathbf{b}=(b_1,b_2,b_3)$, $\mathbf{c}=(c_1,c_2,c_3)$, 则

$$(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}=(a_2b_3-a_3b_2)c_1+(a_3b_1-a_1b_3)c_2+(a_1b_2-a_2b_1)c_3.$$

回忆例 1.10, 给定两个平面向量 $\mathbf{a}=(a_1,a_2)$, $\mathbf{b}=(b_1,b_2)$, 则 $a_1b_2-a_2b_1$ 是以 $\mathbf{a},\mathbf{b}$ 为邻边的平行四边形的 “有向面积”, 我们也称这个值为 $\mathbf{a},\mathbf{b}$ 的二阶行列式, 并记为

$$\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}=a_1b_2-a_2b_1,$$

从而 $$(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}=\begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix}c_1-\begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix}c_2+\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}c_3.$$

相应的, 我们也称 $(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}$ 为 $\mathbf{a},\mathbf{b},\mathbf{c}$ 的三阶行列式, 并记为

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}. \quad \text{由此, 得到公式 } \mathbf{a}\times\mathbf{b} = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ \mathbf{i} & \mathbf{j} & \mathbf{k} \end{vmatrix}.$$

推论 1.23

(1) $\mathbf{a},\mathbf{b},\mathbf{c}$ 共面 $\Longleftrightarrow$ 它们的三阶行列式为 0.

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

例 1.24

求以 $A=(1,2,3)$, $B=(2,1,4)$, $C=(1,3,5)$, $D=(3,2,1)$ 为顶点的四面体的体积.

1.4

直线与平面

▶ **直线的方程.** 本节默认在空间中取定一组直角坐标系. 设直线 L 过空间不同的点 A, B, 非零向量 $\vec{AB}$ 称为 L 的一个**方向向量**. 若已知 L 上一点 $A=(x_0,y_0,z_0)$ 和它的一个方向向量 $\mathbf{u}=(u_1,u_2,u_3)$. 对 L 上任意点 $P=(x,y,z)$, 存在 $t\in\mathbb{R}$ 使得 $\vec{AP}=t\mathbf{u}$, 从而

$$x=x_0+tu_1,\quad y=y_0+tu_2,\quad z=z_0+tu_3, \tag{1.8}$$

称为 L 的**参数方程**. 在 (1.8) 中消去 t 得到

$$\frac{x-x_0}{u_1}=\frac{y-y_0}{u_2}=\frac{z-z_0}{u_3}, \tag{1.9}$$

称为 L 的**标准方程**或**点向式方程**.

▶ **平面的方程.** 给定点 $M=(x_0,y_0,z_0)$ 和一个非零向量 $\mathbf{n}=(A,B,C)$, 存在唯一的平面 π 过点 M 并和向量 $\mathbf{n}$ 正交. 对 π 上任意点 $P=(x,y,z)$, 有 $\vec{MP}\perp\mathbf{n}$, 从而

$$A(x-x_0)+B(y-y_0)+C(z-z_0)=0, \tag{1.10}$$

称为平面的*点法式方程*, $\mathbf{n}$ 称为 π 的*法向量*. (1.10) 可简化为

$$Ax+By+Cz+D=0 \tag{1.11}$$

称为平面的*一般方程*.

例 1.25

求过 y 轴和 $M=(3,-2,7)$ 的平面方程.

空间中不共线的三个点 $M_i = (x_i, y_i, z_i)$ $(i = 0, 1, 2)$ 决定一个平面 π, $P = (x, y, z)$ 在 π 上 $\Longleftrightarrow$ $\overrightarrow{M_0M_1}, \overrightarrow{M_0M_2}, \overrightarrow{M_0P}$ 共面

$\Longleftrightarrow$

$$\begin{vmatrix} x_1 - x_0 & y_1 - y_0 & z_1 - z_0 \\ x_2 - x_0 & y_2 - y_0 & z_2 - z_0 \\ x - x_0 & y - y_0 & z - z_0 \end{vmatrix} = 0, \tag{1.12}$$

称为平面的三点式方程.

例 1.26

(1) 求过 $M_1 = (1, -2, 1)$, $M_2 = (2, 1, 0)$, $M_3 = (3, 1, 5)$ 的平面方程.

(2) 求过 $M = (1, 1, 1)$ 和直线 $x + 1 = 2y + 3 = 3z - 5$ 的平面方程.

两相交平面的联立方程

$$\begin{cases} A_1x + B_1y + C_1z + D_1 = 0, \\ A_2x + B_2y + C_2z + D_2 = 0. \end{cases} \tag{1.13}$$

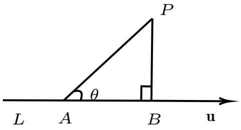
决定了它们的相交直线 L, 称为 L 的**一般方程**.

例 1.27

已知直线 L 的一般方程 $\begin{cases} 2x - 3y - z - 3 = 0, \\ 4x - 6y + 5z + 1 = 0. \end{cases}$. 求它的标准方程和参数方程.

- **点到直线的距离.** 设直线 L 过点 A, 方向向量为 $\mathbf{u}$, P 为空间任意一点. 过点 P 作 L 的垂线, 垂足为 B. 点 P 到 L 的距离为

$$d=\left|\vec{BP}\right|=\left|\vec{AP}\right|\sin\theta=\frac{\left|\vec{AP}\times\mathbf{u}\right|}{|\mathbf{u}|}.$$

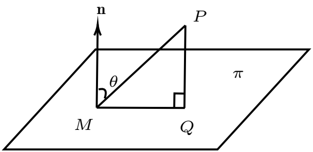


例 1.28

求 $P=(1,1,2)$ 到直线 $\frac{x-2}{2}=\frac{y-3}{1}=\frac{z}{2}$ 的距离.

- **点到平面的距离.** 设平面 π 的一般方程为 $Ax+By+Cz+D=0$, 法向量为 $\mathbf{n}=(A,B,C)$. 在平面上任取一点 M, P 为空间中任意点, 过 P 作平面 π 的垂线, 垂足为 Q. 则 P 到 π 的距离为

$$d=\left|\vec{QP}\right|=\frac{\left|\mathbf{n}\cdot\vec{MP}\right|}{|\mathbf{n}|}=\frac{|A(x-x_0)+B(y-y_0)+C(z-z_0)|}{\sqrt{A^2+B^2+C^2}}$$
$$=\frac{|Ax+By+Cz+D|}{\sqrt{A^2+B^2+C^2}}.$$



例 1.29

求点 $(1,1,1)$ 到平面 $2x+3y+6z=18$ 的距离.

- **两直线的位置关系.** 设 L_1, L_2 上一点分别为 M_1, M_2, 方向向量分别为 $\mathbf{u}_1, \mathbf{u}_2$, 则

1. L_1 和 L_2 共面 $\Longleftrightarrow$ 向量 $\overrightarrow{M_1M_2}, \mathbf{u}_1, \mathbf{u}_2$ 的行列式为 0.
2. L_1 和 L_2 异面 $\Longleftrightarrow$ 向量 $\overrightarrow{M_1M_2}, \mathbf{u}_1, \mathbf{u}_2$ 的行列式非 0.
3. 两直线的夹角的余弦为 $\cos\theta = \dfrac{|\mathbf{u}_1 \cdot \mathbf{u}_2|}{|\mathbf{u}_1||\mathbf{u}_2|}$.

例 1.30

求 a 使得直线 $x-a=y-2=z-3$ 和 $x=2y=3z$ 相交.

▶ **两平面的位置关系.** 设平面 π_1, π_2 的一般方程分别为 $A_i x + B_i y + C_i z + D_i = 0\ (i = 1, 2)$, 法向量为 $\mathbf{n}_1, \mathbf{n}_2$, 则

1. π_1 和 π_2 平行 $\Longleftrightarrow \dfrac{A_1}{A_2} = \dfrac{B_1}{B_2} = \dfrac{C_1}{C_2} \neq \dfrac{D_1}{D_2}$.
2. π_1 和 π_2 重合 $\Longleftrightarrow \dfrac{A_1}{A_2} = \dfrac{B_1}{B_2} = \dfrac{C_1}{C_2} = \dfrac{D_1}{D_2}$.
3. π_1 和 π_2 垂直 $\Longleftrightarrow A_1A_2 + B_1B_2 + C_1C_2 = 0$.
4. 两平面的夹角的余弦为 $\cos\theta = \dfrac{|\mathbf{n}_1 \cdot \mathbf{n}_2|}{|\mathbf{n}_1||\mathbf{n}_2|}$.

- **直线与平面的关系.**　设直线 L 的方向向量为 $\mathbf{u}$, 平面 π 的法向量为 $\mathbf{n}$, 则 L 与 π 的夹角的正弦为

$$\sin\theta = \frac{|\mathbf{u}\cdot\mathbf{n}|}{|\mathbf{u}||\mathbf{n}|}.$$

例 1.31

求直线 $x = 2y = 3z$ 和平面 $x + 2y + 3z = 4$ 的夹角和交点.

- **平面束方程.** 设 L 是平面

$$\pi_i: A_ix+B_iy+C_iz+D_i=0,\quad i=1,2$$

的相交直线. 则过 L 的任意平面可以写为

$$\lambda(A_1x+B_1y+C_1z+D_1)+\mu(A_2x+B_2y+C_2z+D_2)=0,$$

其中 λ,μ 不同时为 0, 称为平面束方程.

例 1.32

求直线 $\begin{cases} x-y+z+2=0, \\ x+2y-z+5=0. \end{cases}$ 在平面 $2x+y-3z+1=0$ 上的投影直线的方程.

第一章复习题:

例 1.33

判断向量 $(-2,3,10)$, $(-1,4,15)$, $(3,6,-15)$ 是否共面.

例 1.34

设 $A=(1,4,-2)$, $B=(-1,5,0)$, 求 $\vec{AB}$ 在 x 轴, y 轴, z 轴上的投影, 并求 $\left|\vec{AB}\right|$ 和 $\vec{AB}$ 的单位向量.

例 1.35

在直角坐标系 $[O;\mathbf{i},\mathbf{j},\mathbf{k}]$ 中, 向量 $\mathbf{a}$ 与 $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ 的夹角 α,β,γ 称为 $\mathbf{a}$ 的**方向角**, $\cos\alpha,\cos\beta,\cos\gamma$ 称为 $\mathbf{a}$ 的**方向余弦**. 证明若 $\mathbf{a}=x\mathbf{i}+y\mathbf{j}+z\mathbf{k}$, 则 $\cos\alpha=\dfrac{x}{|a|}$, $\cos\beta=\dfrac{y}{|a|}$, $\cos\gamma=\dfrac{z}{|a|}$.

1.5

高维向量

现实生活中，很多量无法在三维空间中表示，例如质点在一段时间内的运动轨迹，物体某一点的速度等. 所以有必要将三维向量推广到高维向量.

取定一个坐标系，一个空间向量可以用一个三元数组 (坐标) 来表示. 向量的加法和数乘都可以转化为坐标运算. 因此将三维向量推广到高维的一个直接方法是通过多元数组来定义.

定义 1.36

设 $\mathbb{F}$ 是一个数域，一个 **n 维 (行) 向量** $\mathbf{a}$ 是一个 n 元有序数组

$$\mathbf{a} = (a_1, a_2, \ldots, a_n), \quad a_i \in \mathbb{F}.$$

所有 n 维向量构成的集合称为 $\mathbb{F}$ 上的 n 维**向量空间**，记为 $\mathbb{F}^n$.

在 $\mathbb{F}^n$ 上定义零向量, 负向量, 向量加法和数乘:

$$\mathbf{a}+\mathbf{b}=(a_1+b_1, a_2+b_2, \ldots, a_n+b_n), \tag{1.14}$$

$$\lambda\mathbf{a}=(\lambda a_1, \lambda a_2, \ldots, \lambda a_n), \quad \lambda\in\mathbb{F} \tag{1.15}$$

$$\mathbf{0}=(0,0,\ldots,0), \quad -\mathbf{a}=(-a_1,-a_2,\ldots,-a_n). \tag{1.16}$$

容易验证空间向量情形的性质 (1.1)–(1.5) 仍然成立.

向量有时需要写成列的形式, 如 $\mathbf{a}=\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$, 称为**列向量**. 为了方便, 有时我们用 $[a_1,\ldots,a_n]$ 表示列向量.

定义 1.37

给定一组 n 维向量 $\mathbf{a}_1,\ldots,\mathbf{a}_m$ 及 $\mathbb{F}$ 中一组数 $\lambda_1,\ldots,\lambda_m$, 称

$$\lambda_1\mathbf{a}_1+\cdots+\lambda_m\mathbf{a}_m$$

为 $\mathbf{a}_1,\ldots,\mathbf{a}_m$ 的线性组合. 如果 $\mathbf{b}$ 可以表示成 $\mathbf{a}_1,\ldots,\mathbf{a}_m$ 的一个线性组合, 称 $\mathbf{b}$ 可以由 $\mathbf{a}_1,\ldots,\mathbf{a}_m$ 线性表示.

给定一组 n 维向量 $\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_m$, 如果存在一组不全为 0 的数 $\lambda_1,\ldots,\lambda_m$ 使得

$$\lambda_1\mathbf{a}_1+\cdots+\lambda_m\mathbf{a}_m=0,$$

称 $\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_m$ 线性相关. 否则称为线性无关.

线性代数的研究对象: n 维向量空间, 相关运算及映射.

其核心思想: 用代数工具研究几何问题.

几何应用: “平行六面体” 的有向体积的定义和计算 (§1), 直线与平面 (§3), 向量组的线性相关 (§4), 内积, 距离, 正交 (§7.1, §9.6), 曲线与曲面 (§8), 坐标变换 (§9.1-9.5).

代数工具:

1. 线性方程组 (§5):
   例如一个三元一次方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1, \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2, \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3. \end{cases} \tag{1.17}$$

   一方面方程的求解等价于求三个平面的交点.

另一方面, 考虑列向量

$$\mathbf{a}_1=\begin{pmatrix}a_{11}\\a_{21}\\a_{31}\end{pmatrix},\quad \mathbf{a}_2=\begin{pmatrix}a_{12}\\a_{22}\\a_{32}\end{pmatrix},\quad \mathbf{a}_3=\begin{pmatrix}a_{13}\\a_{23}\\a_{33}\end{pmatrix},\quad \mathbf{b}=\begin{pmatrix}b_1\\b_2\\b_3\end{pmatrix},$$

则方程组可以写成

$$x_1\mathbf{a}_1+x_2\mathbf{a}_2+x_3\mathbf{a}_3=\mathbf{b},$$

等价于求解 $\mathbf{b}$ 是否可以写成 $\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_3$ 的线性组合.

2. 矩阵 (§2, §6, §7.2-7.5): 解线性方程组, 计算行列式, 化简坐标变换都可以化为矩阵运算!

2 线性方程组

消元法

$\mathbb{F}^m$ 的子空间

矩阵的秩

2.1

消元法

我们希望求得如下线性方程组的所有解:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2m}x_m = b_2 \\ \qquad\qquad\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nm}x_m = b_n \end{cases} \tag{2.1}$$

其中 $a_{ij}, b_i \in \mathbb{F}, 1 \leq i \leq n, 1 \leq j \leq m$, 并且对任意 $1 \leq j \leq m$, 存在 i 使得 $a_{ij} \neq 0$.

称 $\{x_i\}_{1\leq i\leq m}$ 为未知元或变量, $\{a_{ij}\}_{1\leq i\leq n}^{1\leq j\leq m}$ 为方程组的系数, $\{b_i\}_{1\leq i\leq n}$ 为常数项.

$\{s_1, s_2, \ldots, s_m, s_i \in \mathbb{F}\}$ 为方程组 (2.1) 的一个**解**, 如果 x_i 取值为 s_i 时方程组成为恒等式.

可以用消元法将方程组化为简单的形式来求解, 我们先看一个例子.

例 2.1

求解线性方程组

$$\begin{cases} x_1 + 2x_2 + 3x_3 = 1 & (1) \\ 2x_1 + 2x_2 + 5x_3 = 2 & (2) \\ 3x_1 + 5x_2 + x_3 = 3 & (3) \end{cases}$$

$$\begin{cases} x_1 + 2x_2 + 3x_3 = 1 \\ 2x_1 + 2x_2 + 5x_3 = 2 \\ 3x_1 + 5x_2 + x_3 = 3 \end{cases} \xrightarrow[(1)\times -3+(3)]{(1)\times -2+(2)} \begin{cases} x_1 + 2x_2 + 3x_3 = 1 \\ -2x_2 - x_3 = 0 \\ -x_2 - 8x_3 = 0 \end{cases}$$

$$\xrightarrow{(3)\times -2+(2)} \begin{cases} x_1 + 2x_2 + 3x_3 = 1 \\ 15x_3 = 0 \\ -x_2 - 8x_3 = 0 \end{cases} \xrightarrow[(3)\times \frac{1}{15}]{\text{交换 (2) 和 (3)}}$$

$$\begin{cases} x_1 + 2x_2 + 3x_3 = 1 \\ -x_2 - 8x_3 = 0 \\ x_3 = 0 \end{cases} \Longrightarrow$$

$x_3 = 0, x_2 = 0, x_1 = 1$

是方程组的解.

例 2.2

求解线性方程组

$$\begin{cases} x_1 - x_2 + x_3 = 1 \\ x_1 - x_2 - x_3 = 3 \\ 2x_1 - 2x_2 - x_3 = 3 \end{cases}$$

$$\begin{cases} x_1 + 2x_2 + 3x_3 + 4x_4 = -3 \\ x_1 + 2x_2 - 5x_4 = 1 \\ 2x_1 + 4x_2 - 3x_3 - 19x_4 = 6 \\ 3x_1 + 6x_2 - 3x_3 - 24x_4 = 7 \end{cases}$$

上面所做的消元法可以归纳为如下三类行变换，它们称为**初等行变换**.

1. 交换两行的位置.
2. 将某行乘以一个非零常数.
3. 将某行的一个倍数加到另一行.

定理 2.3

原线性方程组和经过行变换后的线性方程组有相同的解. 也称经过行变换所得到的方程组和原方程组是**等价**的.

接下来给出一般线性方程组的消元算法.

定理 2.4

可以通过初等行变换将方程组 (2.1) 化为如下**阶梯形**:

$$\begin{cases} a'_{11}x_1 + \cdots + 0 + \cdots + 0 + \cdots + a'_{1m}x_m = b'_1 \\ \qquad a'_{2s_2}x_{s_2} + \cdots + 0 + \cdots + a'_{2m}x_m = b'_2 \\ \qquad\qquad \ddots \qquad\qquad \vdots \\ \qquad\qquad\qquad a'_{rs_r}x_{s_r} + \cdots + a'_{rm}x_m = b'_r \\ \qquad\qquad\qquad\qquad 0 = b'_{r+1} \\ \qquad\qquad\qquad\qquad \vdots \\ \qquad\qquad\qquad\qquad 0 = b'_n \end{cases} \tag{2.2}$$

其中 $1 < s_2 < \cdots < s_r$, 并且 $a'_{11}, a'_{2s_2}, \ldots, a'_{rs_r}$ 均不为 0.

命题 2.5 (方程组相容性判定)

1. 方程组 (2.2) 有解当且仅当 $b'_{r+1} = \cdots = b'_n = 0$.
2. 方程组 (2.2) 有唯一解当且仅当 $r = m$ 并且 $b'_{r+1} = \cdots = b'_n = 0$.
3. 方程组 (2.2) 有无穷多解当且仅当 $r < m$ 并且 $b'_{r+1} = \cdots = b'_n = 0$.

称 $\{x_1, x_{s_2}, \ldots, x_{s_r}\}$ 为主未知元, 其余变量为自由变量. 用行变换将方程组化为阶梯形求解的方法称为消元法. 如果一个方程组有解, 称它是相容的或可解的, 否则称为是不相容的.

若 (2.2) 是相容的, 自由变量可以在 $\mathbb{F}$ 中任意取值, 并且主未知元 $x_1, x_{s_2}, \ldots, x_{s_r}$ 随之唯一确定. 当自由变量均取 0 时,

$$x_1 = \frac{b'_1}{a'_{11}}, x_{s_2} = \frac{b'_2}{a'_{2s_2}}, \ldots, x_{s_r} = \frac{b'_r}{a'_{rs_r}}, \tag{2.3}$$

这个解称为方程组 (2.1) 的一个特解.

一般的, 当自由变量取为 $\lambda_1, \ldots, \lambda_{m-r}$, 则线性方程组 (2.1) 的通解可以写成

$$\begin{cases} x_1 = c_{11}\lambda_1 + \cdots + c_{1m-r}\lambda_{m-r} + d_1, \\ x_2 = c_{21}\lambda_1 + \cdots + c_{2m-r}\lambda_{m-r} + d_2, \\ \quad \vdots \\ x_m = c_{m1}\lambda_1 + \cdots + c_{mm-r}\lambda_{m-r} + d_m, \end{cases} \tag{2.4}$$

回忆 $\mathbb{F}^m$ 是所有 m 维 (列) 向量构成的集合. 记

$$\mathbf{x}=\begin{pmatrix}x_1\\ \vdots\\ x_m\end{pmatrix},\ \mathbf{c}_1=\begin{pmatrix}c_{11}\\ \vdots\\ c_{m1}\end{pmatrix},\cdots,\ \mathbf{c}_{m-r}=\begin{pmatrix}c_{1m-r}\\ \vdots\\ c_{mm-r}\end{pmatrix},\ \mathbf{d}=\begin{pmatrix}d_1\\ \vdots\\ d_m\end{pmatrix}, \tag{2.5}$$

则通解 (2.4) 可以写成向量的线性组合形式:

$$\mathbf{x}=\lambda_1\mathbf{c}_1+\cdots+\lambda_{m-r}\mathbf{c}_{m-r}+\mathbf{d},\ \ \lambda_1,\ldots,\lambda_{m-r}\in\mathbb{F}.$$

其中 $\mathbf{d}$ 是 (2.3) 中的特解. 也称一组解 $\mathbf{x}$ 为**解向量**.

命题 2.6

(2.5) 中的 $\mathbf{c}_1,\ldots,\mathbf{c}_{m-r}$ 线性无关.

与方程组 (2.1) 相关联的是如下齐次线性方程组, 即常数项 $\{b_i\}_{1\le i\le n}$ 均为 0 的线性方程组:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2m}x_m = 0 \\ \qquad\qquad\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nm}x_m = 0 \end{cases} \tag{2.6}$$

这个方程组是一定有解的, 比如零解, 记它的所有解的集合, 简称解集, 为 Ω_0. 若方程组 (2.1) 有解, 且解集为 Ω, 则

命题 2.7

$\Omega_0 = \{\lambda_1\mathbf{c}_1 + \cdots + \lambda_{m-r}\mathbf{c}_{m-r},\ \ \lambda_i \in \mathbb{F}\}$, 并且有双射 $f : \Omega_0 \longleftrightarrow \Omega$, 其中 $f(\mathbf{x}) = \mathbf{x} + \mathbf{d}$.

2.2

$\mathbb{F}^m$ 的子空间

上一节利用消元法, 我们给出了齐次线性方程组 (2.6) 的解集 Ω_0 的如下描述: 存在一组线性无关的解向量 $\mathbf{c}_1, \ldots, \mathbf{c}_{m-r}$, 也称为 (2.6) 的一组**基础解系**, 使得

$$\Omega_0 = \{\lambda_1 \mathbf{c}_1 + \cdots + \lambda_{m-r}\mathbf{c}_{m-r},\ \ \lambda_i \in \mathbb{F}\}.$$

但这个描述依赖于我们做消元法 (行变换) 的过程. 我们期待能够直接从原方程组给出 Ω_0 的描述, 例如

1. 如何直接确定 r? r 是否唯一?
2. 解集 Ω_0 的大小和 r 的大小有何关系?

命题 2.8

对 $\forall\, \mathbf{a}, \mathbf{b} \in \Omega_0$ 及 $\forall\, \lambda, \mu \in \mathbb{F}$, $\lambda\mathbf{a} + \mu\mathbf{b} \in \Omega_0$.

定义 2.9

设 $V \subset \mathbb{F}^m$ 是由一些 m 维向量构成的集合, 如果 V 满足

$$\mathbf{a}, \mathbf{b} \in V,\ \ \lambda, \mu \in \mathbb{F} \Longrightarrow \lambda\mathbf{a} + \mu\mathbf{b} \in V,$$

称 V 是 $\mathbb{F}^m$ 的**子空间**, 也称为 $\mathbb{F}$ 上的**向量空间**.

从而齐次线性方程组 (2.6) 的解集 Ω_0 为 $\mathbb{F}^m$ 的子空间, 称为**解空间**.

定义 2.10

设 $\boldsymbol{\epsilon}_i \in \mathbb{F}^m$ 是第 i 个坐标为 1, 其余坐标为 0 的向量, 则任何

$$\mathbf{a} = (a_1, \dots, a_m) \in \mathbb{F}^m$$

可以表示为 $\{\boldsymbol{\epsilon}_1, \dots, \boldsymbol{\epsilon}_m\}$ 的线性组合

$$\mathbf{a} = a_1\boldsymbol{\epsilon}_1 + \cdots + a_m\boldsymbol{\epsilon}_m,$$

并且这个表示是唯一的. 称 $\{\boldsymbol{\epsilon}_1, \dots, \boldsymbol{\epsilon}_m\}$ 为 $\mathbb{F}^m$ 的标准基.

接下来的问题是对于任意 $\mathbb{F}^m$ 的子空间 V, 是否可以找到它的一个有限子集 $B \subset V$, 使得 V 中任意向量都可以表示为 B 中向量的线性组合, 并且这个表示是唯一的.

引理 2.11

设 $\mathbf{a}_1, \ldots, \mathbf{a}_k \in \mathbb{F}^m$, 则如下两个命题等价:

1. $\mathbf{a}_1 \neq 0$, 且对 $\forall\ \ell = 2, 3, \ldots, k$, $\mathbf{a}_\ell$ 不能表示成 $\mathbf{a}_1, \ldots, \mathbf{a}_{\ell-1}$ 的线性组合.
2. $\mathbf{a}_1, \ldots, \mathbf{a}_k$ 线性无关.

定理 2.12

设 V 是 $\mathbb{F}^m$ 的子空间, 且 $V \neq \{0\}$. 任取 $\mathbf{a}_1 \neq 0 \in V$, 然后取 $\mathbf{a}_2 \in V$ 不能表示为 $\mathbf{a}_1$ 的线性组合, 再取 $\mathbf{a}_3 \in V$ 不能表示为 $\mathbf{a}_1, \mathbf{a}_2$ 的线性组合..... 则在 m 步内该操作就必须停止, 即存在 $1 \leq r \leq m$ 和 $\mathbf{a}_1, \ldots, \mathbf{a}_r$, 使得 V 中任何向量都可以表示成 $\mathbf{a}_1, \ldots, \mathbf{a}_r$ 的线性组合.

命题 2.13

记 $\{\mathbf{a}_1, \ldots, \mathbf{a}_r\}$ 为定理 2.12 中方式取到的 V 中的向量组, 则 V 中任意向量表示成此向量组的线性组合的方式唯一. 换句话说, $\{\mathbf{a}_1, \ldots, \mathbf{a}_r\}$ 是 V 的一组基.

定义 2.14

设 V 是 $\mathbb{F}^m$ 的子空间, $T = \{\mathbf{a}_1, \ldots, \mathbf{a}_k\}$ 是 V 中线性无关的向量组. 如果在 T 中添加任意 V 中向量就变为线性相关向量组, 称 T 是 V 的一个**极大线性无关向量组**.

推论 2.15

V 中任何一组线性无关的向量组均可扩充为 V 的一个极大线性无关向量组.

V 中有着很多不同的极大线性无关向量组. 记 $|U|$ 是有限集合 U 中元素的个数, 或者 U 的势, 则有

命题 2.16

设 T_1, T_2 是 V 的两个极大线性无关向量组, 则 $|T_1| = |T_2|$.

从而一个向量空间 V 的极大线性无关向量组的势是 V 的不变量, 称为是 V 的维数, 记为 $\dim V$. V 的一组极大线性无关向量组 $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ 也称为 V 的一组基, V 中任意向量 $\mathbf{a}$ 都可以唯一表示为一组基的线性组合:

$$\mathbf{a} = \lambda_1 \mathbf{v}_1 + \cdots + \lambda_r \mathbf{v}_r, \ \lambda_1, \ldots, \lambda_r \in \mathbb{F},$$

$(\lambda_1, \ldots, \lambda_r)$ 为 $\mathbf{a}$ 在基 $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ 下的坐标表示.

定义 2.17

设 $T=\{\mathbf{a}_1,\dots,\mathbf{a}_k\}$ 是 $\mathbb{F}^m$ 中向量组. 若 $\mathbf{a}_{i_1},\dots,\mathbf{a}_{i_r}$ 是 T 中线性无关向量, 且任取 T 中向量 $\mathbf{a}_{i_{r+1}}$ 后, $\mathbf{a}_{i_1},\dots,\mathbf{a}_{i_r},\mathbf{a}_{i_{r+1}}$ 线性相关, 就称 $\{\mathbf{a}_{i_1},\dots,\mathbf{a}_{i_r}\}$ 是 T 的极大线性无关向量组.

定义 2.18

设 $T=\{\mathbf{a}_1,\dots,\mathbf{a}_k\}$ 是 $\mathbb{F}^m$ 中向量组, 称如下 $\mathbb{F}^m$ 的子空间

$$\operatorname{span} T=\{\lambda_1\mathbf{a}_1+\cdots+\lambda_k\mathbf{a}_k,\lambda_i\in\mathbb{F},1\le i\le k\}$$

为 T 中向量生成的向量空间. $\operatorname{span} T$ 的维数也称为向量组 T 的秩, 记为 $\operatorname{rank} T$.

命题 2.19

设 $T = \{\mathbf{a}_1, \ldots, \mathbf{a}_k\}$ 是 $\mathbb{F}^m$ 中向量组, 则 T 的任一极大线性无关向量组也是 $\operatorname{span} T$ 的一组极大线性无关向量组. 从而 T 的任一极大线性无关向量组中向量的个数等于 $\operatorname{rank} T$.

例 2.20

在 $\mathbb{F}^3$ 中, 设 $V = \{(x_1, x_2, x_3) \in \mathbb{F}^3 \mid x_1 + x_2 + x_3 = 0\}$. 则 V 是 $\mathbb{F}^3$ 的子空间, 并求它的一组基和维数.

例 2.21

计算向量组 $(3, 4, -2, 5)$, $(2, -5, 0, -3)$, $(5, 0, -1, 2)$, $(3, 3, -3, 5)$ 的秩, 并写出它的一组极大线性无关向量组.

回到齐次线性方程组 (2.6), 设 Ω_0 是它的解空间. 化为阶梯形

$$\begin{cases} a'_{11}x_1 + \cdots + 0 + \cdots + 0 + \cdots + a'_{1m}x_m = 0 \\ \qquad a'_{2s_2}x_{s_2} + \cdots + 0 + \cdots + a'_{2m}x_m = 0 \\ \qquad\qquad \ddots \qquad\qquad \vdots \\ \qquad\qquad\qquad a'_{rs_r}x_{s_r} + \cdots + a'_{rm}x_m = 0 \end{cases} \tag{2.7}$$

后, m 为未知元的个数, r 是主未知元的个数, 自由变量的个数就是解空间 Ω_0 的维数, 从而

$$r + \dim \Omega_0 = m. \tag{2.8}$$

(2.8) 说明了 r 不依赖于行变换的过程, 以及它和解空间 Ω_0 大小的联系.

2.3

矩阵的秩

因为消元法解线性方程组 (2.6) 的过程仅和系数 a_{ij} 相关, 而与未知元无关, 我们将线性方程组 (2.6) 的系数提取出来写成如下矩阵的形式:

$$\mathbf{A}=(a_{ij})_{n\times m}=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix}.$$

$\mathbf{A}$ 称为方程组 (2.6) 的系数矩阵. n 是 $\mathbf{A}$ 的行数, m 是 $\mathbf{A}$ 的列数, a_{ij} 是 $\mathbf{A}$ 中的元素. 因为 $a_{ij}\in\mathbb{F}$, $\mathbf{A}$ 称为 $\mathbb{F}$ 上 $n\times m$ 阶矩阵. 矩阵 $\mathbf{A}$ 包含了方程组 (2.6) 的所有信息.

矩阵 $\mathbf{A}$ 的每一行构成 $\mathbb{F}^m$ 中的行向量

$$\mathbf{a}_i = (a_{i1}, a_{i2}, \ldots, a_{im}),\ \ i = 1, 2, \ldots, n$$

利用三类初等行变换对方程组 (2.6) 做消元法的过程等价于对矩阵 $\mathbf{A}$ 做类似的初等行变换.

定义 2.22

矩阵 A 的所有行向量构成的向量组的秩称为矩阵的**行秩**, 记为 $\operatorname{rrank} \mathbf{A}$.

命题 2.23

初等行变换不改变矩阵的行秩.

设方程组 (2.6) 经过初等行变换化为阶梯形 (2.2), 设阶梯形的系数矩阵为 $\mathbf{A}'$. 则由命题 2.23 知

$$\operatorname{rrank}\ \mathbf{A} = \operatorname{rrank}\ \mathbf{A}',$$

而 $\operatorname{rrank}\ \mathbf{A}'$ 就是阶梯形中主未知元的个数 r, 从而

$$\operatorname{rrank}\ \mathbf{A} + \dim(\Omega_0) = m. \tag{2.9}$$

$\operatorname{rrank}\ \mathbf{A}$ 给出了阶梯形中出现的 r 的确切含义: 方程组中 “真实” 的方程个数 (如果一个方程可以由其他方程线性表示, 那么这个方程就是多余的). (2.9) 说明, 线性方程组中每加上一个 “真实” 的方程, 就对解空间加了一个约束, 解空间就变小了一维.

另一方面, 记 $\mathbf{A}$ 的所有列为 $\mathbb{F}^n$ 中的列向量

$$\mathbf{b}_j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{pmatrix}, \; j = 1, 2, \ldots, m$$

则

$$\mathbf{x} = [x_1, \ldots, x_m] \in \Omega_0 \Longleftrightarrow x_1\mathbf{b}_1 + \cdots + x_m\mathbf{b}_m = 0.$$

不妨设 $\{\mathbf{b}_1, \ldots, \mathbf{b}_s\}$ 是列向量组 $\{\mathbf{b}_1, \ldots, \mathbf{b}_m\}$ 的极大线性无关向量组.

从而 $\{\mathbf{b}_{s+1},\dots,\mathbf{b}_m\}$ 可以由 $\{\mathbf{b}_1,\dots,\mathbf{b}_s\}$ 线性表出:

$$\begin{aligned}\mathbf{b}_{s+1} &= \lambda_{1s+1}\mathbf{b}_1+\lambda_{2s+1}\mathbf{b}_2+\cdots+\lambda_{ss+1}\mathbf{b}_s,\\ \mathbf{b}_{s+2} &= \lambda_{1s+2}\mathbf{b}_1+\lambda_{2s+2}\mathbf{b}_2+\cdots+\lambda_{ss+2}\mathbf{b}_s,\\ &\vdots\\ \mathbf{b}_m &= \lambda_{1m}\mathbf{b}_1+\lambda_{2m}\mathbf{b}_2+\cdots+\lambda_{sm}\mathbf{b}_s.\end{aligned}\tag{2.10}$$

接下来将 (2.10) 代入 $x_1\mathbf{b}_1+\cdots+x_m\mathbf{b}_m=0$, 得到

$$x_1\mathbf{b}_1+\cdots+x_m\mathbf{b}_m=0 \Longleftrightarrow y_1\mathbf{b}_1+\cdots+y_s\mathbf{b}_s=0,$$

其中

$$
\begin{aligned}
y_1 &= x_1 + \lambda_{1s+1}x_{s+1} + \lambda_{1s+2}x_{s+2} + \cdots + \lambda_{1m}x_m,\\
y_2 &= x_2 + \lambda_{2s+1}x_{s+1} + \lambda_{2s+2}x_{s+2} + \cdots + \lambda_{2m}x_m,\\
&\ \vdots\\
y_s &= x_s + \lambda_{ss+1}x_{s+1} + \lambda_{ss+2}x_{s+2} + \cdots + \lambda_{sm}x_m.
\end{aligned}
\tag{2.11}
$$

因为 $\{\mathbf{b}_1, \ldots, \mathbf{b}_s\}$ 线性无关, $y_1\mathbf{b}_1 + \cdots + y_s\mathbf{b}_s = 0 \Longleftrightarrow$

$$
\begin{cases}
x_1 + \lambda_{1s+1}x_{s+1} + \lambda_{1s+2}x_{s+2} + \cdots + \lambda_{1m}x_m = 0\\
x_2 + \lambda_{2s+1}x_{s+1} + \lambda_{2s+2}x_{s+2} + \cdots + \lambda_{2m}x_m = 0\\
\qquad\qquad\qquad\qquad\vdots\\
x_s + \lambda_{ss+1}x_{s+1} + \lambda_{ss+2}x_{s+2} + \cdots + \lambda_{sm}x_m = 0
\end{cases}
\tag{2.12}
$$

而齐次线性方程组 (2.12) 本身就是阶梯型的, 它的解空间也是 Ω_0, 从而 $s + \dim\,\Omega_0 = m$, 但是由定义 s 就是 $\mathbf{A}$ 的列向量组的秩, 也称为矩阵 $\mathbf{A}$ 的列秩, 记为 $\mathrm{crank}\ \mathbf{A}$, 从而

$$\mathrm{crank}\ \mathbf{A} + \dim\,\Omega_0 = m.$$

结合 (2.9), 我们得到

定理 2.24

任何矩阵 $\mathbf{A}$ 的行秩等于列秩, 统称为矩阵的秩, 记为 $\mathrm{rank}\ \mathbf{A}$. 设 m 是矩阵的列数, Ω_0 是系数矩阵为 $\mathbf{A}$ 的齐次线性方程组的解空间, 则

$$\mathrm{rank}\ \mathbf{A} + \dim\,\Omega_0 = m.$$

设 $\mathbf{A} = (a_{ij})_{n\times m}$, 设 $\mathbf{A}^{\mathrm{T}} = (a'_{k\ell})_{m\times n}$, 使得 $a'_{ji} = a_{ij}$, 称为矩阵 $\mathbf{A}$ 的**转置**, 也记为 $\mathbf{A}^{\mathrm{T}} = (a_{ji})_{m\times n}$. 由定理 2.24 得到 $\operatorname{rank}(\mathbf{A}) = \operatorname{rank}(\mathbf{A}^{\mathrm{T}})$.

证明 $\operatorname{rrank} \mathbf{A} = \operatorname{crank} \mathbf{A}$ 的另一种方法是先证明如下命题:

命题 2.25

初等行变换不改变矩阵的列秩. 换句话说, 设 $\mathbf{A} = (a_{ij})_{n\times m}$ 是矩阵, 列向量为 $\mathbf{b}_1, \ldots, \mathbf{b}_m$. $\mathbf{A}'$ 是对 $\mathbf{A}$ 做某个初等行变换后得到的矩阵, 列向量为 $\mathbf{b}'_1, \ldots, \mathbf{b}'_m$. 则对任意 $1 \le i_1 < i_2 < \cdots < i_t \le m$,

$$\mathbf{b}_{i_1}, \ldots, \mathbf{b}_{i_t} \text{ 线性无关} \iff \mathbf{b}'_{i_1}, \ldots, \mathbf{b}'_{i_t} \text{ 线性无关}.$$

当矩阵 $\mathbf{A}$ 通过行变换化为阶梯形

$$\mathbf{A}' = \begin{pmatrix} a'_{11} & \cdots & 0 & \cdots & 0 & \cdots & a'_{1m} \\ & & a'_{2s_2} & \cdots & 0 & \cdots & a'_{2m} \\ & & & \ddots & \vdots & & \vdots \\ & & & & a'_{rs_r} & \cdots & a'_{rm} \\ & & & & & & \end{pmatrix},$$

其中 $a'_{11}, \ldots, a'_{rs_r} \neq 0$. 可以看出 $\mathbf{A}'$ 的第 $1, s_2, \ldots, s_r$ 列即为 $\mathbf{A}'$ 的列向量组的极大线性无关向量组.

命题 (2.25) 给出了求向量组 $T = \{\mathbf{a}_1, \ldots, \mathbf{a}_k\}$ 的秩及其极大线性无关向量组的矩阵算法:

1. 将 $\mathbf{a}_1, \ldots, \mathbf{a}_k$ 作为列向量排列成矩阵 $\mathbf{A}$.
2. 同过行变换将 $\mathbf{A}$ 化为阶梯形 $\mathbf{A}'$. 设 $\mathbf{A}'$ 中非零行为第 $1, 2, \ldots, r$ 行, 第 ℓ 行中的第一个非零元素在第 s_ℓ 列.
3. 则 $\operatorname{rank} T = r$, $\mathbf{a}_{s_1}, \ldots, \mathbf{a}_{s_r}$ 是 T 的极大线性无关向量组.

例 2.26

利用矩阵算法求向量组 $(3, 4, -2, 5)$, $(2, -5, 0, -3)$, $(5, 0, -1, 2)$, $(3, 3, -3, 5)$ 的秩以及一组极大线性无关向量组.

最后我们给出解判断一般线性方程组 (2.1) 是否有解的矩阵方法: (2.1) 的系数矩阵为 $\mathbf{A} = (a_{ij})_{n\times m}$, 常数项可写为列向量 $\mathbf{b} = [b_1, \ldots, b_n]$. 方程组的全部信息由如下增广矩阵给出:

$$\mathbf{B} = (\mathbf{A}|\mathbf{b}) = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1m} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2m} & b_2 \\ & & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} & b_n \end{array}\right).$$

命题 2.27

1. 若 $\operatorname{rank} \mathbf{A} = \operatorname{rank} \mathbf{B} = m$, 则 (2.1) 相容且有唯一解.
2. 若 $\operatorname{rank} \mathbf{A} = \operatorname{rank} \mathbf{B} < m$, 则 (2.1) 相容且有无穷多解.
3. 若 $\operatorname{rank} \mathbf{A} < \operatorname{rank} \mathbf{B}$, 则 (2.1) 不相容.

在相容时要写出 (2.1) 的通解, 仍需将 $\mathbf{B}$ 化为阶梯形然后求解.

第二章复习题:

例 2.28

当参数 λ 取不同值时, 判断如下方程组是否可解, 并在可解时求出它的通解:

$$\begin{cases} \lambda x_1 + x_2 + x_3 = 1 \\ x_1 + \lambda x_2 + x_3 = \lambda \\ x_1 + x_2 + \lambda x_3 = \lambda^2 \end{cases}$$

例 2.29

设 $T_1 = \{\mathbf{a}_1, \ldots, \mathbf{a}_r\}$, $T_2 = \{\mathbf{b}_1, \ldots, \mathbf{b}_s\}$, 且 T_1 中向量可以由 T_2 中向量线性表示, 证明 $\operatorname{rank} T_1 \leq \operatorname{rank} T_2$.

例 2.30

设 $\mathbf{a}_1=(1,0,2,3)$, $\mathbf{a}_2=(1,1,3,5)$, $\mathbf{a}_3=(1,-1,\lambda+2,1)$, $\mathbf{a}_4=(1,2,4,\lambda+8)$, $\mathbf{b}=(1,1,\mu+3,5)$.

(1) 求 λ,μ 使得 $\mathbf{b}$ 不能由 $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$, $\mathbf{a}_4$ 线性表出.

(2) 求 λ,μ 使得 $\mathbf{b}$ 可以由 $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$, $\mathbf{a}_4$ 线性表出, 且表达方式唯一.

例 2.31

(1) 设 V,W 是 $\mathbb{F}^m$ 的子空间, 且 $V\subset W$, 则 $\dim V\leq\dim W$.

(2) 设 V,W 是 $\mathbb{F}^m$ 的子空间, 且 $V\subset W$, $\dim V=\dim W$, 则 $V=W$.

例 2.32

设 V, W 是如下给定的 $\mathbb{R}^4$ 的子空间:

$$V = \{(x_1, x_2, x_3, x_4) \mid x_1 - x_2 + x_3 - x_4 = 0\},$$
$$W = \{(x_1, x_2, x_3, x_4) \mid x_1 + x_2 + x_3 + x_4 = 0\},$$

$V \cap W$ 也是 $\mathbb{R}^4$ 的子空间, 求 $V \cap W$ 的一组基.

例 2.33

求 λ 取不同值时矩阵 $\mathbf{A} = \begin{pmatrix} \lambda & 1 & \cdots & 1 \\ 1 & \lambda & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 1 & \cdots & 1 & \lambda \end{pmatrix}$ 的秩.

❸ 矩阵运算

矩阵乘法

可逆矩阵

行列式: 定义, 计算, 应用

3.1

矩阵乘法

上一章我们通过研究齐次线性方程组系数矩阵的秩确定了解空间的维数. 事实上, 如果引入矩阵运算, 线性方程组可以表达为关于矩阵的方程.

回忆一般线性方程组的模样:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2m}x_m = b_2 \\ \qquad\qquad\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nm}x_m = b_n \end{cases} \tag{3.1}$$

回忆向量内积在直角坐标系下的公式, 方程的第 i 行可以表示为 $(a_{i1}, \ldots, a_{im}) \cdot (x_1, \ldots, x_m) = b_i$.

定义 3.1

设 $\mathbf{A}=(a_{ij})_{n\times m}$ 是 $n\times m$ 阶矩阵, $\mathbf{x}=[x_1,\ldots,x_m]$ 是 m 维列向量. 定义 $\mathbf{A}$ 和 $\mathbf{x}$ 的乘积为

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} = \begin{pmatrix} a_{11}x_1+\cdots+a_{1m}x_m \\ a_{21}x_1+\cdots+a_{2m}x_m \\ \vdots \\ a_{n1}x_1+\cdots+a_{nm}x_m \end{pmatrix}.$$

记 $\mathbf{b}=[b_1,\ldots,b_n]$, 则线性方程组 (3.1) 可以表达为

$$\mathbf{A}\cdot\mathbf{x}=\mathbf{b}. \tag{3.2}$$

定义 3.2

设 $\mathbf{A} = (a_{ij})_{n\times m}$, $\mathbf{B} = (b_{jk})_{m\times \ell}$, $\mathbf{A}$ 的行向量为 $\mathbf{a}_1, \ldots, \mathbf{a}_n$, $\mathbf{B}$ 的列向量为 $\mathbf{b}_1, \ldots, \mathbf{b}_\ell$. 定义 $\mathbf{A}$ 和 $\mathbf{B}$ 的矩阵乘积为矩阵

$$\mathbf{C} = (c_{ik})_{n\times \ell} = \begin{pmatrix} \mathbf{a}_1\cdot\mathbf{b}_1 & \mathbf{a}_1\cdot\mathbf{b}_2 & \cdots & \mathbf{a}_1\cdot\mathbf{b}_\ell \\ \mathbf{a}_2\cdot\mathbf{b}_1 & \mathbf{a}_2\cdot\mathbf{b}_2 & \cdots & \mathbf{a}_2\cdot\mathbf{b}_\ell \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{a}_n\cdot\mathbf{b}_1 & \mathbf{a}_n\cdot\mathbf{b}_2 & \cdots & \mathbf{a}_n\cdot\mathbf{b}_\ell \end{pmatrix},$$

并记作 $\mathbf{A}\cdot\mathbf{B}$, 或简记为 $\mathbf{AB}$.

内积就是 $1\times n$ 阶矩阵和 $n\times 1$ 阶矩阵的矩阵乘积.

例 3.3

计算 $\begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} \begin{pmatrix} b_1 & \cdots & b_m \end{pmatrix}$, $\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{pmatrix}$.

记 $\mathsf{M}_{n\times m}(\mathbb{F})$ 为 $\mathbb{F}$ 上 $n \times m$ 阶矩阵的集合. 则矩阵乘积定义了从 $\mathsf{M}_{n\times m}(\mathbb{F}) \times \mathsf{M}_{m\times \ell}(\mathbb{F})$ 到 $\mathsf{M}_{n\times \ell}(\mathbb{F})$ 的映射.

对 $\mathbf{A}, \mathbf{B} \in \mathsf{M}_{n\times m}(\mathbb{F}), \lambda \in \mathbb{F}$, 还可以定义矩阵加法和数乘:

$$\mathbf{A} + \mathbf{B} = (a_{ij} + b_{ij})_{n\times m}, \tag{3.3}$$

$$\lambda\mathbf{A} = (\lambda a_{ij})_{n\times m}. \tag{3.4}$$

事实上, 只看加法和数乘, $\mathsf{M}_{n\times m}(\mathbb{F})$ 是 $\mathbb{F}$ 上 nm 维向量空间.

$n\times n$ 阶矩阵也称为 n **阶方阵**, 所有 n 阶方阵构成的空间记为 $\mathsf{M}_n(\mathbb{F})$. 记 $\mathbf{I}_n$ 是对角线上为 1, 其他位置均为 0 的 n 阶方阵, 称为 n **阶单位阵**. 形如 $\lambda\cdot\mathbf{I}_n$ $(\lambda\in\mathbb{F})$ 的方阵称为**纯量阵**.

引理 3.4

矩阵乘法具有如下性质:

1. 乘法结合律: $(\mathbf{AB})\mathbf{C}=\mathbf{A}(\mathbf{BC})$.
2. 乘法单位元: $\mathbf{I}_n\mathbf{A}=\mathbf{A}=\mathbf{A}\mathbf{I}_m$, $\mathbf{A}\in\mathsf{M}_{n\times m}(\mathbb{F})$.
3. 分配律: $(\mathbf{A}+\mathbf{B})\mathbf{C}=\mathbf{AC}+\mathbf{BC}$, $\mathbf{A}(\mathbf{B}+\mathbf{C})=\mathbf{AB}+\mathbf{AC}$.
4. 数乘结合律: $\lambda(\mathbf{AB})=(\lambda\mathbf{A})\mathbf{B}=\mathbf{A}(\lambda\mathbf{B})$.

特别的, $\mathrm{M}_n(\mathbb{F})$ 在矩阵加法, 乘法下构成一个环, 称为**矩阵环**. 只要 $n>1$, 则 $\mathrm{M}_n(\mathbb{F})$ 是非交换环, 并且两个非零矩阵的乘积可能是零矩阵.

$$\begin{pmatrix}0&1\\0&0\end{pmatrix}\begin{pmatrix}0&0\\1&0\end{pmatrix}=\begin{pmatrix}1&0\\0&0\end{pmatrix},\ \begin{pmatrix}0&0\\1&0\end{pmatrix}\begin{pmatrix}0&1\\0&0\end{pmatrix}=\begin{pmatrix}0&0\\0&1\end{pmatrix},\ \begin{pmatrix}0&1\\0&0\end{pmatrix}^2=\mathbf{0}.$$

记 $\mathbf{E}_{ij}\in\mathrm{M}_n(\mathbb{F})$ 是 (i,j) 位置为 1, 其他位置全为 0 的矩阵, 则 $\mathbf{A}=(a_{ij})_{n\times n}$ 可以表示为 $\mathbf{A}=\sum_{i=1}^n\sum_{j=1}^n a_{ij}\mathbf{E}_{ij}$, 并且有公式

$$\mathbf{E}_{ij}\mathbf{E}_{k\ell}=\delta_{jk}\mathbf{E}_{i\ell}. \tag{3.5}$$

例 3.5

设 $\mathbf{A}\in\mathrm{M}_{n\times m}(\mathbb{F})$, $\mathbf{B}\in\mathrm{M}_{m\times\ell}(\mathbb{F})$, 证明 $(\mathbf{AB})^{\mathrm{T}}=\mathbf{B}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}$.

例 3.6

证明与任意 n 阶方阵乘法可交换的方阵一定是纯量阵.

矩阵 $\mathbf{A} = (a_{ij})_{n\times m}$ 的零空间为

$$\mathrm{Null}\ \mathbf{A} = \{\mathbf{a} \in \mathbb{F}^m \mid \mathbf{A}\cdot\mathbf{a} = \mathbf{0}\},$$

由 (3.2), $\mathrm{Null}\ \mathbf{A}$ 是齐次线性方程组 $\mathbf{A}\cdot\mathbf{x} = \mathbf{0}$ 的解空间.

矩阵 $\mathbf{A} = (a_{ij})_{n\times m}$ 的列空间为

$$\mathrm{Col}\ \mathbf{A} = \{\mathbf{b} \in \mathbb{F}^n \mid \exists\ \mathbf{a} \in \mathbb{F}^m,\ \mathbf{A}\cdot\mathbf{a} = \mathbf{b}\},$$

矩阵 $\mathbf{A}$ 的零空间和列空间都是子空间.

命题 3.7

$\mathrm{Col}\ \mathbf{A}$ 就是 $\mathbf{A}$ 的所有列向量张成的 $\mathbb{F}^n$ 的子空间. 从而 $\dim\ \mathrm{Col}\ \mathbf{A} = \mathrm{rank}\ \mathbf{A}$.

定理 3.8

设 $\mathbf{A} \in \mathsf{M}_{n\times m}(\mathbb{F})$, $\dim\ \mathrm{Null}\ \mathbf{A} + \dim\ \mathrm{Col}\ \mathbf{A} = m$.

例 3.9

(1) 设 $\mathbf{A}, \mathbf{B} \in \mathsf{M}_{n\times m}(\mathbb{F})$, 则 $\mathrm{rank}(\mathbf{A}+\mathbf{B}) \leq \mathrm{rank}\ \mathbf{A} + \mathrm{rank}\ \mathbf{B}$.

(2) 设 $\mathbf{A} \in \mathsf{M}_{n\times m}(\mathbb{F}), \mathbf{B} \in \mathsf{M}_{m\times \ell}(\mathbb{F})$, 则

$$\mathrm{rank}\ \mathbf{AB} \leq \min\{\mathrm{rank}\ \mathbf{A}, \mathrm{rank}\ \mathbf{B}\}.$$

(3) 设 $\mathbf{A} \in \mathsf{M}_{n\times m}(\mathbb{R})$, 证明 $\mathrm{rank}\ \mathbf{A}^{\mathrm{T}}\mathbf{A} = \mathrm{rank}\ \mathbf{A}$.

3.2

可逆矩阵

$\mathrm{M}_n(\mathbb{F})$ 构成一个环, 且有单位元 $\mathbf{I}_n$, 考察 $\mathrm{M}_n(\mathbb{F})$ 中可逆的元素:

定义 3.10

设 $\mathbf{A} \in \mathrm{M}_n(\mathbb{F})$, 如果存在 $\mathbf{B} \in \mathrm{M}_n(\mathbb{F})$ 使得

$$\mathbf{AB} = \mathbf{BA} = \mathbf{I}_n,$$

称 $\mathbf{A}$ 可逆, $\mathbf{B}$ 是 $\mathbf{A}$ 的逆矩阵, 记作 $\mathbf{A}^{-1}$.

命题 3.11

(1) 若 $\mathbf{A}$ 可逆, 则其逆矩阵唯一.

(2) $(\mathbf{A}^{-1})^{-1} = \mathbf{A}, \quad (\lambda\mathbf{A})^{-1} = \lambda^{-1}\mathbf{A}^{-1}$.

(3) $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}, \quad (\mathbf{A}^{\mathrm{T}})^{-1} = (\mathbf{A}^{-1})^{\mathrm{T}}$.

定理 3.12

设 $\mathbf{A} \in \mathsf{M}_n(\mathbb{F})$, 则 $\mathbf{A}$ 可逆 $\Longleftrightarrow \operatorname{rank} \mathbf{A} = n$.

推论 3.13

设 $\mathbf{A} \in \mathsf{M}_n(\mathbb{F})$, 则如下命题等价:

(1) $\mathbf{A}$ 是可逆矩阵.

(2) 存在 $\mathbf{B} \in \mathsf{M}_n(\mathbb{F})$, 使得 $\mathbf{AB} = \mathbf{I}_n$.

(3) 存在 $\mathbf{C} \in \mathsf{M}_n(\mathbb{F})$, 使得 $\mathbf{CA} = \mathbf{I}_n$.

例 3.14

设 $\mathbf{A} = (a_{ij})_{n\times n} \in \mathsf{M}_n(\mathbb{R})$, 若 $a_{ii} > 0$, $a_{ii} > \sum_{j\neq i} |a_{ij}|$ 对任意 $1 \leq i \leq n$ 成立, 称 $\mathbf{A}$ 是对角占优的. 证明对角占优方阵可逆.

对任意 $1 \le i, j \le n$, 定义如下三类初等矩阵:

1. $\mathbf{P}_{ij} = \mathbf{I}_n - \mathbf{E}_{ii} - \mathbf{E}_{jj} + \mathbf{E}_{ij} + \mathbf{E}_{ji}$, 即将单位矩阵 $\mathbf{I}_n$ 的 i 行和 j 行交换得到的矩阵. 易知 $\mathbf{P}_{ij}^{-1} = \mathbf{P}_{ij}$.
2. $\mathbf{D}_i(\lambda) = \mathbf{I}_n + (\lambda - 1)\mathbf{E}_{ii}$, $\lambda \neq 0$, 即将单位矩阵 $\mathbf{I}_n$ 的第 i 行乘以非零数 λ 得到的矩阵. 易知 $\mathbf{D}_i(\lambda)^{-1} = \mathbf{D}_i(\lambda^{-1})$.
3. $\mathbf{T}_{ij}(\lambda) = \mathbf{I}_n + \lambda\mathbf{E}_{ij}$, 即将单位矩阵 $\mathbf{I}_n$ 的第 j 行的 λ 倍加到第 i 行得到的矩阵. 易知 $\mathbf{T}_{ij}(\lambda)^{-1} = \mathbf{T}_{ij}(-\lambda)$.

初等矩阵的逆仍是初等矩阵. 初等矩阵和初等行、列变换的关系如下:

命题 3.15

设 $\mathbf{A}\in \mathsf{M}_{n\times p}(\mathbb{F}), \mathbf{B}\in \mathsf{M}_{q\times n}(\mathbb{F})$, 则

1. $\mathbf{P}_{ij}\cdot \mathbf{A}$ 是将 $\mathbf{A}$ 的 i 行和 j 行交换得到的矩阵.
   $\mathbf{B}\cdot \mathbf{P}_{ij}$ 是将 $\mathbf{B}$ 的 i 列和 j 列交换得到的矩阵.
2. $\mathbf{D}_i(\lambda)\cdot \mathbf{A}$ 是将 $\mathbf{A}$ 的第 i 行乘以非零数 λ 得到的矩阵.
   $\mathbf{B}\cdot \mathbf{D}_i(\lambda)$ 是将 $\mathbf{B}$ 的第 i 列乘以非零数 λ 得到的矩阵.
3. $\mathbf{T}_{ij}(\lambda)\cdot \mathbf{A}$ 是将 $\mathbf{A}$ 的第 j 行的 λ 倍加到第 i 行得到的矩阵.
   $\mathbf{B}\cdot \mathbf{T}_{ij}(\lambda)$ 是将 $\mathbf{B}$ 的第 i 列的 λ 倍加到第 j 列得到的矩阵.

于是对一个矩阵左乘初等矩阵等价于对原矩阵作初等行变换,
右乘初等矩阵等价于对原矩阵作初等列变换.

通过初等行变换可以将矩阵化为阶梯形, 如果继续对阶梯形矩阵进行初等列变换, 则最终可以将矩阵化为如下简单的形式

$$\begin{pmatrix} \mathbf{I}_r & 0 \\ 0 & 0 \end{pmatrix}.$$

因为行、列变换不改变矩阵的秩, 上面的 r 就是矩阵的秩.

定理 3.16

设 $\mathbf{A} \in \mathsf{M}_{n\times m}(\mathbb{F})$, 且 $\operatorname{rank} \mathbf{A} = r$, 则存在 n 阶初等矩阵 $\mathbf{P}_1, \ldots, \mathbf{P}_k$, 以及 m 阶初等矩阵 $\mathbf{Q}_1, \ldots, \mathbf{Q}_\ell$, 使得

$$\mathbf{P}_1 \cdots \mathbf{P}_k \cdot \mathbf{A} \cdot \mathbf{Q}_1 \cdots \mathbf{Q}_\ell = \begin{pmatrix} \mathbf{I}_r & 0 \\ 0 & 0 \end{pmatrix}.$$

推论 3.17

可逆矩阵是一些初等矩阵的乘积. 特别的, 若 $\mathbf{A}$ 可逆, 则存在初等矩阵 $\mathbf{P}_1, \ldots, \mathbf{P}_k$ 使得

$$\mathbf{P}_1 \cdots \mathbf{P}_k \cdot \mathbf{A} = \mathbf{I}_n \text{ (也可以直接由初等行变换看出).}$$

另一方面, $\mathbf{A}^{-1} = \mathbf{P}_1 \cdots \mathbf{P}_k$, 从而 $\mathbf{A}^{-1}$ 可以看作将单位矩阵 $\mathbf{I}_n$ 依次作行变换 $\mathbf{P}_k, \mathbf{P}_{k-1}, \ldots, \mathbf{P}_1$ 得到. 这样就得到了逆矩阵 $\mathbf{A}^{-1}$ 的有效算法: 同时对 $\mathbf{A}$ 和 $\mathbf{I}_n$ 做相同的初等行变换, 当 $\mathbf{A}$ 变为 $\mathbf{I}_n$ 时, $\mathbf{I}_n$ 就变为了 $\mathbf{A}^{-1}$. 过程如下图所示:

$$(\mathbf{A} \mid \mathbf{I}_n) \xrightarrow{\mathbf{P}_k} (\mathbf{P}_k\mathbf{A} \mid \mathbf{P}_k) \xrightarrow{\mathbf{P}_{k-1}} \cdots \xrightarrow{\mathbf{P}_1} (\mathbf{I}_n \mid \mathbf{P}_1 \cdots \mathbf{P}_k)$$

例 3.18

(1) 求矩阵 $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 2 \\ 1 & 1 & 1 & -1 \\ 1 & 0 & -2 & -6 \end{pmatrix}$ 的逆.

(2) 设 $\mathbf{A}$ 的秩为 r, 证明 $\mathbf{A}$ 可以分解为 r 个秩为 1 的矩阵之和.

在上面表示 $\begin{pmatrix} \mathbf{I}_r & 0 \\ 0 & 0 \end{pmatrix}$ 时, 以及更早的我们将 $\mathbf{A}$ 表示成它的行向量 $\mathbf{A} = \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_n \end{pmatrix}$ 以及列向量 $\mathbf{A} = \begin{pmatrix} \mathbf{b}_1 & \cdots & \mathbf{b}_m \end{pmatrix}$ 的形式时我们都用到了分块矩阵的技巧.

一般的, 对任意分解 $n = n_1 + n_2$, $m = m_1 + m_2$, $\ell = \ell_1 + \ell_2$, 有分块矩阵乘积:

$$\begin{array}{cc} & \begin{array}{cc} m_1 & m_2 \end{array} \\ \begin{array}{c} n_1 \\ n_2 \end{array} & \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} \end{array} \begin{array}{cc} \begin{array}{cc} \ell_1 & \ell_2 \end{array} & \\ \begin{pmatrix} \mathbf{A}' & \mathbf{B}' \\ \mathbf{C}' & \mathbf{D}' \end{pmatrix} & \begin{array}{c} m_1 \\ m_2 \end{array} \end{array} = \begin{array}{cc} & \begin{array}{cc} \ell_1 & \ell_2 \end{array} \\ \begin{array}{c} n_1 \\ n_2 \end{array} & \begin{pmatrix} \mathbf{AA}' + \mathbf{BC}' & \mathbf{AB}' + \mathbf{BD}' \\ \mathbf{CA}' + \mathbf{DC}' & \mathbf{CB}' + \mathbf{DD}' \end{pmatrix} \end{array}.$$

例 3.19

(1) 求 $\begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}^n$.

(2) 设 $\mathbf{A}, \mathbf{B} \in \mathrm{M}_n(\mathbb{F})$, 且 $\mathbf{AB} = 0$, 求 $\begin{pmatrix} \mathbf{I}_n & \mathbf{A} \\ \mathbf{B} & \mathbf{I}_n \end{pmatrix}^{-1}$.

3.3

行列式: 定义, 计算, 应用

回忆二阶方阵行列式的公式及几何含义:

设 $\mathbf{A}=\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, $\mathbf{a}_1$, $\mathbf{a}_2$ 为 $\mathbf{A}$ 的行向量. 则

$$|\mathbf{A}| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{21}a_{12}, \tag{3.6}$$

等于以 $\mathbf{a}_1, \mathbf{a}_2$ 为邻边的平行四边形的 “有向” 面积.

回忆三阶方阵行列式的公式及几何含义:

设 $\mathbf{A}=\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$, $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$ 为 $\mathbf{A}$ 的行向量. 则

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{31}\begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} - a_{32}\begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix} + a_{33}\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \tag{3.7}$$

等于以 $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ 为邻边的平行六面体的 "有向" 体积.

设 $\mathbf{A} \in \mathsf{M}_n(\mathbb{F})$, 期望赋予方阵 $\mathbf{A}$ 一个 $\mathbb{F}$ 中的值, 称为 $\mathbf{A}$ 的 n 阶行列式, 记作 $\det_n \mathbf{A}$, $\det \mathbf{A}$ 或 $|\mathbf{A}|$. 行列式是从 $\mathsf{M}_n(\mathbb{F})$ 到 $\mathbb{F}$ 的一个函数 $\det\colon \mathsf{M}_n(\mathbb{F}) \to \mathbb{F}$.

$\det_n$ 也可以看作关于 n 个 n 维向量 $\mathbf{a}_1, \ldots, \mathbf{a}_n \in \mathbb{F}^n$ 的函数, 记作 $\det_n(\mathbf{a}_1, \ldots, \mathbf{a}_n)$.

二阶行列式满足如下性质:

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = -\begin{vmatrix} c & d \\ a & b \end{vmatrix}, \quad \begin{vmatrix} \lambda a + \mu a' & \lambda b + \mu b' \\ c & d \end{vmatrix} = \lambda \begin{vmatrix} a & b \\ c & d \end{vmatrix} + \mu \begin{vmatrix} a' & b' \\ c & d \end{vmatrix},$$

$$\begin{vmatrix} a & b \\ \lambda c + \mu c' & \lambda d + \mu d' \end{vmatrix} = \lambda \begin{vmatrix} a & b \\ c & d \end{vmatrix} + \mu \begin{vmatrix} a & b \\ c' & d' \end{vmatrix}, \quad \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1.$$

将行列式看成关于行向量的函数, 等价的有

$$\begin{aligned}
&(1) \quad \det{}_2(\lambda \mathbf{a}_1 + \mu \mathbf{a}_2, \mathbf{b}) = \lambda \det{}_2(\mathbf{a}_1, \mathbf{b}) + \mu \det{}_2(\mathbf{a}_2, \mathbf{b}), \\
&(2) \quad \det{}_2(\mathbf{a}, \lambda \mathbf{b}_1 + \mu \mathbf{b}_2) = \lambda \det{}_2(\mathbf{a}, \mathbf{b}_1) + \mu \det{}_2(\mathbf{a}, \mathbf{b}_2), \\
&(3) \quad \det{}_2(\mathbf{a}, \mathbf{b}) = -\det{}_2(\mathbf{b}, \mathbf{a}), \\
&(4) \quad \det{}_2(\boldsymbol{\epsilon}_1, \boldsymbol{\epsilon}_2) = 1, \quad \boldsymbol{\epsilon}_1 = (1, 0), \boldsymbol{\epsilon}_2 = (0, 1).
\end{aligned}$$

(1), (2) 称为**线性**, (3) 称为**交错性**, (4) 称为**规范性**.

引理 3.20

设 $D\colon \mathbb{F}^2 \times \mathbb{F}^2 \to \mathbb{F}$ 满足线性, 则 D 是交错的当且仅当对 $\forall\, \mathbf{a} \in \mathbb{F}^2$, 有 $D(\mathbf{a}, \mathbf{a}) = 0$.

定理 3.21

设 $D\colon \mathbb{F}^2 \times \mathbb{F}^2 \to \mathbb{F}$ 满足线性, 交错性以及规范性, 则 $D = \det_2$.

对于 $\mathbb{F}^n$ 上的 n 元函数 $D_n\colon \overbrace{\mathbb{F}^n \times \cdots \times \mathbb{F}^n}^{n} \to \mathbb{F}$, 仍然可以定义线性, 交错性以及规范性:

$$
\begin{aligned}
\text{线性}\quad & D_n(\ldots, \lambda \mathbf{a}_i + \mu \mathbf{a}_i', \ldots) \\
& = \lambda D_n(\ldots, \mathbf{a}_i, \ldots) + \mu D_n(\ldots, \mathbf{a}_i', \ldots), \quad \forall\, 1 \le i \le n \\
\text{交错性}\quad & D_n(\ldots, \mathbf{a}_i, \ldots, \mathbf{a}_j, \ldots) \\
& = -D_n(\ldots, \mathbf{a}_j, \ldots, \mathbf{a}_i, \ldots), \quad \forall\, 1 \le i < j \le n \\
\text{规范性}\quad & D_n(\boldsymbol{\epsilon}_1, \ldots, \boldsymbol{\epsilon}_n) = 1.
\end{aligned}
$$

定义 3.22

满足线性, 交错性以及规范性的 $\mathbb{F}^n$ 上的 n 元函数称为 n 阶行列式函数, 记作 $\det_n$. 以 $\mathbf{a}_1, \ldots, \mathbf{a}_n \in \mathbb{F}^n$ 为行向量的方阵 $\mathbf{A} \in \mathsf{M}_n(\mathbb{F})$ 的行列式记为 $\det \mathbf{A}$, 或 $|\mathbf{A}|$.

例 3.23

验证 3 阶行列式函数定义的合理性.

我们将证明行列式函数存在且唯一, 首先引入置换的概念.

定义 3.24

一个 n 元置换是从 $\{1,2,\ldots,n\}$ 到它自身的一个一一映射

$$\sigma\colon \{1,2,\ldots,n\} \longleftrightarrow \{1,2,\ldots,n\}.$$

记为 $\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix}$, 或 $\sigma = (k_1 k_2 \cdots k_n)$.

一个 n 元置换也可以看作 $\{1,2,\ldots,n\}$ 的一个排列. 所有 n 元置换构成的集合记为 $\mathfrak{S}_n$.

在 $\sigma = (k_1 k_2 \cdots k_n)$ 中, 交换 i, j 位置的两个数 k_i, k_j 的操作称为对换, 记为 (ij). 因为 $k_1, \ldots, k_n$ 是 $1, \ldots, n$ 的一个排列, 可以通过如下一系列对换将 σ 变为单位置换 $(12 \cdots n)$:

定理 3.25 (冒泡排序)

1. 比较 k_1, k_2, 若 $k_1 > k_2$, 则做对换 (12), 否则不做操作. 继续在新得到的置换中比较 k_2, k_3, 若 $k_2 > k_3$, 则做对换 (23), 否则不做操作. 以此类推, 比较并操作至 k_{n-1}, k_n. 则在得到的置换中 $k_n = n$.
2. 对 $k_1, \ldots, k_{n-1}$ 进行第一步的操作, 从 k_1, k_2 比较并操作至 k_{n-2}, k_{n-1}, 则在得到的置换中 $k_{n-1} = n - 1$.
3. 重复以上操作, $n - 1$ 步后, 得到置换 $(12 \cdots n)$.

定义 3.26

设 $\sigma = (k_1 k_2 \cdots k_n)$, 记

$$\mathrm{inv}(\sigma) = \#\{(k_i, k_j) \mid i < j,\ k_i > k_j\},$$

称为 σ 的**逆序数**, 其中 $\#$ 表示集合元素的个数. $(-1)^{\mathrm{inv}(\sigma)}$ 称为 σ 的**符号**, 记为 $\mathrm{sign}(\sigma)$. 若 $\mathrm{sign}(\sigma) = 1$, σ 称为**偶置换**; 若 $\mathrm{sign}(\sigma) = -1$, σ 称为**奇置换**.

例 3.27

设 $\sigma = (4231)$, 则

$$\mathrm{inv}(\sigma) = \#\{(4,2), (4,3), (4,1), (2,1), (3,1)\} = 5.$$

$$\mathrm{sign}(\sigma) = (-1)^5 = 1, \ \sigma \text{ 是奇置换}.$$

引理 3.28

利用冒泡排序将 $\sigma = (k_1 k_2 \cdots k_n)$ 变为 $(12\cdots n)$ 所需的对换操作次数为 $\mathrm{inv}(\sigma)$.

命题 3.29

设 D 是满足交错性的 n 元函数, $\sigma \in \mathfrak{S}_n$, 则

$$D(x_{\sigma(1)}, x_{\sigma(2)}, \ldots, x_{\sigma(n)}) = \mathrm{sign}(\sigma) D(x_1, x_2, \ldots, x_n).$$

接下来给出 n 阶行列式的公式. 设 $\mathbf{A} = (a_{ij}) \in \mathsf{M}_n(\mathbb{F})$, $\mathbf{a}_1, \ldots, \mathbf{a}_n$ 是 $\mathbf{A}$ 的行向量. 则 $\mathbf{a}_i = \sum_{j=1}^n a_{ij} \boldsymbol{\epsilon}_j$, 从而

$$
\begin{aligned}
\det\ \mathbf{A} &= \det(\mathbf{a}_1, \ldots, \mathbf{a}_n) \\
&= \det\Big(\sum_{k_1=1}^{n} a_{1k_1}\boldsymbol{\epsilon}_{k_1}, \ldots, \sum_{k_n=1}^{n} a_{nk_n}\boldsymbol{\epsilon}_{k_n}\Big) \\
\text{(线性)} &= \sum_{1\le k_1,\ldots,k_n\le n} a_{1k_1}\cdots a_{nk_n}\det(\boldsymbol{\epsilon}_{k_1}, \ldots, \boldsymbol{\epsilon}_{k_n}) \\
\text{(交错性)} &= \sum_{(k_1\cdots k_n)\in\mathfrak{S}_n} a_{1k_1}\cdots a_{nk_n}\det(\boldsymbol{\epsilon}_{k_1}, \ldots, \boldsymbol{\epsilon}_{k_n}) \\
\text{(命题 3.29)} &= \sum_{(k_1\cdots k_n)\in\mathfrak{S}_n} \operatorname{sign}(k_1\cdots k_n)a_{1k_1}\cdots a_{nk_n}\det(\boldsymbol{\epsilon}_1, \ldots, \boldsymbol{\epsilon}_n) \\
\text{(规范性)} &= \sum_{\sigma\in\mathfrak{S}_n} \operatorname{sign}(\sigma)a_{1\sigma(1)}a_{2\sigma(2)}\cdots a_{n\sigma(n)}.
\end{aligned}
$$

定理 3.30

行列式函数 $\det_n$ 存在且唯一. 设 $\mathbf{A} = (a_{ij})_{n\times n}$, 公式为

$$\det\ \mathbf{A} = \sum_{\sigma\in\mathfrak{S}_n} \operatorname{sign}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}.$$

例 3.31

按照定义计算如下行列式:

$$\begin{vmatrix} 0 & 0 & 0 & a_1 \\ 0 & 0 & a_2 & 0 \\ 0 & a_3 & 0 & 0 \\ a_4 & 0 & 0 & 0 \end{vmatrix}, \qquad \begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ 0 & a_{22} & a_{23} & a_{24} \\ 0 & 0 & a_{33} & a_{34} \\ 0 & 0 & 0 & a_{44} \end{vmatrix}.$$

定理 3.30 给出了行列式的公式, 但并不适用于计算. 我们已经看到上三角矩阵的行列式是可以直接计算的, 而每个方阵可以通过一系列初等行变换化为上三角矩阵, 需要研究初等行变换对行列式的影响.

命题 3.32

设 $\mathbf{A} \in \mathsf{M}_n(\mathbb{F})$, $\mathbf{P} \in \mathsf{M}_n(\mathbb{F})$ 是初等矩阵.

1. 若 $\mathbf{P} = \mathbf{P}_{ij}$, 则 $\det \mathbf{PA} = -\det \mathbf{A}$.
2. 若 $\mathbf{P} = \mathbf{D}_i(\lambda)$ $(\lambda \neq 0)$, 则 $\det \mathbf{PA} = \lambda \det \mathbf{A}$.
3. 若 $\mathbf{P} = \mathbf{T}_{ij}(\lambda)$, 则 $\det \mathbf{PA} = \det \mathbf{A}$.

例 3.33

计算如下行列式: $\begin{vmatrix} 1 & 1 & 2 \\ 2 & 3 & 4 \\ 3 & 4 & 7 \end{vmatrix}, \qquad \begin{vmatrix} 0 & 0 & 1 \\ 0 & 2 & 3 \\ 3 & 4 & 5 \end{vmatrix}.$

推论 3.34

设 $\mathbf{A} \in \mathsf{M}_n(\mathbb{F})$, $\mathbf{P} \in \mathsf{M}_n(\mathbb{F})$ 是初等矩阵, 则

$$\det\ \mathbf{PA} = \det\ \mathbf{P} \det\ \mathbf{A}.$$

特别的, 设 $\mathbf{P}_1, \ldots, \mathbf{P}_k$ 是初等矩阵, 则

$$\det\ \mathbf{P}_1 \cdots \mathbf{P}_k = \det\ \mathbf{P}_1 \cdots \det\ \mathbf{P}_k.$$

定理 3.35

设 $\mathbf{A} \in \mathsf{M}_n(\mathbb{F})$, 则 $\mathbf{A}$ 可逆 $\Longleftrightarrow \det\ \mathbf{A} \neq 0$.

定理 3.36

设 $\mathbf{A}, \mathbf{B} \in \mathsf{M}_n$, 则 $\det\ \mathbf{AB} = \det\ \mathbf{A} \det\ \mathbf{B}$.

命题 3.37

设 $\mathbf{A}_1 \in \mathsf{M}_{n_1}(\mathbb{F})$, $\mathbf{A}_2 \in \mathsf{M}_{n_2}(\mathbb{F})$, $n_1 + n_2 = n$, 则

$$\begin{vmatrix} \mathbf{A}_1 & \mathbf{A}_3 \\ \mathbf{0} & \mathbf{A}_2 \end{vmatrix} = |\mathbf{A}_1||\mathbf{A}_2|,$$

$$\begin{vmatrix} \mathbf{A}_3 & \mathbf{A}_1 \\ \mathbf{A}_2 & \mathbf{0} \end{vmatrix} = (-1)^{n_1 n_2}|\mathbf{A}_1||\mathbf{A}_2|.$$

引理 3.38

设 $\sigma \in \mathfrak{S}_n$, σ^{-1} 是 σ 的逆, 则 $\operatorname{sign}(\sigma) = \operatorname{sign}(\sigma^{-1})$.

命题 3.39

设 $\mathbf{A} \in \mathsf{M}_n(\mathbb{F})$, 则 $\det \mathbf{A} = \det \mathbf{A}^{\mathrm{T}}$.

命题 3.39 表明 det 既可以看作关于行向量的函数, 也可以看作关于列向量的函数. 关于列向量, 也满足线性, 交错性和规范性, 可以通过列变换来简化计算.

例 3.40

计算 $\begin{vmatrix} 1 & a_1 & \cdots & a_1^{n-1} \\ 1 & a_2 & \cdots & a_2^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & a_n & \cdots & a_n^{n-1} \end{vmatrix}$ (范德蒙德行列式).

命题 3.41

设 $\mathbf{A} \in \mathrm{M}_n(\mathbb{F})$, 且第 j 列 (或第 i 行) 除了 a_{ij} 均为 0, 并且 $\mathbf{M}_{ij}$ 是去掉 $\mathbf{A}$ 的第 i 行和第 j 列得到的 $n-1$ 阶方阵. 则

$$\det \mathbf{A} = (-1)^{i+j} a_{ij} \det \mathbf{M}_{ij}.$$

定义 3.42

设 $\mathbf{A} \in \mathrm{M}_n(\mathbb{F})$, $\mathbf{M}_{ij}$ 是去掉 $\mathbf{A}$ 的第 i 行和第 j 列得到的 $n-1$ 阶方阵, 记

$$A_{ij} = (-1)^{i+j} \det \mathbf{M}_{ij},$$

称为 $\mathbf{A}$ 关于 a_{ij} 的代数余子式.

定理 3.43

设 $\mathbf{A} \in \mathsf{M}_n(\mathbb{F})$，对任意 $1 \leq i \leq n$，

$$\det\ \mathbf{A} = a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in},$$

$$\det\ \mathbf{A} = a_{1i}A_{1i} + a_{2i}A_{2i} + \cdots + a_{ni}A_{ni}.$$

分别称为行列式关于第 i 行和第 i 列的展开.

推论 3.44

设 $\mathbf{A} \in \mathsf{M}_n(\mathbb{F})$，$1 \leq i, j \leq n$，则

$$a_{i1}A_{j1} + a_{i2}A_{j2} + \cdots + a_{in}A_{jn} = \delta_{ij}\det\ \mathbf{A},$$

$$a_{1i}A_{1j} + a_{2i}A_{2j} + \cdots + a_{ni}A_{nj} = \delta_{ij}\det\ \mathbf{A}.$$

设 $\mathbf{A} \in \mathsf{M}_n(\mathbb{F})$, 定义 $\mathbf{A}$ 的伴随矩阵为

$$\mathbf{A}^* = (A_{ij})^{\mathrm{T}} = \begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix}.$$

命题 3.45

设 $\mathbf{A} \in \mathsf{M}_n(\mathbb{F})$, 则 $\mathbf{A}\mathbf{A}^* = (\det \mathbf{A})\mathbf{I}_n$. 特别的, 如果 $\mathbf{A}$ 可逆, 则

$$\mathbf{A}^{-1} = \frac{\mathbf{A}^*}{\det \mathbf{A}}.$$

例 3.46

设 $\mathbf{A}=\begin{pmatrix} 2 & 2 & 3 \\ 1 & -1 & 0 \\ -1 & 2 & 1 \end{pmatrix}$, 求 $\mathbf{A}^*$ 以及 $\mathbf{A}^{-1}$.

定理 3.47 (克拉默法则)

如果线性方程组 $\mathbf{A}\mathbf{x}=\mathbf{b}$ 的系数矩阵 $\mathbf{A}=(a_{ij})_{n\times n}$ 可逆, 则它有唯一解. 设 $\mathbf{A}$ 的列向量为 $\mathbf{a}_1,\ldots,\mathbf{a}_n$, 则解的公式为:

$$x_k=\frac{\det(\mathbf{a}_1,\ldots,\mathbf{b},\ldots,\mathbf{a}_n)}{\det(\mathbf{a}_1,\ldots,\mathbf{a}_k,\ldots,\mathbf{a}_n)},\quad k=1,2,\ldots,n$$

例 3.48

利用克拉默法则解线性方程组
$$\begin{cases} x_1+2x_2+3x_3=1, \\ 2x_1+2x_2+5x_3=2, \\ 3x_1+5x_2+x_3=3. \end{cases}$$

下面介绍行列式与矩阵的秩的关系:

定义 3.49

设 $\mathbf{A} \in \mathrm{M}_{n\times m}(\mathbb{F})$, $\mathbf{A}$ 的第 $i_1,\ldots,i_k$ 行与 $j_1,\ldots,j_k$ 列交叉处形成的 k 阶方阵的行列式称为 $\mathbf{A}$ 的一个 k 阶子式, 记为

$$A\begin{pmatrix} i_1 & \cdots & i_k \\ j_1 & \cdots & j_k \end{pmatrix} = \det \begin{pmatrix} a_{i_1j_1} & \cdots & a_{i_1j_k} \\ \vdots & \ddots & \vdots \\ a_{i_kj_1} & \cdots & a_{i_kj_k} \end{pmatrix}.$$

命题 3.50

设 $\mathbf{A} \in \mathsf{M}_{n\times m}(\mathbb{F})$, 则 $\operatorname{rank} \mathbf{A} = r \iff \mathbf{A}$ 的任意 $r+1$ 阶子式为 0, 且存在 $\mathbf{A}$ 的一个 r 阶子式不为 0 .

例 3.51

设 $\mathbf{A} \in \mathsf{M}_{n}(\mathbb{F})$, 则 $\operatorname{rank} \mathbf{A}^* = \begin{cases} n, & \operatorname{rank} \mathbf{A} = n. \\ 1, & \operatorname{rank} \mathbf{A} = n-1. \\ 0, & \operatorname{rank} \mathbf{A} < n-1. \end{cases}$

行列式还有一些几何应用:

命题 3.52

过两点 $(x_1, y_1), (x_2, y_2) \in \mathbb{R}^2$ 的直线方程为 $\begin{vmatrix} x & y & 1 \\ x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \end{vmatrix} = 0.$

命题 3.53

过三点的圆的方程为 $\begin{vmatrix} x^2+y^2 & x & y & 1 \\ x_1^2+y_1^2 & x_1 & y_1 & 1 \\ x_2^2+y_2^2 & x_2 & y_2 & 1 \\ x_3^2+y_3^2 & x_3 & y_3 & 1 \end{vmatrix} = 0$, 其中 (x_1, y_1), (x_2, y_2), $(x_3, y_3) \in \mathbb{R}^2$.

第三章复习题:

例 3.54

计算 (1) $\begin{pmatrix} 1 & 1 & & \\ & 1 & \ddots & \\ & & \ddots & 1 \\ & & & 1 \end{pmatrix}_{n\times n}^{k}$. (2) $\begin{vmatrix} x & & & a_n \\ -1 & \ddots & & \vdots \\ & \ddots & x & a_2 \\ & & -1 & x+a_1 \end{vmatrix}$.

例 3.55

设 $\mathbf{A}, \mathbf{B} \in \mathsf{M}_n(\mathbb{F})$, $\lambda \in \mathbb{F}$, 证明

$$(\lambda\mathbf{A})^* = \lambda^{n-1}\mathbf{A}^*, \quad |\mathbf{A}^*| = |\mathbf{A}|^{n-1}, \quad (\mathbf{A}^*)^{\mathrm{T}} = (\mathbf{A}^{\mathrm{T}})^*,$$
$$(\mathbf{AB})^* = \mathbf{B}^*\mathbf{A}^*, \quad (\mathbf{A}^*)^* = |\mathbf{A}|^{n-2}\mathbf{A}, \ n \geq 2.$$

例 3.56

设 $\mathbf{A} \in \mathsf{M}_n(\mathbb{R})$, 是对角占优矩阵, 证明 $|\mathbf{A}| > 0$.

例 3.57

证明 $\operatorname{rank} \begin{pmatrix} \mathbf{A} & \mathbf{C} \\ \mathbf{0} & \mathbf{B} \end{pmatrix} \geq \operatorname{rank} \mathbf{A} + \operatorname{rank} \mathbf{B}$, 其中 $\mathbf{A} \in \mathsf{M}_{n_1\times m_1}(\mathbb{F})$, $\mathbf{B} \in \mathsf{M}_{n_2\times m_2}(\mathbb{F})$, $\mathbf{C} \in \mathsf{M}_{n_1\times m_2}(\mathbb{F})$.

例 3.58

设 $\mathbf{A} \in \mathsf{M}_n(\mathbb{F})$, 且 $\mathbf{A}^{\mathrm{T}}\mathbf{A} = \mathbf{I}_n$, $|\mathbf{A}| < 0$. 证明 $|\mathbf{A} + \mathbf{I}_n| = 0$.

例 3.59

设 $\mathbf{A} \in \mathsf{M}_n(\mathbb{F})$, $k \in \mathbb{N}$, 如果 $\operatorname{rank} \mathbf{A}^k = \operatorname{rank} \mathbf{A}^{k+1}$, 则 $\operatorname{rank} \mathbf{A}^k = \operatorname{rank} \mathbf{A}^{k+1} = \operatorname{rank} \mathbf{A}^{k+2} = \cdots$.

4 线性变换

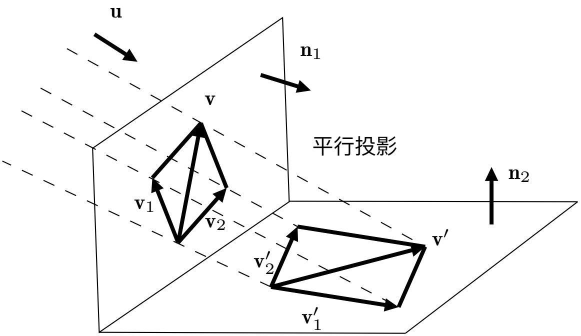
线性变换及其矩阵

特征值, 特征向量与对角化

4.1

线性变换及其矩阵

我们已经研究了向量空间 $\mathbb{F}^n$ 及其子空间的结构, 接下来更为重要的 (无论从实际应用还是数学理论上) 是研究向量空间之间的映射. 我们熟知的平面或空间中的几何变换, 例如伸缩, 旋转和反射, 都是线性变换. 从一个平面到另一个平面的平行投影也是线性变换.



定义 4.1

如果映射 $\mathcal{A}\colon \mathbb{F}^n \longrightarrow \mathbb{F}^n$ 满足

$$\forall\ \mathbf{u},\mathbf{v}\in V,\ \ \lambda,\mu\in\mathbb{F},\quad \mathcal{A}(\lambda\mathbf{u}+\mu\mathbf{v})=\lambda\mathcal{A}(\mathbf{u})+\mu\mathcal{A}(\mathbf{v}),\quad (4.1)$$

则称 $\mathcal{A}$ 为 $\mathbb{F}^n$ 上的线性变换.

例 4.2

设 $\mathbf{A}\in \mathrm{M}_n(\mathbb{F})$, 定义映射 $\varphi_{\mathbf{A}}\colon \mathbb{F}^n \longrightarrow \mathbb{F}^n$ 为

$$\varphi_{\mathbf{A}}(\mathbf{x})=\mathbf{A}\mathbf{x},\quad \forall\ \mathbf{x}\in\mathbb{F}^n.$$

由矩阵乘法的性质可知 $\varphi_{\mathbf{A}}$ 是线性变换, 称为矩阵 $\mathbf{A}$ 的线性变换.

例 4.3

在平面 $\mathbb{R}^2$ 上取直角坐标系 Oxy.

(1) 设 $\mathcal{S}_{\lambda,\mu}$ 是沿 x 轴系数为 λ, 沿 y 轴系数为 μ 的伸缩变换, 则

$$\mathcal{S}_{\lambda,\mu}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \lambda a \\ \mu b \end{pmatrix} = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix}. \tag{4.2}$$

(2) 设 $\mathcal{R}_\theta$ 是绕原点逆时针旋转 θ 角度的旋转变换, 则

$$\mathcal{R}_\theta\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \cos\theta a - \sin\theta b \\ \sin\theta a + \cos\theta b \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix}. \tag{4.3}$$

(3) 设 $\mathbf{u}=(\cos\theta,\sin\theta)$ 是以原点为起点的单位向量, L 是以 $\mathbf{u}$ 为方向向量的直线, r_θ 是以直线 L 为轴的反射变换. 设 $\mathbf{v}=[a,b]$ 是以原点为起点的向量, 则 $r_\theta(\mathbf{v})=2(\mathbf{v}\cdot\mathbf{u})\mathbf{u}-\mathbf{v}$, 坐标表示为

$$\begin{aligned} r_\theta\begin{pmatrix} a \\ b \end{pmatrix} &= 2(\cos\theta a+\sin\theta b)\begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}-\begin{pmatrix} a \\ b \end{pmatrix} \\ &= \begin{pmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix}. \end{aligned} \tag{4.4}$$

综上, 平面上的伸缩, 旋转, 反射都是二阶矩阵的线性变换.

命题 4.4

设 $\mathcal{A}\colon \mathbb{F}^n \longrightarrow \mathbb{F}^n$ 是线性变换, 则 $\mathcal{A}(\mathbf{0}) = \mathbf{0}$.

例 4.5

判断下列映射是否是线性变换:

$$\mathcal{A}\colon \mathbb{F}^3 \longrightarrow \mathbb{F}^3, \quad \mathcal{A}(a,b,c) = (a-b, c, a+1).$$

例 4.6

设 $\mathcal{T}_{\mathbf{a}}$ 是平面 $\mathbb{R}^2$ 上关于向量 $\mathbf{a}$ 的平移, 即

$$\mathcal{T}_{\mathbf{a}}(\mathbf{x}) = \mathbf{x} + \mathbf{a}, \quad \forall\, \mathbf{x} \in \mathbb{R}^2.$$

若 $\mathbf{a} \neq \mathbf{0}$, 则 $\mathcal{T}_{\mathbf{a}}$ 不是线性变换.

设 $\mathcal{A}\colon \mathbb{F}^n \longrightarrow \mathbb{F}^n$ 是线性变换. 对 $\forall$ $\mathbf{x} = [x_1, \ldots, x_n] \in \mathbb{F}^n$, 由定义

$$\mathcal{A}(\mathbf{x}) = x_1\mathcal{A}(\boldsymbol{\epsilon}_1) + \cdots + x_n\mathcal{A}(\boldsymbol{\epsilon}_n),$$

从而 $\mathcal{A}$ 由它在标准基 $\boldsymbol{\epsilon}_1, \ldots, \boldsymbol{\epsilon}_n$ 上的取值所决定.

设 $\mathcal{A}(\boldsymbol{\epsilon}_1) = \mathbf{a}_1, \ldots, \mathcal{A}(\boldsymbol{\epsilon}_n) = \mathbf{a}_n$, 且

$$\mathbf{A} = \begin{pmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{pmatrix}$$

是以 $\mathbf{a}_1, \ldots, \mathbf{a}_n$ 为列向量的矩阵, 则

$$\mathcal{A}(\mathbf{x}) = x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n = \mathbf{A}\mathbf{x}.$$

定理 4.7

设 $\mathcal{A}\colon \mathbb{F}^n \longrightarrow \mathbb{F}^n$ 是线性变换, 且

$$\mathbf{A} = \begin{pmatrix} \mathcal{A}(\boldsymbol{\epsilon}_1) & \cdots & \mathcal{A}(\boldsymbol{\epsilon}_n) \end{pmatrix}.$$

则 $\mathcal{A} = \varphi_{\mathbf{A}}$, 即 $\mathcal{A}$ 是矩阵 $\mathbf{A}$ 的线性变换. $\mathbf{A}$ 称为 *$\mathcal{A}$ 在标准基下的矩阵*.

综上, $\mathbb{F}^n$ **上的线性变换一一对应于 $\mathrm{M}_n(\mathbb{F})$ 中的矩阵.**

例 4.8

求线性变换 $\mathcal{A}\colon \mathbb{F}^3 \to \mathbb{F}^3$ 在标准基下的矩阵, 其中

$$\mathcal{A}[x, y, z] = [2x + y + z, x + 3y + 2z, -x - y + 3z].$$

设 $\boldsymbol{\alpha} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ 是 $\mathbb{F}^n$ 的任何一组基. 则线性变换 $\mathcal{A}\colon \mathbb{F}^n \longrightarrow \mathbb{F}^n$ 也由 $\mathcal{A}$ 在 $\mathbf{v}_1, \ldots, \mathbf{v}_n$ 上的取值决定. 设

$$\begin{aligned}
\mathcal{A}(\mathbf{v}_1) &= a_{11}\mathbf{v}_1 + a_{21}\mathbf{v}_2 + \cdots + a_{n1}\mathbf{v}_n, \\
\mathcal{A}(\mathbf{v}_2) &= a_{12}\mathbf{v}_1 + a_{22}\mathbf{v}_2 + \cdots + a_{n2}\mathbf{v}_n, \\
&\vdots \\
\mathcal{A}(\mathbf{v}_n) &= a_{1n}\mathbf{v}_1 + a_{2n}\mathbf{v}_2 + \cdots + a_{nn}\mathbf{v}_n.
\end{aligned} \tag{4.5}$$

则 $\mathcal{A}$ 由矩阵 $\mathbf{A} = (a_{ij})_{n\times n}$ 唯一决定, 称为 $\mathcal{A}$ 在基 $\boldsymbol{\alpha}$ 下的矩阵. (4.5) 可以写为如下更为简单的形式:

$$\begin{pmatrix} \mathcal{A}(\mathbf{v}_1) & \cdots & \mathcal{A}(\mathbf{v}_n) \end{pmatrix} = \begin{pmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_n \end{pmatrix} \cdot \mathbf{A}. \tag{4.6}$$

定理 4.9

设 $\mathcal{A}\colon \mathbb{F}^n \longrightarrow \mathbb{F}^n$ 是线性变换, $\mathbf{A}$ 是 $\mathcal{A}$ 在基 $\boldsymbol{\alpha} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ 下的矩阵. 若 $\mathbf{v}$ 在基 $\boldsymbol{\alpha}$ 下的坐标是

$$\mathbf{x} = [x_1, \ldots, x_n],$$

则 $\mathcal{A}(\mathbf{v})$ 在基 $\boldsymbol{\alpha}$ 下的坐标是 $\mathbf{A}\mathbf{x}$.

命题 4.10

设 $\mathcal{A}, \mathcal{B}$ 是 $\mathbb{F}^n$ 上线性变换, 且在一组基 $\boldsymbol{\alpha} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ 下的矩阵分别是 $\mathbf{A}$, $\mathbf{B}$. 则复合映射 $\mathcal{A}\mathcal{B}\colon \mathbb{F}^n \longrightarrow \mathbb{F}^n$ 也是线性变换, 且在基 $\boldsymbol{\alpha}$ 下的矩阵为 $\mathbf{A}\mathbf{B}$.

定义 4.11

设 $\mathcal{A}$ 是 $\mathbb{F}^n$ 上的线性变换,

$$\mathrm{Ker}\ \mathcal{A} = \{\mathbf{v} \in \mathbb{F}^n \mid \mathcal{A}(\mathbf{v}) = \mathbf{0}\}, \tag{4.7}$$

$$\mathrm{Im}\ \mathcal{A} = \{\mathbf{w} \in \mathbb{F}^n \mid \exists\ \mathbf{v} \in \mathbb{F}^n \text{ 使得 } \mathcal{A}(\mathbf{v}) = \mathbf{w}\}, \tag{4.8}$$

分别称为 $\mathcal{A}$ 的**核空间**和**像空间**.

例 4.12

设 $\varphi_{\mathbf{A}}$ 是 $\mathbf{A}$ 的线性变换, 则 $\varphi_{\mathbf{A}}$ 的核空间 $\mathrm{Ker}\ \varphi_{\mathbf{A}}$ 就是 $\mathbf{A}$ 的零空间 $\mathrm{Null}\ \mathbf{A}$, $\varphi_{\mathbf{A}}$ 的像空间 $\mathrm{Im}\ \varphi_{\mathbf{A}}$ 就是 $\mathbf{A}$ 的列空间 $\mathrm{Col}\ \mathbf{A}$.

定理 4.13 (像核维数定理)

设 $\mathcal{A}$ 是 $\mathbb{F}^n$ 上的线性变换, 则

$$\dim\ \mathrm{Ker}\ \mathcal{A} + \dim\ \mathrm{Im}\ \mathcal{A} = n.$$

像核维数定理是线性代数课程中的核心定理, 将线性方程组, 矩阵, 线性变换统一的联系起来.

例 4.14

设 $\mathcal{A}$ 是如下定义的 $\mathbb{F}^3$ 上的线性变换

$$\mathcal{A}[x_1, x_2, x_3] = [x_1 + 2x_3, 2x_1 + x_2 + x_3, -x_1 + x_2 - 5x_3].$$

求 $\mathrm{Ker}\ \mathcal{A}$ 和 $\mathrm{Im}\ \mathcal{A}$.

我们已经看到, 一个线性变换在不同的基下的矩阵是不同的. 一个自然的问题是, 若线性变换 $\mathcal{A}$ 在基 $\boldsymbol{\alpha} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ 下的矩阵为 $\mathbf{A}$, 在另一组基 $\boldsymbol{\alpha}' = \{\mathbf{v}_1', \ldots, \mathbf{v}_n'\}$ 下的矩阵为 $\mathbf{A}'$, 则 $\mathbf{A}$ 和 $\mathbf{A}'$ 之间的关联是什么?

不同基之间的联系由如下矩阵 $\mathbf{P} \in \mathsf{M}_n(\mathbb{F})$ 给出:

$$\begin{aligned}\begin{pmatrix}\mathbf{v}_1' & \mathbf{v}_2' & \cdots & \mathbf{v}_n'\end{pmatrix} &= \begin{pmatrix}\mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n\end{pmatrix}\begin{pmatrix}p_{11} & p_{12} & \cdots & p_{1n}\\ p_{21} & p_{22} & \cdots & p_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ p_{n1} & p_{n2} & \cdots & p_{nn}\end{pmatrix}\\ &= \begin{pmatrix}\mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n\end{pmatrix}\cdot\mathbf{P}.\end{aligned} \tag{4.9}$$

$\mathbf{P}$ 称为从 α 到 α' 的过渡矩阵.

命题 4.15

从一组基 $\boldsymbol{\alpha}$ 到另一组基 $\boldsymbol{\alpha}'$ 的过渡矩阵 $\mathbf{P}$ 是可逆矩阵，且 $\mathbf{P}^{-1}$ 是从 $\boldsymbol{\alpha}'$ 到 $\boldsymbol{\alpha}$ 的过渡矩阵.

由 (4.6), $\mathcal{A}$ 在基 $\boldsymbol{\alpha}$ 下的矩阵可以表示为

$$\begin{pmatrix} \mathcal{A}(\mathbf{v}_1) & \cdots & \mathcal{A}(\mathbf{v}_n) \end{pmatrix} = \begin{pmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_n \end{pmatrix} \cdot \mathbf{A}. \tag{4.10}$$

将 (4.9) 代入 (4.10) 得到

$$\begin{pmatrix} \mathcal{A}(\mathbf{v}_1') & \cdots & \mathcal{A}(\mathbf{v}_n') \end{pmatrix} = \begin{pmatrix} \mathbf{v}_1' & \cdots & \mathbf{v}_n' \end{pmatrix} \cdot \mathbf{P}^{-1}\mathbf{A}\mathbf{P}, \tag{4.11}$$

从而 $\mathbf{A}' = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$.

定义 4.16

设 $\mathbf{A}, \mathbf{B} \in \mathsf{M}_n(\mathbb{F})$, 若存在可逆矩阵 $\mathbf{P}$, 使得 $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{B}$, 则称 $\mathbf{A}$ 与 $\mathbf{B}$ 相似. 相似的方阵可以看作同一个线性变换在不同基下的矩阵.

例 4.17

(1) 设 $\mathcal{A}\colon \mathbb{F}^3 \to \mathbb{F}^3$ 在标准基下矩阵为 $\mathbf{A} = \begin{pmatrix} 3 & -20 & 11 \\ 0 & -16 & 10 \\ 5 & -15 & 7 \end{pmatrix}$. 求 $\mathcal{A}$ 在基 $\mathbf{a}_1 = [2,3,5]$, $\mathbf{a}_2 = [0,1,2]$, $\mathbf{a}_3 = [1,0,0]$ 下的矩阵.

(2) 设 $\mathbf{A} \in \mathsf{M}_n(\mathbb{F})$, 且 $\operatorname{rank} \mathbf{A} = r$, $\mathbf{A}^2 = \mathbf{A}$. 证明存在可逆矩阵 $\mathbf{P} \in \mathsf{M}_n(\mathbb{F})$, 使得 $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \begin{pmatrix} \mathbf{I}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$.

4.2

特征值, 特征向量与对角化

设 $\mathcal{A}$ 是 $\mathbb{F}^n$ 上的线性变换, 在标准基下的矩阵是 $\mathbf{A}$. 希望找到一组 "好" 的基 $\boldsymbol{\alpha} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$, 使得 $\mathcal{A}$ 在基 $\boldsymbol{\alpha}$ 下的矩阵尽可能简单. 或者等价的, 找到和 $\mathbf{A}$ 相似的尽可能简单的矩阵.

容易想到的简单的矩阵是对角矩阵, 即除对角线位置外都为 0 的方阵:

$$\mathrm{diag}(\lambda_1, \ldots, \lambda_n) = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}.$$

但并不是每一个方阵都相似于对角矩阵, 例如: $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$.

定义 4.18

设 $\mathbf{A} \in \mathsf{M}_n(\mathbb{F})$, 若存在可逆矩阵 $\mathbf{P} \in \mathsf{M}_n(\mathbb{F})$ 使得

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n), \tag{4.12}$$

就称 $\mathbf{A}$ 在数域 $\mathbb{F}$ 上是可对角化的.

定义 4.19

设 $\mathcal{A}$ 是 $\mathbb{F}^n$ 上线性变换, 若 $\mathcal{A}$ 在一组基 $\boldsymbol{\alpha} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ 下矩阵为对角矩阵 $\mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, 即 $\mathcal{A}$ 在基 $\boldsymbol{\alpha}$ 上的作用为

$$\mathcal{A}(\mathbf{v}_i) = \lambda_i \mathbf{v}_i, \quad 1 \leq i \leq n, \tag{4.13}$$

就称 $\mathcal{A}$ 是可对角化的.

定义 4.20

设 $\mathcal{A}\colon \mathbb{F}^n \longrightarrow \mathbb{F}^n$ 是线性变换, 如果非零向量 $\mathbf{w} \in V$ 满足 $\mathcal{A}(\mathbf{w}) = \lambda\mathbf{w}$, 其中 $\lambda \in \mathbb{F}$, 就称 $\mathbf{w}$ 是 $\mathcal{A}$ 的一个**特征向量**, λ 是 $\mathcal{A}$ 的一个**特征值**.

如何求 $\mathcal{A}\colon \mathbb{F}^n \longrightarrow \mathbb{F}^n$ 的特征向量和特征值? 设 $\mathcal{A}$ 在标准基下矩阵为 $\mathbf{A}$. 若 $\mathbf{w}$ 是 $\mathcal{A}$ 的特征值为 λ 的特征向量, 则 $\mathcal{A}(\mathbf{w}) = \lambda\mathbf{w} \Longleftrightarrow \mathbf{A}\mathbf{w} = \lambda\mathbf{w}$.

定义 4.21

设 $\mathbf{A} \in \mathsf{M}_n(\mathbb{F})$, 若 $\mathbf{x} = [x_1, \ldots, x_n] \in \mathbb{F}^n$ 满足 $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$, 其中 $\lambda \in \mathbb{F}$, 则称 $\mathbf{x}$ 是 $\mathbf{A}$ 的一个**特征向量**, λ 是 $\mathbf{A}$ 的一个**特征值**.

若存在 $\mathbf{x} \neq \mathbf{0}$ 使得 $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$, 则 $(\lambda\mathbf{I}_n - \mathbf{A})\mathbf{x} = \mathbf{0}$, 即 $\mathbf{x}$ 是齐次线性方程组

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

的非零解, 从而有 $\det(\lambda\mathbf{I}_n - \mathbf{A}) = 0$, 即

$$\begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{vmatrix} = 0.$$

对任意 $\mathbf{A} \in \mathrm{M}_n(\mathbb{F})$, 当 λ 看作变量时, $\det(\lambda\mathbf{I}_n - \mathbf{A})$ 是关于 λ 的 $\mathbb{F}$ 上的一元 n 次多项式, 记为 $p_{\mathbf{A}}(\lambda)$, 称为 $\mathbf{A}$ 的**特征多项式**.

命题 4.22

λ_0 是 $\mathbf{A}$ 的特征值当且仅当 λ_0 是 $\mathbf{A}$ 的特征多项式 $p_{\mathbf{A}}(\lambda)$ 的根.

为了保证特征多项式 $p_{\mathbf{A}}(\lambda)$ 的根都在 $\mathbb{F}$ 中, 根据代数基本定理, 以下**设数域 $\mathbb{F}$ 为复数域 $\mathbb{C}$**. $\mathbf{A} \in \mathrm{M}_n(\mathbb{C})$ 的特征值和特征向量的计算方法如下:

1. 计算特征多项式 $p_{\mathbf{A}}(\lambda) = |\lambda \mathbf{I}_n - \mathbf{A}|$, 设

$$p_{\mathbf{A}}(\lambda) = (\lambda - \lambda_1)^{n_1} \cdots (\lambda - \lambda_s)^{n_s},$$

   则 $\lambda_1, \ldots, \lambda_s$ 为 $\mathbf{A}$ 的全部特征值.
2. 对每个 λ_i, 求解线性方程组 $(\lambda_i \mathbf{I}_n - \mathbf{A})\mathbf{x} = \mathbf{0}$, 其解空间 V_{λ_i} 中任意向量就是 $\mathbf{A}$ 关于 λ_i 的特征向量.

例 4.23

求 $\mathbf{A}=\begin{pmatrix}3 & -1 & -2\\ 2 & 0 & -2\\ 2 & -1 & -1\end{pmatrix}$ 的所有特征值和特征向量.

虽然方阵不一定可对角化, 但可以证明 $\mathrm{M}_n(\mathbb{C})$ 中的方阵都可以上三角化.

定义 4.24

设 $\mathbf{U}=(u_{ij})_{n\times n}\in \mathrm{M}_n(\mathbb{C})$ 满足对任意 $i>j$ 都有 $u_{ij}=0$, 即 $\mathbf{U}$ 是形如 $\begin{pmatrix}u_{11} & u_{12} & \cdots & u_{1n}\\ & u_{22} & \cdots & u_{2n}\\ & & \ddots & \vdots\\ & & & u_{nn}\end{pmatrix}$ 的方阵, 称 $\mathbf{U}$ 是上三角矩阵.

我们已经看到上三角矩阵在解线性方程组, 计算矩阵的秩, 计算行列式中发挥了关键作用.

引理 4.25

设 $\mathbf{A}, \mathbf{B} \in \mathsf{M}_n(\mathbb{C})$ 是相似的方阵, 则 $p_{\mathbf{A}}(\lambda) = p_{\mathbf{B}}(\lambda)$. 特别的, 相似矩阵有相同的特征值.

定理 4.26

任何 $\mathbf{A} \in \mathsf{M}_n(\mathbb{C})$ 都相似于一个上三角矩阵, 且上三角矩阵中主对角线上的元素就是 $\mathbf{A}$ 的全部特征值.

推论 4.27

设 $\mathbf{A} \in \mathsf{M}_n(\mathbb{C})$, $\lambda_1, \ldots, \lambda_n$ 是 $\mathbf{A}$ 的全部特征值, 则 $\det \mathbf{A} = \lambda_1 \cdots \lambda_n$. 从而 $\mathbf{A}$ 可逆 $\Longleftrightarrow$ 所有特征值 $\lambda_i \neq 0$.

接下来研究 $\mathbf{A} \in \mathsf{M}_n(\mathbb{C})$ 可对角化的充要条件. 首先按照定义, $\mathbf{A} \in \mathsf{M}_n(\mathbb{C})$ 可对角化 $\iff$ $\mathbf{A}$ 有 n 个线性无关的特征向量.

例 4.28

证明 $\mathbf{A} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$ 不可对角化.

定义 4.29

设 $\mathbf{A} \in \mathsf{M}_n(\mathbb{C})$, $p_{\mathbf{A}}(\lambda) = (\lambda - \lambda_1)^{n_1} \cdots (\lambda - \lambda_s)^{n_s}$. 对每个 i,

(1) n_i 称为 λ_i 的代数重数.

(2) 线性方程组 $(\lambda_i \mathbf{I}_n - \mathbf{A})\mathbf{x} = \mathbf{0}$ 的解空间 V_{λ_i} 称为 $\mathbf{A}$ 关于 λ_i 的特征子空间, 维数 $\dim V_{\lambda_i}$ 称为 λ_i 的几何重数.

命题 4.30

设 λ 是 $\mathbf{A}$ 的特征值, 则它的几何重数不超过它的代数重数.

命题 4.31

设 $\mathbf{v}_1, \ldots, \mathbf{v}_s$ 是 $\mathbf{A}$ 的特征值分别为 $\lambda_1, \ldots, \lambda_s$ 的特征向量, 且 $\lambda_1, \ldots, \lambda_s$ 互不相同, 则 $\mathbf{v}_1, \ldots, \mathbf{v}_s$ 线性无关.

定理 4.32

$\mathbf{A} \in \mathrm{M}_n(\mathbb{C})$ 可对角化 $\iff$ $\mathbf{A}$ 的每个特征值的代数重数与几何重数相同.

若 $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. 设 $\mathbf{P} = \begin{pmatrix} \mathbf{p}_1 & \cdots & \mathbf{p}_n \end{pmatrix}$, 则

$$\mathbf{A}\begin{pmatrix} \mathbf{p}_1 & \cdots & \mathbf{p}_n \end{pmatrix} = \begin{pmatrix} \lambda_1\mathbf{p}_1 & \cdots & \lambda_n\mathbf{p}_n \end{pmatrix}.$$

例 4.33

$\mathbf{A} = \begin{pmatrix} 3 & -1 & -2 \\ 2 & 0 & -2 \\ 2 & -1 & -1 \end{pmatrix}$ 是否可对角化? 若可以, 求可逆矩阵 $\mathbf{P}$ 使得 $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ 是对角矩阵.

例 4.34

平面 $\mathbb{R}^2$ 上的伸缩 $\mathcal{S}_{\lambda,\mu}$, 旋转 $\mathcal{R}_\theta$ 以及反射 r_θ 在实数域 $\mathbb{R}$ 上是否可对角化?

例 4.35

设 $\mathbf{A} = \begin{pmatrix} 0 & 0 & 1 \\ x & 1 & y \\ 1 & 0 & 0 \end{pmatrix}$ 有三个线性无关的特征向量, 求 x, y.

5 欧几里得空间

内积

正交变换与对称变换

二次型

5.1

内积

基于向量的加法及数乘运算, 我们研究了 $\mathbb{F}^n$ 中的向量组的线性相关性, 子空间及线性映射. 为了研究 $\mathbb{F}^n$ 中的几何, 还需要引入**向量的长度**以及**向量间的角度**. 在 $\mathbb{R}^3$ 中, 将这两者统一起来的概念是**内积**:

$$\mathbf{a}\cdot\mathbf{b}=|\mathbf{a}||\mathbf{b}|\cos\theta.$$

本章将固定 $\mathbb{F}=\mathbb{R}$, 研究 $\mathbb{R}$ 上 n 维线性空间 (实空间) 的几何.

在 $\mathbb{R}^3$ 中选取直角坐标系 $[O;\mathbf{i},\mathbf{j},\mathbf{k}]$, 内积在坐标下的公式为

$$[a_1,a_2,a_3]\cdot[b_1,b_2,b_3]=a_1b_1+a_2b_2+a_3b_3.$$

在 $\mathbb{R}^n$ 中, 有标准基 $\boldsymbol{\epsilon}_1,\ldots,\boldsymbol{\epsilon}_n$ 及在这组基下的坐标表示, 可以先定义内积, 然后定义向量的长度及向量间的夹角.

定义 5.1

在 $\mathbb{R}^n$ 中选取标准基 $\boldsymbol{\epsilon}_1, \ldots, \boldsymbol{\epsilon}_n$, 并在坐标下定义

$$\mathbf{a} \cdot \mathbf{b} = [a_1, \ldots, a_n] \cdot [b_1, \ldots, b_n] = a_1 b_1 + \cdots + a_n b_n,$$

称为 $\mathbb{R}^n$ 的标准内积.

在标准内积下, 定义向量 $\mathbf{a}$ 的长度为

$$|\mathbf{a}| = \sqrt{\mathbf{a} \cdot \mathbf{a}} = \sqrt{a_1^2 + \cdots + a_n^2},$$

定义向量 $\mathbf{a}$, $\mathbf{b}$ 的夹角的余弦为

$$\cos\theta = \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}||\mathbf{b}|} = \frac{a_1 b_1 + \cdots + a_n b_n}{\sqrt{a_1^2 + \cdots + a_n^2}\sqrt{b_1^2 + \cdots + b_n^2}}.$$

标准内积的公式依赖于基的选取，需要研究内积在不同基下的公式之间的联系，并给出内积不依赖于基的选取的定义．和定义行列式一样，我们已经从标准内积那里收集了内积应当满足的关键性质，可以定义内积为满足这些性质的运算．

定义 5.2

设在 $\mathbb{R}^n$ 上定义了一种运算，使得任何两个向量 $\mathbf{a}$, $\mathbf{b}$ 都对应于一个实数，记为 $(\mathbf{a},\mathbf{b})$，并且满足：

1. **对称性：** $(\mathbf{a},\mathbf{b})=(\mathbf{b},\mathbf{a})$.
2. **线性：** $(\lambda\mathbf{a}_1+\mu\mathbf{a}_2,\mathbf{b})=\lambda(\mathbf{a}_1,\mathbf{b})+\mu(\mathbf{a}_2,\mathbf{b})$.
3. **正定性：** $(\mathbf{a},\mathbf{a})\geq 0$, 等号成立 $\iff$ $\mathbf{a}=\mathbf{0}$.

则称 $(\mathbf{a},\mathbf{b})$ 是 $\mathbf{a}$, $\mathbf{b}$ 的**内积**．$\mathbb{R}^n$ 在内积 $(\cdot,\cdot)$ 下成为一个**欧几里得空间**，简称**欧式空间**．

设 $(\cdot,\cdot)$ 是 $\mathbb{R}^n$ 上的内积, 则由正定性可以定义 $\mathbf{a}$ 的长度为 $|\mathbf{a}|=\sqrt{(\mathbf{a},\mathbf{a})}$. 若 $|\mathbf{a}|=1$, 称 $\mathbf{a}$ 是单位向量. 另一方面,

引理 5.3 (柯西-施瓦茨不等式)

$|(\mathbf{a},\mathbf{b})|^2 \le (\mathbf{a},\mathbf{a})(\mathbf{b},\mathbf{b})$, 等号成立 $\iff$ $\mathbf{a},\mathbf{b}$ 线性相关.

从而可以定义 $\mathbf{a},\mathbf{b}$ 的夹角的余弦为 $\cos\theta=\dfrac{(\mathbf{a},\mathbf{b})}{\sqrt{(\mathbf{a},\mathbf{a})(\mathbf{b},\mathbf{b})}}$. 若 $(\mathbf{a},\mathbf{b})=0$, 称 $\mathbf{a},\mathbf{b}$ 正交.

推论 5.4 (三角不等式)

在欧式空间 $\mathbb{R}^n$ 中, $|\mathbf{a}+\mathbf{b}|\le|\mathbf{a}|+|\mathbf{b}|$.

下面给出内积 $(\cdot,\cdot)$ 在一组基 $\boldsymbol{\alpha}=\{\mathbf{v}_1,\ldots,\mathbf{v}_n\}$ 下的坐标公式:

设 $\mathbf{a}$, $\mathbf{b}$ 在基 $\boldsymbol{\alpha}$ 下的坐标为 $\mathbf{x} = [x_1, \ldots, x_n]$, $\mathbf{y} = [y_1, \ldots, y_n]$. 则由内积的线性可知

$$(\mathbf{a}, \mathbf{b}) = (\sum_{i=1}^{n} x_i \mathbf{v}_i, \sum_{j=1}^{n} y_j \mathbf{v}_j) = \sum_{1 \leq i,j \leq n} x_i y_j (\mathbf{v}_i, \mathbf{v}_j).$$

矩阵 $\mathbf{G} = ((\mathbf{v}_i, \mathbf{v}_j))_{n \times n}$ 称为 $(\cdot, \cdot)$ 在基 $\mathbf{v}_1, \ldots, \mathbf{v}_n$ 下的**度量矩阵**, 并且 $(\cdot, \cdot)$ 在基 $\mathbf{v}_1, \ldots, \mathbf{v}_n$ 下的坐标公式为

$$(\mathbf{a}, \mathbf{b}) = \mathbf{x}^{\mathrm{T}} \mathbf{G} \mathbf{y} = \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} \begin{pmatrix} (\mathbf{v}_1, \mathbf{v}_1) & \cdots & (\mathbf{v}_1, \mathbf{v}_n) \\ \vdots & \ddots & \vdots \\ (\mathbf{v}_n, \mathbf{v}_1) & \cdots & (\mathbf{v}_n, \mathbf{v}_n) \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

若度量矩阵 $\mathbf{G}$ 为单位阵 $\mathbf{I}_n$, 就得到标准内积公式.

定义 5.5

若欧式空间 $\mathbb{R}^n$ 中的一组基 $\boldsymbol{\eta} = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ 满足

$$(\mathbf{u}_i, \mathbf{u}_j) = \delta_{ij}, \quad \forall\ 1 \leq i \leq j \leq n$$

即内积 $(\cdot, \cdot)$ 在基 $\boldsymbol{\eta}$ 下的度量矩阵为 $\mathbf{I}_n$, 则 $\boldsymbol{\eta}$ 称为 $\mathbb{R}^n$ 的一组标准正交基.

命题 5.6

在欧氏空间 $\mathbb{R}^n$ 中, 设 $\mathbf{a}_1, \ldots, \mathbf{a}_r$ 为非零向量且两两正交, 则它们线性无关.

定理 5.7 (格拉姆-施密特正交化)

从欧式空间 $\mathbb{R}^n$ 的任意一组基 $\boldsymbol{\alpha} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ 出发, 可以构造一组标准正交基 $\boldsymbol{\eta} = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$. 特别的, 欧氏空间中存在标准正交基.

1. 取 $\mathbf{u}_1 = \mathbf{v}_1/|\mathbf{v}_1|$.
2. 取 $\mathbf{w}_2 = \mathbf{v}_2 - (\mathbf{v}_2, \mathbf{u}_1)\mathbf{u}_1$, 并令 $\mathbf{u}_2 = \mathbf{w}_2/|\mathbf{w}_2|$.
3. 假设已得到 $\mathbf{u}_1, \ldots, \mathbf{u}_k$, 考虑

$$\mathbf{w}_{k+1} = \mathbf{v}_{k+1} - (\mathbf{v}_{k+1}, \mathbf{u}_1)\mathbf{u}_1 - \cdots - (\mathbf{v}_{k+1}, \mathbf{u}_k)\mathbf{u}_k,$$

   则 $\mathbf{w}_{k+1} \neq \mathbf{0}$, 并令 $\mathbf{u}_{k+1} = \mathbf{w}_{k+1}/|\mathbf{w}_{k+1}|$.

例 5.8

从 $\mathbb{R}^3$ 的一组基 $\mathbf{v}_1=(1,-1,0)$, $\mathbf{v}_2=(0,1,-1)$, $\mathbf{v}_3=(1,0,1)$ 出发构造 $\mathbb{R}^3$ 在标准内积下的一组标准正交基.

推论 5.9

(1) 设欧式空间 V 中的内积在一组基 $\boldsymbol{\alpha}$ 下的度量矩阵为 $\mathbf{G}$, 则 $\mathbf{G}=\mathbf{P}^{\mathrm{T}}\mathbf{P}$, 其中 $\mathbf{P}$ 是可逆矩阵.

(2) 欧式空间 $\mathbb{R}^n$ 中的一组两两正交的单位向量 $\mathbf{u}_1,\dots,\mathbf{u}_k$ 可以扩充为 $\mathbb{R}^n$ 的一组标准正交基.

(3) 设 $\mathbf{u}_1,\dots,\mathbf{u}_n$ 是一组标准正交基, 则对 $\forall\ \mathbf{v}\in V$,

$$\mathbf{v}=(\mathbf{v},\mathbf{u}_1)\mathbf{u}_1+\cdots+(\mathbf{v},\mathbf{u}_n)\mathbf{u}_n.$$

5.2

正交变换与对称变换

欧式空间就是带有内积 $(\cdot,\cdot)$ 的向量空间 $\mathbb{R}^n$, 本节将介绍两类和内积相关的 $\mathbb{R}^n$ 上的线性变换.

定义 5.10

设 $\mathcal{A}$ 是欧式空间 $\mathbb{R}^n$ 上的线性变换, 且满足

$$(\mathcal{A}(\mathbf{v}),\mathcal{A}(\mathbf{w}))=(\mathbf{v},\mathbf{w}),\quad \forall\ \mathbf{v},\mathbf{w}\in\mathbb{R}^n,$$

则 $\mathcal{A}$ 称为正交变换. 正交变换就是保持内积的线性变换.

命题 5.11

设 $\mathcal{A}$ 是欧式空间 $\mathbb{R}^n$ 上的正交变换, 则 $|\mathcal{A}(\mathbf{v})|=|\mathbf{v}|$, 对任意 $\mathbf{v}\in\mathbb{R}^n$ 成立. 即正交变换保持向量长度, 因此正交变换也称为保距变换.

命题 5.12

设 $\mathcal{A}$ 是欧式空间 $\mathbb{R}^n$ 上的正交变换, $\mathbf{u}_1, \ldots, \mathbf{u}_n$ 是 $\mathbb{R}^n$ 的一组标准正交基, 则 $\mathcal{A}(\mathbf{u}_1), \ldots, \mathcal{A}(\mathbf{u}_n)$ 也是 $\mathbb{R}^n$ 的一组标准正交基.

命题 5.13

设 $\boldsymbol{\eta} = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ 是 $\mathbb{R}^n$ 的一组标准正交基, 正交变换 $\mathcal{A}$ 在基 $\boldsymbol{\eta}$ 下的矩阵为 $\mathbf{A} = (a_{ij})_{n\times n}$, 即 $\mathcal{A}(\mathbf{u}_j) = \sum_{i=1}^n a_{ij}\mathbf{u}_i$, 或者

$$\begin{pmatrix} \mathcal{A}(\mathbf{u}_1) & \cdots & \mathcal{A}(\mathbf{u}_n) \end{pmatrix} = \begin{pmatrix} \mathbf{u}_1 & \cdots & \mathbf{u}_n \end{pmatrix} \mathbf{A},$$

则 $\mathbf{A}$ 满足 $\mathbf{A}^{\mathrm{T}}\mathbf{A} = \mathbf{I}_n$.

定义 5.14

满足 $\mathbf{A}^{\mathrm{T}}\mathbf{A} = \mathbf{I}_n$ 的 $\mathrm{M}_n(\mathbb{R})$ 中的方阵称为**正交方阵**.

定理 5.15

设 $\mathcal{A}$ 是欧式空间 $\mathbb{R}^n$ 上的线性变换, 则以下四者等价:

(1) $\mathcal{A}$ 是正交变换.

(2) $(\mathcal{A}(\mathbf{v}), \mathcal{A}(\mathbf{v})) = (\mathbf{v}, \mathbf{v})$ 对任意 $\mathbf{v} \in \mathbb{R}^n$ 成立.

(3) $\mathcal{A}$ 将一组标准正交基映为一组标准正交基.

(4) $\mathcal{A}$ 在一组标准正交基下的矩阵为正交方阵.

例 5.16

证明二阶正交方阵 $\mathbf{A} \in \mathsf{M}_2(\mathbb{R})$ 只能为 $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ 或 $\begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}$. 即平面上的正交变换只有旋转和反射.

命题 5.17

设 $\mathbf{A} \in \mathsf{M}_n(\mathbb{R})$ 是正交方阵, 则

(1) $\det \mathbf{A} = \pm 1$.

(2) 若 $\lambda \in \mathbb{C}$ 是 $\mathbf{A}$ 的特征值, 则 $|\lambda| = 1$.

例 5.18

设 $\mathcal{A}\colon \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ 是正交变换, 且在标准基 $\{\boldsymbol{\epsilon}_1, \boldsymbol{\epsilon}_2, \boldsymbol{\epsilon}_3\}$ 下的矩阵 $\mathbf{A}$ 满足 $\det \mathbf{A} = 1$. 则存在一组标准正交基 $\boldsymbol{\eta}$, $\mathcal{A}$ 在 $\boldsymbol{\eta}$ 下的矩阵为

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix},$$

即 $\mathcal{A}$ 是空间中以某条过原点直线为轴的旋转变换.

定义 5.19

设 $\mathcal{A}$ 是欧式空间 $\mathbb{R}^n$ 上的线性变换, 满足

$$(\mathcal{A}(\mathbf{v}), \mathbf{w}) = (\mathbf{v}, \mathcal{A}(\mathbf{w})), \quad \forall\ \mathbf{v}, \mathbf{w} \in \mathbb{R}^n,$$

则 $\mathcal{A}$ 称为对称变换.

定理 5.20

设 $\mathcal{A}$ 是欧式空间 $\mathbb{R}^n$ 上的线性变换, 则 $\mathcal{A}$ 是对称变换 $\Longleftrightarrow$ $\mathcal{A}$ 在一组标准正交基下的矩阵是实对称方阵.

命题 5.21

设 $\mathcal{A}$ 是欧式空间 $\mathbb{R}^n$ 上的对称变换, 则 $\mathcal{A}$ 的属于不同特征值的特征向量彼此正交.

命题 5.22

实对称方阵的特征值都是实数.

定理 5.23

设 $\mathbf{A} \in \mathrm{M}_n(\mathbb{R})$ 是实对称方阵, 则存在正交方阵 $\mathbf{O}$, 使得 $\mathbf{O}^{-1}\mathbf{A}\mathbf{O}$ 是对角矩阵 $\mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. 特别的, 实对称方阵可对角化.

求将对称阵 $\mathbf{A}$ 化为对角阵的正交方阵 $\mathbf{O}$ $\Longleftrightarrow$ 求一组 $\mathbf{A}$ 的特征向量, 它们构成 $\mathbb{R}^n$ 的一组标准正交基.

例 5.24

设 $\mathbf{A} = \begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{pmatrix}$, 求正交方阵 $\mathbf{O}$ 使得 $\mathbf{O}^{-1}\mathbf{A}\mathbf{O}$ 是对角矩阵.

5.3

二次型

定义 5.25

一个 $\mathbb{R}$ 上含有 n 个变元的二次齐次多项式

$$Q(x_1,\ldots,x_n)=\sum_{1\le i\le j\le n}\bar{a}_{ij}x_ix_j,\quad \bar{a}_{ij}\in\mathbb{R}$$

称为一个 n 阶二次型.

将 $Q(x_1,\ldots,x_n)$ 看作关于列向量 $\mathbf{x}=[x_1,\ldots,x_n]$ 的函数 $Q(\mathbf{x})$, 则二次型唯一对应一个对称方阵 $\mathbf{A}=(a_{ij})_{n\times n}$, 其中

$$a_{ii}=\bar{a}_{ii},\quad 1\le i\le n,$$

$$a_{ij}=a_{ji}=\bar{a}_{ij}/2,\quad 1\le i<j\le n.$$

并且 $Q(\mathbf{x})=\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}$, $\mathbf{A}$ 称为二次型的矩阵.

例 5.26

写出二次型 $x_1x_2+x_2x_3+x_1x_3$, 以及 $x_1^2+2x_2^2+3x_3^2$ 的矩阵.

例 5.27

设 $(\cdot,\cdot)$ 是 $\mathbb{R}^n$ 上的内积, $\mathbf{G}$ 是内积在一组基 $\boldsymbol{\alpha}$ 下的度量矩阵. 对任意 $\mathbf{x}\in\mathbb{R}^n$, 定义 $Q(\mathbf{x})=(\mathbf{x},\mathbf{x})=\mathbf{x}^{\mathrm{T}}\mathbf{G}\mathbf{x}$, 是 n 阶二次型. 由内积的正定性, $Q(\mathbf{x})>0$ 对任意 $\mathbf{x}\neq\mathbf{0}$ 成立.

定义 5.28

设 $Q(\mathbf{x})$ 是 n 阶二次型.

若 $Q(\mathbf{x})>0$ $(Q(\mathbf{x})\geq 0)$ 对 $\forall\ \mathbf{x}\neq\mathbf{0}$ 成立, 则称 $Q(\mathbf{x})$ 是**正定的** (**半正定的**); 若 $Q(\mathbf{x})<0$ $(Q(\mathbf{x})\leq 0)$ 对 $\forall\ \mathbf{x}\neq\mathbf{0}$ 成立, 则称 $Q(\mathbf{x})$ 是**负定的** (**半负定的**); 否则称 $Q(\mathbf{x})$ 是**不定的**.

设 $Q(\mathbf{x}) = \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}$ 是二次型. 因为 $\mathbf{A}$ 是对称方阵, 由定理 5.23, 存在正交方阵 $\mathbf{O}$ 使得 $\mathbf{O}^{\mathrm{T}}\mathbf{A}\mathbf{O} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, 其中 λ_i 都是实数, 是 $\mathbf{A}$ 的全部特征值. 对任意 $\mathbf{x} \in \mathbb{R}^n$, 设 $\mathbf{y} = \mathbf{O}^{\mathrm{T}}\mathbf{x}$, 则

$$Q(\mathbf{x}) = \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} = \mathbf{y}^{\mathrm{T}}(\mathbf{O}^{\mathrm{T}}\mathbf{A}\mathbf{O})\mathbf{y} = \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2.$$

因为 $\mathbf{O}^{\mathrm{T}}$ 是可逆矩阵, 故 $\mathbf{x} \to \mathbf{O}^{\mathrm{T}}\mathbf{x} = \mathbf{y}$ 诱导了从 $\mathbb{R}^n$ 到自身的一一映射. 从而 $Q(\mathbf{x})$ 的值域和二次型

$$\tilde{Q}(\mathbf{y}) = \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2$$

的值域相同. 这说明每个二次型都可以通过一组坐标变换化为简单的平方和形式, $\tilde{Q}(\mathbf{y})$ 称为 $Q(\mathbf{x})$ 的**标准形**.

例 5.29

求正交方阵 $\mathbf{O}$ 使得二次型

$$Q(\mathbf{x}) = x_1^2 + 4x_2^2 + 4x_3^2 - 4x_1x_2 + 4x_1x_3 - 8x_2x_3$$

化为标准形, 并写出它的标准形.

命题 5.30

设二次型 $Q(\mathbf{x})$ 的标准形为 $\lambda_1 y_1^2 + \cdots + \lambda_n y_n^2$, 则

(1) $Q(\mathbf{x})$ 是正定的 (半正定的)$\Longleftrightarrow$ 所有 $\lambda_i > 0$ $(\lambda_i \geq 0)$.

(2) $Q(\mathbf{x})$ 是负定的 (半负定的)$\Longleftrightarrow$ 所有 $\lambda_i < 0$ $(\lambda_i \leq 0)$.

(3) $Q(\mathbf{x})$ 是不定的 $\Longleftrightarrow$ 既存在 $\lambda_i > 0$, 又存在 $\lambda_j < 0$.

命题 5.31

设 $\mathbf{A}$ 是对称方阵, 则存在可逆矩阵 $\mathbf{P}$, 使得

$$\mathbf{P}^{\mathrm{T}}\mathbf{A}\mathbf{P} = \begin{pmatrix} \mathbf{I}_r & & \\ & -\mathbf{I}_s & \\ & & \mathbf{0} \end{pmatrix},$$

等式右边的对角阵称为 $\mathbf{A}$ 的规范形. 对应的, 实二次型 $Q(\mathbf{x})$ 可以经过坐标变换化为如下规范形:

$$y_1^2 + \cdots + y_r^2 - y_{r+1}^2 - \cdots - y_{r+s}^2.$$

定理 5.32

实对称矩阵 $\mathbf{A}$ 的规范形唯一.

对称方阵 $\mathbf{A}$ 称为是正定的 (负定的, 不定的), 如果其定义的二次型 $\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}$ 是正定的 (负定的, 不定的).

推论 5.33

(1) 二次型 $Q(\mathbf{x})$ 的规范形唯一. r 和 s 分别称为 $Q(\mathbf{x})$ 的正惯性指数和负惯性指数.

(2) $\mathbf{A}$ 是正定的 $\Longleftrightarrow$ 存在可逆矩阵 $\mathbf{P}$ 使得 $\mathbf{A}=\mathbf{P}^{\mathrm{T}}\mathbf{P}$.

定义 5.34

设 $\mathbf{A}=(a_{ij})_{n\times n}$ 是实对称方阵, $\begin{vmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{k1} & \cdots & a_{kk} \end{vmatrix}$ 称为 $\mathbf{A}$ 的 k 阶顺序主子式, 其中 $k=1,\ldots,n$.

下面给出判别矩阵正定 (负定, 不定) 的实用方法:

定理 5.35

(1) 实对称方阵 $\mathbf{A}$ 是正定的 $\iff$ $\mathbf{A}$ 的各阶顺序主子式 > 0.

(2) 实对称方阵 $\mathbf{A}$ 是负定的 $\iff$ $\mathbf{A}$ 的奇数阶顺序主子式 < 0, 偶数阶顺序主子式 > 0.

例 5.36

求如下二次型的规范形:

(1) $Q(\mathbf{x}) = x_1^2 + 2x_2^2 + 6x_3^2 + 2x_1x_2 + 2x_1x_3 + 6x_2x_3$.

(2) $Q(\mathbf{x}) = x_1^2 - x_2^2 + x_3^2 + 4x_1x_3 + 2x_2x_3$.

谢谢大家!